EECS 16A Designing Information Devices and Systems I
Spring 2018 Discussion 11A

Reference Definitions

Inner products: An inner product is a function that associates each pair of two vectors in a vector space V with a real number (called the inner product). For any $\vec{x},\vec{y},\vec{z} \in V$ and $c \in \mathbb{R}$, the inner product satisfies the following three properties:

(a) **Symmetry:** $\langle \vec{x},\vec{y}\rangle = \langle \vec{y},\vec{x}\rangle$

(b) **Linearity:**

i. $\langle \vec{x},\vec{y}+\vec{z}\rangle = \langle \vec{x},\vec{y}\rangle + \langle \vec{x},\vec{z}\rangle$

ii. $\langle c\vec{x},\vec{y}\rangle = c\langle \vec{x},\vec{y}\rangle$

(c) **Positive-definiteness:** $\langle \vec{x},\vec{x}\rangle \geq 0$ with $\langle \vec{x},\vec{x}\rangle = 0$ if and only if $\vec{x} = \vec{0}$

Norm: The norm of a vector $\vec{x} \in V$ is defined to be:

$$\|\vec{x}\|^2 = \langle \vec{x},\vec{x}\rangle \implies \|\vec{x}\| = \sqrt{\langle \vec{x},\vec{x}\rangle}$$

1. From Inner Products To Projections

Given that $\langle \vec{x},\vec{y}\rangle$ is a measure of similarity between two vectors, let's try to use this to find how much of one vector $\vec{y}$ is in the direction of another vector $\vec{x}$.

(a) Let's start with $\langle \vec{x},\vec{y}\rangle$. We want a quantity that is independent of the norm of $\vec{x}$, $\|\vec{x}\|$. Is $\langle \vec{x},\vec{y}\rangle$ independent of the norm? Consider $\langle \vec{x},\vec{y}\rangle$ for the examples below.

$\vec{x} = \begin{bmatrix}1\\1\end{bmatrix}, \vec{y} = \begin{bmatrix}2\\1\end{bmatrix}, \vec{x} = \begin{bmatrix}2\\2\end{bmatrix}, \vec{y} = \begin{bmatrix}2\\1\end{bmatrix}$

Answer:

For the first example above, we find that $\langle \vec{x},\vec{y}\rangle = 2+1 = 3$.

For the second example, we find that $\langle \vec{x},\vec{y}\rangle = 4+2 = 6$.

Notice that in both cases, the direction of the vector $\vec{x}$ did not change. However, the value of $\langle \vec{x},\vec{y}\rangle$ was dependent on the norm of $\|\vec{x}\|$, not just the direction.

(b) Suppose we divide $\langle \vec{x},\vec{y}\rangle$ by the norm of $\vec{x}$, $\|\vec{x}\|$, to get $\frac{\langle \vec{x},\vec{y}\rangle}{\|\vec{x}\|}$. Is this new quantity independent of the norm of $\vec{x}$? Test it on the examples above.

Answer:

For the first case: $\|\vec{x}\| = \sqrt{1+1} = \sqrt{2}$, so $\frac{\langle \vec{x},\vec{y}\rangle}{\|\vec{x}\|} = \frac{3}{\sqrt{2}}$.

For the second case: $\|\vec{x}\| = \sqrt{4+4} = 2\sqrt{2}$, so $\frac{\langle \vec{x},\vec{y}\rangle}{\|\vec{x}\|} = \frac{6}{2\sqrt{2}} = \frac{3}{\sqrt{2}}$.

Notice here that we have removed the dependence on the norm of $\vec{x}$.

(c) We now have a scalar quantity that represents how much of $\vec{y}$ is in the direction of $\vec{x}$. Let's try to find a vector that is how much of $\vec{y}$ is in the $\vec{x}$ direction. That is, we are looking for a vector $\vec{z}$ that has a norm of $\frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|}$ and points in the same direction as $\vec{x}$.

Answer:

If we multiply our answer from the last part by the vector $\vec{x}$, we get a quanity that point in the correct direction. However, its norm is $\frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|} \|\vec{x}\| = \langle \vec{x}, \vec{y} \rangle$. In order to get the expected norm, we need to scale our vector $\vec{x}$ by $\frac{1}{\|\vec{x}\|}$.

Thus, we find $\vec{z} = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|^2} \vec{x}$. This is also known as the projection. We will use it a lot in the upcoming dicussions on least squares.

(d) Given the projection between two vectors, defined as $\text{proj}_{\vec{x}} \vec{y} = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|^2} \vec{x}$, prove the Cauchy-Schwarz inequality, $|\langle \vec{x}, \vec{y} \rangle| \le \|\vec{x}\| \|\vec{y}\|$.

Answer:

Let the vector $\vec{p}$ be the projection of $\vec{y}$ on the vector $\vec{x}$, that is, $\vec{p} = \text{proj}_{\vec{x}} \vec{y} = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|^2} \vec{x}$.

We know that the magnitude of the projection of $\vec{y}$ onto $\vec{x}$ must be less than or equal to the magnitude of $\vec{y}$.

$$\|\vec{y}\| \ge \|\vec{p}\|$$
$$\|\vec{y}\| \ge \left| \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|^2} \right| \|\vec{x}\|$$
$$\|\vec{y}\| \ge \frac{|\langle \vec{x}, \vec{y} \rangle|}{\|\vec{x}\|}$$
$$\|\vec{y}\| \|\vec{x}\| \ge |\langle \vec{x}, \vec{y} \rangle|$$

(e) Consider the quantity $\frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|}$. What is the maximum this quantity could be? When does this occur? What is the minimum this quantity could be? When does this occur?

Answer:

Using the Cauchy-Schwarz inequality, we know that $|\langle \vec{x}, \vec{y} \rangle| \le \|\vec{x}\| \|\vec{y}\|$.

$$\frac{|\langle \vec{x}, \vec{y} \rangle|}{\|\vec{x}\| \|\vec{y}\|} \le 1 \implies -1 \le \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|} \le 1$$

The maximum occurs when $\vec{x}$ and $\vec{y}$ are parallel. The quantity becomes $\frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|} = 1$.

The minimum occurs when $\vec{x}$ and $\vec{y}$ are antiparallel. The quantity becomes $\frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|} = -1$.

(f) We define the angle between two vectors as $\cos(\theta) = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|}$. When do two vectors have an angle of 90° between them? When do they have an angle of 0°? When do they have an angle of 180°?

Answer:

Vectors have an angle of 90° when their inner product is zero. The have an angle of 0° when they point in the same direction and an angle of 180° when they point in opposite directions.

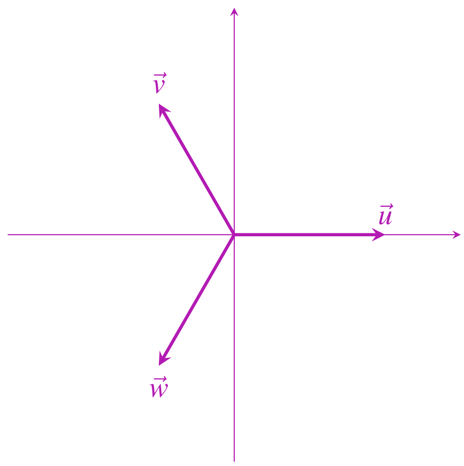
2. Packings

(a) Can three vectors in the $\mathbb{R}^2$ plane have only negative pairwise inner-products? That is, do there exist vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^2$ such that $\langle \vec{u}, \vec{v} \rangle < 0$, $\langle \vec{v}, \vec{w} \rangle < 0$, and $\langle \vec{u}, \vec{w} \rangle < 0$?

Hint: Draw a picture!

Answer:

Yes, this is possible. Geometrically, we want vectors in $\mathbb{R}^2$ such that the pairwise angles between them are all greater than $90°$. Take for example three vectors oriented symmetrically, with pairwise angles of $120°$.



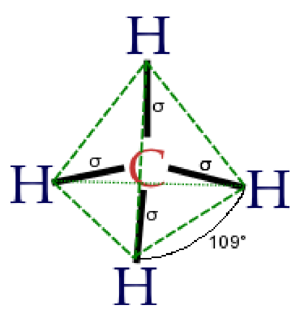
(b) What about four vectors in $\mathbb{R}^2$? That is, do there exist four vectors $\vec{u}, \vec{v}, \vec{w}, \vec{x} \in \mathbb{R}^2$ such that for every pair of vectors $\vec{a}, \vec{b}$: $\langle \vec{a}, \vec{b} \rangle < 0$?

Bonus: What about four vectors in $\mathbb{R}^3$?

Answer:

This is not possible in $\mathbb{R}^2$. Say we have four vectors satisfying the above. Order them by the angle they make with (say) the x-axis. By the same geometric reasoning above, the angles between consecutive vectors should be greater than $90°$, but this is impossible – there are only 360 degrees in total.

However, this is possible in $\mathbb{R}^3$, for example the structure of a methane molecule:



3. Orthogonal Subspaces

Two vectors are $\vec{x}$ and $\vec{y}$ are said to be orthogonal if their inner product is zero. That is $\langle \vec{x}, \vec{y} \rangle = 0$.

Two subspaces $\mathbb{S}_1$ and $\mathbb{S}_2$ of $\mathbb{R}^N$ are said to be orthogonal if all vectors in $\mathbb{S}_1$ are orthogonal to all vectors in $\mathbb{S}_2$. That is,

$$\langle \vec{v}_1, \vec{v}_2 \rangle = 0 \; \forall \vec{v}_1 \in \mathbb{S}_1, \vec{v}_2 \in \mathbb{S}_2.$$

(a) Recall that the *row space* of an $M \times N$ matrix $\mathbf{A}$ is the subspace spanned by the rows of $\mathbf{A}$ and that the *null space* of $\mathbf{A}$ is the subspace of all vectors $\vec{v}$ such that $\mathbf{A}\vec{v} = \vec{0}$.

Prove that the row space and null space of any matrix are orthogonal subspaces. This can be denoted by $\text{Col}(\mathbf{A}^T) \perp \text{Null}(\mathbf{A}) \; \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.

Answer:

First, we denote the rows of $\mathbf{A}$ as $\vec{a}_1^T, \vec{a}_2^T, \ldots, \vec{a}_M^T$. Now consider any vector $\vec{v} \in \text{Null}(\mathbf{A})$ which means that $\mathbf{A}\vec{v} = \vec{0}$. Note that matrix multiplication can be viewed as many inner products between the rows of $\mathbf{A}$ and the vector $\vec{v}$.

$$\mathbf{A}\vec{v} = \begin{bmatrix} \langle \vec{a}_1, \vec{v} \rangle \\ \langle \vec{a}_2, \vec{v} \rangle \\ \vdots \\ \langle \vec{a}_M, \vec{v} \rangle \end{bmatrix} = \vec{0}$$

Therefore, any vector $\vec{v} \in \text{Null}(\mathbf{A})$ is orthogonal to all rows of $\mathbf{A}$. From the bilinearity of the inner product, it follows that $\vec{v}$ is orthogonal to any linear combination of the rows of $\mathbf{A}$ and thus, any vector in $\text{Null}(\mathbf{A})$ is orthogonal to any vector in $\text{Col}(\mathbf{A}^T)$, proving that $\text{Col}(\mathbf{A}^T) \perp \text{Null}(\mathbf{A})\ \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.

(b) Recall that the *column space* of an $M \times N$ matrix $\mathbf{A}$ is the subspace spanned by the columns of $\mathbf{A}$ and that the *left null space* of $\mathbf{A}$ is the subspace of all vectors $\vec{v}$ such that $\vec{v}^T\mathbf{A} = \vec{0}^T \iff \mathbf{A}^T\vec{v} = \vec{0}$.
Prove that the column space and left null space of any matrix are orthogonal subspaces. This can be denoted by $\text{Col}(\mathbf{A}) \perp \text{Null}(\mathbf{A}^T)\ \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.

Answer:

We can define a new matrix $\mathbf{B} \triangleq \mathbf{A}^T$ and denote its rows as $\vec{b}_1^T, \vec{b}_2^T, \ldots, \vec{b}_N^T$. Using the same steps as in part (a), we can conclude that $\text{Col}(\mathbf{B}^T) \perp \text{Null}(\mathbf{B})\ \forall \mathbf{B} \in \mathbb{R}^{N \times M}$. Changing $\mathbf{B}$ back to $\mathbf{A}^T$ yields $\text{Col}(\mathbf{A}) \perp \text{Null}(\mathbf{A}^T)\ \forall \mathbf{A} \in \mathbb{R}^{M \times N}$, which is what we wanted to prove.

Alternately, we can view $\vec{v}^T\mathbf{A}$ as many inner products between $\vec{v}$ and the columns of $\mathbf{A}$. Let's denote the columns of $\mathbf{A}$ as $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_N$.

$$\vec{v}^T\mathbf{A} = \begin{bmatrix} \langle \vec{v}, \vec{a}_1 \rangle & \langle \vec{v}, \vec{a}_2 \rangle & \cdots & \langle \vec{v}, \vec{a}_N \rangle \end{bmatrix} = \vec{0}^T$$

Therefore, any vector $\vec{v} \in \text{Null}(\mathbf{A}^T)$ is orthogonal to all columns of $\mathbf{A}$. It follows that $\text{Col}(\mathbf{A}) \perp \text{Null}(\mathbf{A}^T)\ \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.