Lecture 14 - Inner product space

Tuesday, December 28, 2021 16:29

Small recap: In the beginning of our lecture course, we studied the basic concepts of linear algebra in the concrete context of $\mathbb{R}^n$. Since all this concepts were defined in terms of two operations: multiplication by $r \in \mathbb{R}$ and addition, we saw that it can be both natural and useful to apply the language developed to objects other then elements of $\mathbb{R}^n$ (for example, functions). We introduced the notion of a vector space over a field, for set that behaves like $\mathbb{R}^n$ as far as addition and scalar multiplication are concerned.

In this lecture, another operation for elements of $\mathbb{R}^n$ is examined: the dot product. We will see that it is very useful to define a product analogous to the dot product in a vector space (linear space) other then $\mathbb{R}^n$. Generalized dot products are called inner products. Once we have inner product in a vector space, we can define what it means for vectors to be orthogonal and define the length of vectors, just like in $\mathbb{R}^n$.

For this lecture, we will only be considering vector spaces over real numbers.

Definition: An inner product in a vector space (over $\mathbb{R}$) V is a function $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$ such that

Definition: An inner product on a vector space (over $\mathbb{R}$)

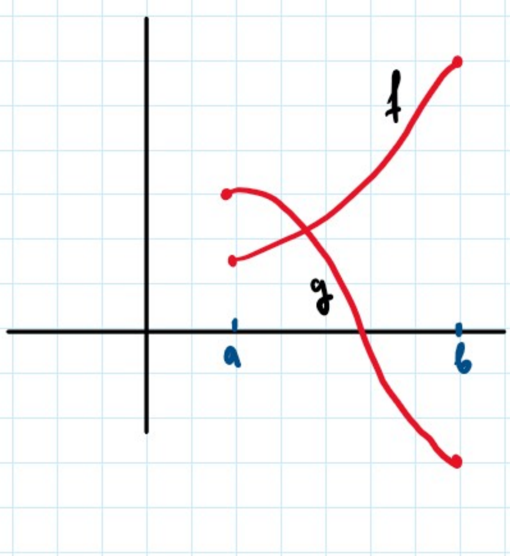
V is a function $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$, such that

1) $\langle f, g \rangle = \langle g, f \rangle$, for all $f, g \in V$.
2) $\langle f+h, g \rangle = \langle f, g \rangle + \langle h, g \rangle$, for all $f, g, h \in V$.
3) $\langle cf, g \rangle = c \langle f, g \rangle$, for all $c \in \mathbb{R}$, $f, g \in V$.
4) $\langle f, f \rangle > 0$, for all $f \in V$ and $f \neq 0$ (positive definitness property)

A vector space endowed with an inner product is called an inner product space.

Roughly, an inner product space $(V, \langle \cdot , \cdot \rangle)$ behaves like $\mathbb{R}^n$ as far as addition, scalar multiplication and dot product are concerned.

Example: Let $V = C[a,b] = \{ f : [a,b] \to \mathbb{R} \mid f \text{ continuous} \}$ for $a < b$, $a, b \in \mathbb{R}$.



For functions $f, g \in C[a,b]$, define

$$\langle f, g \rangle = \int_a^b f(t) \cdot g(t) \, dt .$$

This is an inner product. For example,

$$\langle f, g \rangle = \int_a^b f(t) \cdot g(t) \, dt = \int_a^b g(t) \, f(t) \, dt = \langle g, f \rangle$$

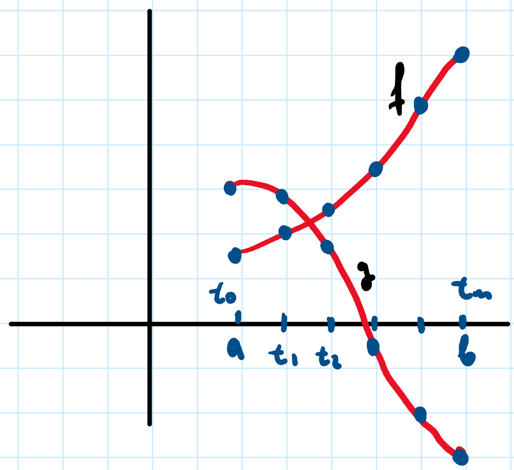
We will verify property 4) in CE.

Now the Riemann integral $\int_a^b f(t) \, g(t) \, dt$ is the we will verify property 4) in CE.

Now the Riemann integral $\int_a^b f(t)g(t)\,dt$ is the limit of the Riemann sum

$$\sum_{i=1}^{m} f(t_k)\,g(t_k)\,\Delta t$$

where the t_k can be chosen as equally spaced points on the interval $[a,b]$.



Then,

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt \approx \sum_{k=1}^{m} f(t_k)\,g(t_k)\,\Delta t =$$

$$= \left(\begin{pmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_m) \end{pmatrix} \cdot \begin{pmatrix} g(t_1) \\ g(t_2) \\ \vdots \\ g(t_m) \end{pmatrix} \right) \Delta t$$

for large m. This shows that inner product $\langle f,g\rangle$ for functions is a continuous version of the dot product. The more subdivisions we have, better will dot product approximate $\langle f,g\rangle$.

Example: Let $\mathbb{R}^{\mathbb{N}}$ denote the vector space of infinite sequences of real numbers.

$$\ell_2(\mathbb{N}) = \left\{ x \in \mathbb{R}^{\mathbb{N}} \mid \sum_{i=1}^{\infty} x_i^2 = x_1^2 + x_2^2 + \dots \text{ converges} \right\}$$

In this space we can define the inner product

$$\langle x,y\rangle = \sum_{i=1}^{\infty} x_i y_i = x_0 y_0 + x_1 y_1 + \dots$$

Example: Let $Mat_{n\times m}(\mathbb{R})$ be the vector space of all $n\times m$ matrices. We define

Example: Let $Mat_{n\times m}(\mathbb{R})$ be the vector space of all $n\times m$ matrices. We can define

$$\langle A, B\rangle = tr(A^T \cdot B)$$

Where A^T denotes the transpose of A, and $tr(A)$ is the trace of A. This is an inner product space (CE).

Definition: Let V be a (real) vector space. A norm on V is a function $\|\cdot\| : V \longrightarrow \mathbb{R}$ such that

a) $\|k\cdot v\| = |k|\cdot\|v\|$ for all $v\in V$ and $k\in\mathbb{R}$.

b) $\|v\| \geq 0$ (Positive)

c) $\|v\| = 0 \Rightarrow v = 0$ (Non-degenerate)

d) $\|u+v\| \leq \|u\| + \|v\|$ (Triangle inequality)

Lemma: Let V be a real inner product space.

Then

$$\|\cdot\| : V \longrightarrow \mathbb{R}$$

defined by

$$\|v\| := \sqrt{\langle v, v\rangle} \quad \text{is a norm.}$$

Proof: a) $\|k\cdot v\| = \sqrt{\langle k\cdot v, k\cdot v\rangle} = \sqrt{k^2\langle v, v\rangle}$

$= |k|\cdot\sqrt{\langle v, v\rangle} = |k|\cdot\|v\|.$

b) $\|v\| = \sqrt{\langle v, v\rangle} \geq 0$

c) Say $\|u\| = 0$. Then $\sqrt{\langle u, u\rangle} = 0 \Rightarrow \langle u, u\rangle = 0$

$\Rightarrow u = 0$ by positive definitness.

$\Rightarrow u = 0$ by positive definitness.

d) To prove this, we will use what is called

Caushy - Schwartz inequality: Let V be a (real) inner product space. If $u, v \in V$, then

$$\langle u, v \rangle \leq \|u\| \cdot \|v\|.$$

Then,

$$\|u+v\|^2 = \langle u+v, u+v \rangle = \langle u,u \rangle + 2\langle u,v \rangle + \langle v,v \rangle \overset{\text{by C.S.}}{\leq} \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$$

which is the triangle inequality.

We only have to prove Caushy - Schwartz, which we now do: $u=0$ case is obvious. Say $u \neq 0$.

Consider $u, v \in V$ and a function $p: \mathbb{R} \to \mathbb{R}$ defined by $p(t) = \langle tu+v, tu+v \rangle$. Since inner product is positive definite, $Im(p(t)) \subseteq \mathbb{R}_{\geq 0}$. On the other hand

$$p(t) = \langle tu+v, tu+v \rangle = \|u\|^2 t^2 + 2t\langle u,v \rangle + \|v\|^2$$

Since p is a polynomial of degree 2, such that sign of $p(t)$ does not change, the discriminant must be non-positive

$$D = 4(\langle u,v \rangle^2 - \|u\|^2 \|v\|^2) \leq 0$$ therefore

$$\mathcal{D} = 4\left(\langle u,v\rangle^2 - \|u\|^2\|v\|^2\right) \leq 0 \quad \text{therefore}$$

$$\|u\|\cdot\|v\| \geq \langle u,v\rangle. \qquad \square.$$

Definition: We will say that two vectors $u, v \in V$ in an inner product space V are **orthogonal**, if

$$\langle u, v\rangle = 0$$

In this case, we will write $u \perp v$. This corresponds to the familiar notion of perpendicularity in $\mathbb{R}^n$ (see CE).

Example: Say $f(t) = t^2 \in C[0,1]$, where inner product is given by $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$. Then

$$\|f\| = \sqrt{\langle f, f\rangle} = \sqrt{\int_0^1 t^4\, dt} = \sqrt{\frac{1}{5}}$$

Example: Show that $f(t) = \sin(t)$ and $g(t) = \cos(t)$ are orthogonal in the inner product space $C[0, 2\pi]$.

$$\langle f, g\rangle = \int_0^{2\pi} \sin(t)\cos(t)\, dt = \left(\frac{1}{2}\sin^2(t)\right)\Big|_0^{2\pi} = 0.$$

Definition: We say that the basis $v_1, \dots, v_i, \dots$ is an orthogonal basis of V, if vectors $v_1, \dots, v_i, \dots$ are pairwise orthogonal. If in addition vectors v_i have $\|v_i\| = 1$, then we say that basis $v_1, \dots, v_i, \dots$ is **orthonormal**.

Example: The canonical basis in $\mathbb{R}^n$ is orthonormal.

Example: The vectors $\left(\begin{smallmatrix} 1/\sqrt{2} \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1/\sqrt{2} \end{smallmatrix}\right)$ is an orthonormal

Example: The canonical basis in $\mathbb{R}^n$ is orthonormal.

Example: The vectors $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ and $\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ is an orthonormal basis of $\mathbb{R}^2$.

We say that list of vectors $v_1, \dots, v_m, \dots$ is orthogonal, if these vectors are pairwise orthogonal.

Lemma: If the nonzero vectors $v_1, \dots, v_m$ are pairwise orthogonal, then they are independent. In particular if $\dim V = m$, then $v_1, \dots, v_m$ form an orthogonal basis.

Proof: Suppose that

$$r_1 v_1 + \cdots + r_n v_n = 0$$

taking inner product with v_j on both sides:

$$0 = \langle r_1 v_1 + \cdots + r_n v_n, v_j \rangle =$$

$$= \sum_{i=1}^{m} r_i \langle v_i, v_j \rangle = r_j \langle v_j, v_j \rangle$$

as $\langle v_j, v_j \rangle \neq 0$, we get $r_j = 0$. Same for all i, j so $r_i = 0$, $i = 1, \dots, m$. Thus $v_1, \dots, v_m$ are linearly independent.

Lemma: If $(v_1, \dots, v_n)$ is orthonormal basis of V and $v \in V$, then

$$v = \sum_{i=1}^{n} r_i v_i \quad \text{where} \quad r_i = \langle v, v_i \rangle.$$

Proof: Now, since $(v_1, \dots, v_n)$ is a basis, we know

Proof: Now, since $(v_1, .., v_n)$ is a basis, we know that

$$v = \sum_{i=1}^{n} r_i v_i .$$

Again, taking scalar product with v_j on both sides

$$\langle v, v_j \rangle = \sum r_i \langle v_i, v_j \rangle = r_j \langle v_j, v_j \rangle = r_j \quad \square$$

Very useful property of inner product is that we get canonically defined complementary linear subspaces.

Lemma: Let V be a finite dimentional real inner product space. If $U \subset V$ is a subspace, then let

$$U^{\perp} = \{ w \in V \mid \langle w, u \rangle = 0 \text{ for all } u \in U \}$$

the set of all vectors orthogonal to all vectors in U. Then

a) $U^{\perp}$ is a subspace of V.

b) $U \cap U^{\perp} = \{0\}$

c) U and $U^{\perp}$ span V. (No proof of c))

Proof: a) Suppose w_1 and w_2 are in $U^{\perp}$. Pick $u \in U$.

Then $\langle w_1 + w_2, u \rangle = \langle w_1, u \rangle + \langle w_2, u \rangle = 0 + 0$

$= 0$

Thus $w_1 + w_2 \in U^{\perp}$. So $U^{\perp}$ is closed under addition.

Now suppose $w \in U^{\perp}$ and $\lambda \in \mathbb{R}$. Then

$$(\lambda w, u) = \lambda (w, u) = \lambda \cdot 0 = 0$$

Therefor $\lambda w \in U^{\perp}$ and $U^{\perp}$ is also closed under scalar

Therefor $\lambda w \in U^\perp$ and $U^\perp$ is also closed under scalar multiplication.

b) Say $w \in U \cap U^\perp$. Then

$$\langle w, w \rangle = 0$$

and by positive-definitness $w = 0$. $\square$.

Lemma. Let A be (real) matrix and $W = \mathrm{Im}(A)$. Then $W^\perp = \ker(A^T)$

Proof: Say

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \end{pmatrix}, \text{ then}$$

$$A^T = \begin{pmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \\ - & v_m^T & - \end{pmatrix}. \text{ Then } A^T x = \begin{pmatrix} \langle x, v_1 \rangle \\ \langle x, v_2 \rangle \\ \vdots \\ \langle x, v_m \rangle \end{pmatrix}$$

So, $x \in \ker A^T \iff \langle x, v_i \rangle = 0$ for all $i = 1, \dots, m$.

Now if $\langle x, v_i \rangle = 0$ for all $i = 1, \dots, m$. Then

$$\langle x, w \rangle = \langle x, r_1 v_1 + \dots + r_m v_m \rangle = r_1 \langle x, v_1 \rangle + \dots + r_m \langle x, v_m \rangle$$

$$= 0 \text{ for all } w \in W \text{ and } w = \sum_{i=1}^m r_i v_i.$$

Therfore $x \in W^\perp$. So $\ker(A^T) \subseteq W^\perp$

On the other hand if $x \in W^\perp$. in particular

On the other hand if $x \in W^{\perp}$, in particular $\langle x, v_i \rangle = 0$ for all $i = 1, \dots, m$ and $A^T x = \begin{pmatrix} \langle x, v_1 \rangle \\ \vdots \\ \langle x, v_m \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.
and $W^{\perp} \subseteq \ker(A^T)$. $\square$