Sec 6.7 Inner product spaces

Idea: Generalize the notions of length, distance, orthogonality, the Gram-Schmidt process, and "best approximation" from $\mathbb{R}^n$ to general vector spaces.

I. Def Let V be a vector space.
An inner product on V is a function

"Cartesian product
V cross V"

$$V \times V \longrightarrow \mathbb{R}$$

$\{$ tupples (a,b) where $a,b \in V\}$

$$(u,v) \longmapsto \langle u,v \rangle$$

that satisfies the following axioms for all $u,v,w \in V$ and $c \in \mathbb{R}$:

(1) $\langle u,v\rangle = \langle v,u\rangle$ (symmetry)

(2) $\langle u+v, w\rangle = \langle u,w\rangle + \langle v,w\rangle$ (additivity)

(3) $\langle cu, v\rangle = c\langle u,v\rangle$

(4) $\langle u,u\rangle \geq 0$ (positivity)

(5) $\langle u,u\rangle = 0$ if and only if $u =$ (zero element)

Def A vector space with an inner product is called an inner product space

Ex $\mathbb{R}^n$ with $\langle u,v\rangle$ defined by the usual dot product $u \cdot v$ is an inner product space (by Thm I of Sec 6.1)

Ex $\mathbb{R}^2$ with $\left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right\rangle = 4u_1v_1 + 5u_2v_2$ is an inner product space.

Proof (partial)

(Axiom 3) $\langle c\vec{u}, \vec{v}\rangle = 4(cu_1)v_1 + 5(cu_2)v_2 = c\left[4u_1v_1 + 5u_2v_2\right] = c\langle \vec{u},\vec{v}\rangle$

(Axiom 4) $\langle \vec{u},\vec{u}\rangle = 4u_1^2 + 5u_2^2 \geq 0$

Ex $\mathbb{P}_2$ with $\langle p,q\rangle \overset{\text{def}}{=} p(0)\,q(0) + p(\tfrac{1}{2})\,q(\tfrac{1}{2}) + p(1)\,q(1)$

(Example 3) is an inner product space.

If $p(t) = 12t^2$ and $q(t) = 2t-1$, we have ...

$$\langle p,q\rangle = p(0)\,q(0) + p(\tfrac{1}{2})\,q(\tfrac{1}{2}) + p(1)\,q(1)$$

$$= 0\,(-1) + 12(\tfrac{1}{4})^2\underbrace{\left(2(\tfrac{1}{2})-1\right)}_{0} + 12(1)\,(2(1)-1)$$

$$= 0 + 0 + 12\,(1) = 12$$

and $\langle q,q\rangle = q(0)\,q(0) + q(\tfrac{1}{2})\,q(\tfrac{1}{2}) + q(1)\,q(1)$

$$= (-1)(-1) + (0)(0) + (1)(1) = 2$$

Ex In fact, given any distinct $t_0, t_1, t_2 \in \mathbb{R}$,

(Example 2) $\mathbb{P}_2$ with $\langle p,q\rangle \overset{\text{def}}{=} p(t_0)\,q(t_0) + p(t_1)\,q(t_1) + p(t_2)\,q(t_2)$

is an inner product space.

In the previous example, $t_0 = 0$, $t_1 = \frac{1}{2}$, $t_2 = 1$

II Lengths, distances, & orthogonality

If V is an inner product space with inner product denoted by $\langle u,v\rangle$

- define the length (or norm) of an element v in V to be

$$\|v\| \overset{\text{def}}{=} \sqrt{\underbrace{\langle v,v\rangle}}$$

Note: This number is 0 or positive

note: this is a scalar

Take the square of both sides:

$$\|v\|^2 = \langle v,v\rangle$$

- The distance between elements u and v in V is $\|u-v\|$.

- We say $u, v \in V$ are orthogonal if $\langle u,v\rangle = 0$.

Ex From previous ex, $\langle q,q\rangle = 2$, so $\|q\| = \sqrt{2}$

(Example 4) $\langle p,q\rangle = 12 \neq 0$ so p and q are not orthogonal

III The Orthogonal Decomposition Theorem & Best Approximation

(Sec 6.3 Thm 8, generalized) (Sec 6.3 Thm 9, generalized)

Let V be an inner product space with inner product denoted by $\langle u, v \rangle$. Let W be a subspace of V, and let $\vec{y}$ be an element of V.

(a) Then we can decompose $\vec{y}$ into the sum of 2 vectors:

$$\vec{y} = \hat{y} + \vec{z} \qquad (*)$$

$\hat{y}$ is in W and $\underbrace{\vec{z} \text{ is orthogonal to } W}_{z \text{ is in } W^{\perp}}$

(b) If $\{u_1, \dots, u_p\}$ is an orthogonal basis of W,

then $\hat{y} = \underbrace{\frac{\langle y, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1}_{\text{proj}_{u_1} y} + \dots + \underbrace{\frac{\langle y, u_p \rangle}{\langle u_p, u_p \rangle} u_p}_{\text{proj}_{u_p} y}$ and $\underset{\text{immediate from eq } (*)}{z = y - \hat{y}}$

(c) The closest point in W to y is

$\hat{y} = \text{proj}_W y$, the orthogonal projection of y onto W,

i.e. $\| y - \hat{y} \| < \| y - v \|$ for any v in W that's not equal to $\hat{y}$.

IV. An inner product on $C[a,b]$.

Take any polynomial and define its length using the "evaluation" inner product from earlier.

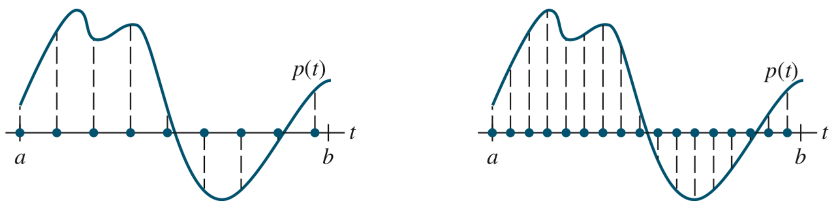


FIGURE 4 Using different numbers of evaluation points in $[a, b]$ to compute $\|p\|^2$.

using 9 points $t_0, t_1, \dots, t_8$ using more points $t_0, t_1, \dots, t_n$

Each interval is of length $\Delta t = \dfrac{b-a}{n+1}$ $\left(\text{so } \dfrac{1}{n+1} = \dfrac{\Delta t}{b-a} \right)$

If n is large, we get a large value of $\langle p, p \rangle$, so scale it down by dividing by $n+1$

$$\langle p, q \rangle = \frac{1}{n+1} \left[p(t_0) q(t_0) + \dots + p(t_n) q(t_n) \right]$$

$$= \frac{1}{b-a} \left[\sum_{j=0}^{n} p(t_j) q(t_j) \Delta t \right]$$

As n approaches infinity, we get

$$\frac{1}{b-a} \int_a^b p(t) q(t) \, dt$$

We can use this to define an inner product on $C[a,b]$, the vector space of all continuous functions on $[a,b]$.

For simplicity, remove the factor $\dfrac{1}{b-a}$

Ex $\langle f, g \rangle \stackrel{\text{def}}{=} \displaystyle\int_a^b f(t) g(t) \, dt$

defines an inner product on $C[a,b]$.

V The Gram-Schmidt Process

Ex Let $V = C[0,1]$ with inner product $\langle f, g\rangle = \int_0^1 f(t)\,g(t)\,dt$.

Let $f(t) = 1$, $g(t) = 2t-1$, $h(t) = 12t^2$.

Then $\langle f, g\rangle = \int_0^1 (1)(2t-1)\,dt = \frac{2t^2}{2} - t \Big|_0^1 = (1-1)-(0-0) = 0$

$\langle f, h\rangle = \int_0^1 (1)\,12t^2\,dt = \frac{12t^3}{3}\Big|_0^1 = 4-0 = 4$

So f, g are orthogonal, but f and h are not.

Let $W = \mathrm{Span}\{f, g, h\}$. Find an orthogonal basis $\{q_1, q_2, q_3\}$ for W.

Sol: ① Take $q_1 = f$ and ② $q_2 = g$ since f and g are already orthogonal.

Let $W_2 = \mathrm{Span}\{q_1, q_2\}$.

③ Set $q_3 = h - \mathrm{proj}_{W_2} h$

$\mathrm{proj}_{W_2} h = \mathrm{proj}_{q_1} h + \mathrm{proj}_{q_2} h$

$= \frac{\langle h, q_1\rangle}{\langle q_1, q_1\rangle} q_1 + \frac{\langle h, q_2\rangle}{\langle q_2, q_2\rangle} q_2 = \frac{4}{1} q_1 + \frac{2}{(1/3)} q_2 = 4 + 6(2t-1)$

$= 12t - 2$

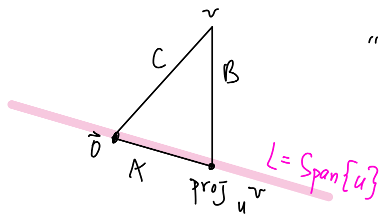
$$\boxed{\begin{aligned}&\langle h, q_1\rangle = \langle h, f\rangle = 4\\ &\langle q_1, q_1\rangle = \int_0^1 (1)(1)\,dt = t\Big|_0^1 = 1\end{aligned}}$$

$$\boxed{\begin{aligned}\langle h, q_2\rangle &= \int_0^1 12t^2(2t-1)\,dt\\ &= \int_0^1 24t^3 - 12t^2\,dt = 2\\ \langle q_2, q_2\rangle &= \int_0^1 (2t-1)^2\,dt\\ &= \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}\end{aligned}}$$

So $q_3 = \overbrace{12t^2}^{h} - \overbrace{(12t-2)}^{\mathrm{proj}_{W_2} h} = \boxed{12t^2 - 12t + 2}$

So an orthogonal basis for W is $\{1,\ 2t-1,\ 12t^2 - 12t + 2\}$. □

VI. Cauchy-Schwarz Inequality



"Pythagorean Theorem":

$$\underbrace{\| v \|^2}_{C} = \underbrace{\| \text{proj}_u v \|^2}_{A} + \underbrace{\| v - \text{proj}_u v \|^2}_{B}$$

<u>Observation</u> $\| \text{proj}_u v \|$ is no bigger than $\|v\|$ itself.

Suppose $u \neq$ the zero elt. Then ...

$$\| \text{proj}_u v \| = \left\| \frac{\langle v, u \rangle}{\langle u, u \rangle} u \right\| \quad \text{by def of } \text{proj}_u v$$

$$= \frac{|\langle v, u \rangle|}{|\langle u, u \rangle|} \|u\| \quad \text{since } \|cu\| = |c| \, \|u\| \text{ for any scalar } c$$

$$= \frac{|\langle v, u \rangle|}{\|u\|^2} \|u\| \quad \text{by def of } \|u\|^2$$

$$= \frac{|\langle u, v \rangle|}{\|u\|}$$

Since $\| \text{proj}_u v \| \leq \|v\|$ by above "Observation",

we have $\dfrac{|\langle u, v \rangle|}{\|u\|} \leq \|v\|$

so $\boxed{|\langle u, v \rangle| \leq \|u\| \, \|v\|}$

This is the <u>Cauchy-Schwarz inequality</u>, useful in many branches of math.

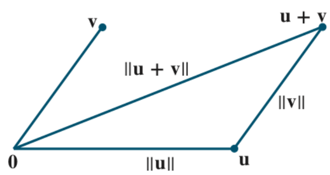
(Extra)

The Cauchy–Schwarz inequality is useful in many branches of mathematics. A few simple applications are presented in the exercises. Our main need for this inequality here is to prove another fundamental inequality involving norms of vectors. See Figure 3.

The Triangle Inequality (Theorem 17)

For all **u**, **v** in V,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$



PROOF

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v} \rangle| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \, \|\mathbf{v}\| + \|\mathbf{v}\|^2 \qquad \text{Cauchy–Schwarz} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

The triangle inequality follows immediately by taking square roots of both sides. ■

— end of PDF —