EL3370 Mathematical Methods in Signals, Systems and Control

Topic 2: Inner Product Spaces

Cristian R. Rojas

Division of Decision and Control Systems
KTH Royal Institute of Technology

Outline

Outline

Motivation and Definitions

Consider the space $\mathbb{C}^n$. It has:

1. *Vector space* (algebraic) structure:
   Given $x, y \in \mathbb{C}^n$, their sum $x + y$ and scalar multiplication αx ($\alpha \in \mathbb{C}$) are defined.
2. *Inner product* structure:

$$(x, y) = \sum_{i=1}^{n} x_i \overline{y}_i, \qquad x = (x_1, \ldots, x_n), \quad y = (y_1, \ldots, y_n) \in \mathbb{C}^n. \quad (\overline{x}\text{: } \textit{complex conjugate} \text{ of } x \in \mathbb{C})$$

Many physical properties (*e.g.*, work) can be defined in terms of inner products.
Also, $(\cdot, \cdot)$ can define: *distances* (metrics), *length* (norms), *angles*, *limits* (topologies), ...

Goal: Extend inner products to general (possibly infinite dimensional) vector spaces.

Motivation and Definitions (cont.)

Definition

Let ℓ_2 denote the vector space over $\mathbb{C}$ of all complex sequences $x = (x_n)$ which are *square summable*, *i.e.*, that satisfy $\sum_{n=1}^{\infty} |x_n|^2 < \infty$, with componentwise addition and scalar multiplication:

$$\begin{aligned} x + y &:= (x_n + y_n), \qquad && x = (x_n),\ y = (y_n) \in \ell_2, \\ \alpha x &:= (\alpha x_n), && \alpha \in \mathbb{C}, \end{aligned}$$

and inner product: $(x, y) := \displaystyle\sum_{n=1}^{\infty} x_n \overline{y}_n$.

Observation

Need to verify that these operations (sum, scalar multiplication, inner product) are valid!

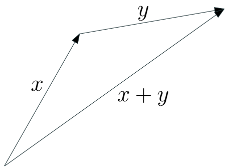
(We will do it later, using the Cauchy-Schwarz inequality.)

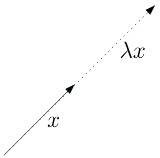
Definition (reminder)

A *vector space* V over a field F (*e.g.*, $\mathbb{R}$ or $\mathbb{C}$) is a set with two operations, *sum* ($x+y \in V$, for $x, y \in V$) and *scalar multiplication* ($\lambda x \in V$, for $x \in V$ and $\lambda \in F$) s.t. for all $x, y, z \in V$, $\alpha, \beta \in F$:

1. $x+y = y+x$, (*commutativity*)
2. $(x+y)+z = x+(y+z)$, (*associativity*)
3. There is a null vector $0 \in V$ s.t. $0+x = x$,
4. $\alpha(x+y) = \alpha x + \alpha y$, (*distributivity*)
5. $(\alpha+\beta)x = \alpha x + \beta x$, (*distributivity*)
6. $(\alpha\beta)x = \alpha(\beta x)$, (*associativity*)
7. $1x = x$.

A *field* F is a set with operations + and $\cdot$ which are: associative and commutative; F has additive and multiplicative identities (0 and 1, respectively); every $a \in F$ has an additive inverse ($-a$) and, if $a \neq 0$, a multiplicative inverse too ($a^{-1} \in F$); and $\cdot$ is distributive with respect to +: $a \cdot (b+c) = a \cdot b + a \cdot c$ for all $a, b, c \in F$.





Definition (reminder)

Let V be a vector space over F; $\alpha_1,\dots,\alpha_n \in F$; $x_1,\dots,x_n \in V$; and $X \subseteq V$.

- *Linear subspace* B of V: subset of V s.t., if $x, y \in B$, $\alpha \in F$, then $x + y \in B$ and $\alpha x \in B$.
- *Affine subspace* (or *linear variety*) B of V: subset of V of the form $x + M := \{x + m : m \in M\}$, where $x \in V$ and M is a linear subspace of V.
- *Linear combination* of $x_1,\dots,x_n$: an element $\alpha_1 x_1 + \cdots + \alpha_n x_n \in V$ (for *finite* n).
- lin X (*span* of X): set of all linear combinations of elements of X.
  Note. lin X is the intersection of all linear subspaces of V containing X (*why?*).
- If for every linear combination $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$, $x_1,\dots,x_n \in X$, we have that $\alpha_1 = \cdots = \alpha_n = 0$, X is *linearly independent* (l.i.). If not, X is *linearly dependent* (l.d.).
- *Basis* of V: A linearly independent set $X \subseteq V$ which spans V (*i.e.*, lin $X = V$).
- $\dim V$ (*dimension* of V): number of elements of some basis of V. (All bases of V have the same number of elements; *why?*).
- If $\dim V < \infty$, V is a *finite-dimensional vector space*. (**Obs**: V is not necessarily finite!)

Motivation and Definitions (cont.)

Definition (inner products)

An *inner product* (*scalar product*) on a vector space V over $\mathbb{C}$ is a mapping $(\cdot,\cdot)\colon V \times V \to \mathbb{C}$ s.t. for all $x, y, z \in V$ and $\lambda \in \mathbb{C}$:

1. $(x, y) = \overline{(y, x)}$,
2. $(\lambda x, y) = \lambda(x, y)$,
3. $(x + y, z) = (x, z) + (y, z)$,
4. $(x, x) > 0$ when $x \neq 0$.

$(V, (\cdot,\cdot))$ is an *inner product space* (or *pre-Hilbert space*).

Examples

1. Complex vector space $C[0,1] := \{f : [0,1] \to \mathbb{C} \colon f \text{ is continuous}\}$, with point-wise addition and scalar multiplication ($(f+g)(t) = f(t) + g(t)$, $(\lambda f)(t) = \lambda f(t)$ for $f, g \in C[0,1]$, $\lambda \in \mathbb{C}$ and $t \in [0,1]$), and inner product $(f, g) = \int_0^1 f(t)\overline{g(t)}dt$.
2. Space $\mathbb{C}^{m \times n}$ of $m \times n$ complex matrices, with inner product $(A, B) = \operatorname{tr}(AB^H)$.

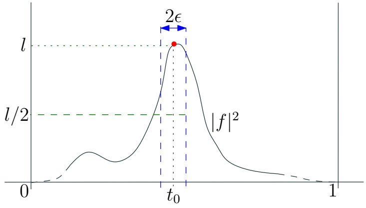
Motivation and Definitions (cont.)

Proof for Example 1:

Since $C[0,1]$ is a complex vector space (*exercise!*), we need to verify that $(\cdot,\cdot)$ satisfies the axioms of an inner product. Let $f,g,h \in C[0,1]$ and $\lambda \in \mathbb{C}$:

1. $(f,g) = \int_0^1 f(t)\overline{g(t)}dt = \overline{\int_0^1 g(t)\overline{f(t)}dt} = \overline{(g,f)}$.
2. $(\lambda f,g) = \int_0^1 \lambda f(t)\overline{g(t)}dt = \lambda \int_0^1 f(t)\overline{g(t)}dt = \lambda(f,g)$.
3. $(f+g,h) = \int_0^1 [f(t)+g(t)]\overline{h(t)}dt = \int_0^1 f(t)\overline{h(t)}dt + \int_0^1 g(t)\overline{h(t)}dt = (f,h)+(g,h)$.
4. If $f \neq 0$, then there is a $t_0 \in [0,1]$ s.t. $l := |f(t_0)|^2 \neq 0$. Since $|f|^2$ is continuous, there is an $\varepsilon > 0$ s.t. $|f(t)|^2 > l/2$ whenever $|t-t_0| < \varepsilon$.
   Therefore,

$$
\begin{aligned}
(f,f) &= \int_0^1 |f(t)|^2 dt \\
&\geqslant \int_{\{t\in[0,1]:\ |t-t_0|<\varepsilon\}} |f(t)|^2 dt \\
&\geqslant \varepsilon \frac{l}{2} > 0.
\end{aligned}
$$



□

Theorem

For every $\lambda \in \mathbb{C}$ and x, y, z in an inner product space V,

(i) $(x, y+z) = (x, y) + (x, z)$,

(ii) $(x, \lambda y) = \overline{\lambda}(x, y)$,

(iii) $(x, 0) = (0, x) = 0$,

(iv) If $(x, z) = (y, z)$ for all $z \in V$, then $x = y$.

Proof

(i) By definition: $(x, y+z) = \overline{(y+z, x)} = \overline{(y, x)} + \overline{(z, x)}$.

(ii) Similar to (i).

(iii) Notice that $(x, 0) = (x, 0y)$, and use (ii).

(iv) Since $(x, z) = (y, z)$, then $(x-y, z) = 0$. Since this holds for every z, take $z = x-y$, which gives $(x-y, x-y) = 0$. By the last axiom of an inner product, this implies $x-y = 0$. $\square$

Outline

Inner Product Spaces as Normed Spaces

Idea: Inner products $\implies$ lengths (*norms*) $\implies$ distances (*metrics*).

Example: In $\mathbb{R}^n$, $(x,y) = x^T y \implies$ length $= \|x\| = \sqrt{x^T x} = \sqrt{(x,x)} \implies$ distance $= \|x-y\|$.

Definition
In an inner product space V, the *norm* of a vector $x \in V$ is $\|x\| := \sqrt{(x,x)}$.

Examples

1. For $x = (x_1, \dots, x_n) \in \mathbb{C}^n$: $\|x\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$.
2. For $f \in C[0,1]$: $\|f\| = \sqrt{\int_0^1 |f(t)|^2 dt}$.

Inner Product Spaces as Normed Spaces (cont.)

Theorem. For every x, y in an inner product space V, and $\lambda \in \mathbb{C}$:

(i) $\|x\| \geqslant 0$, and $\|x\| = 0$ iff $x = 0$,

(ii) $\|\lambda x\| = |\lambda| \|x\|$,

(iii) $|(x, y)| \leqslant \|x\| \|y\|$, with equality iff $x = \alpha y$ for some $\alpha \in \mathbb{C}$, (*Cauchy-Schwarz inequality*)

(iv) $\|x + y\| \leqslant \|x\| + \|y\|$. (*triangle inequality*)

Proof. (i) Direct from last axiom of an inner product.

(ii) $\|\lambda x\| = \sqrt{(\lambda x, \lambda x)} = \sqrt{\lambda (x, \lambda x)} = |\lambda| \sqrt{(x, x)} = |\lambda| \|x\|$.

(iii) For every $\alpha \in \mathbb{C}$: $0 \leqslant (x - \alpha y, x - \alpha y) = \|x\|^2 - 2\mathrm{Re}\{\overline{\alpha}(x, y)\} + |\alpha|^2 \|y\|^2$.

Take $\alpha = tu$, where $t \in \mathbb{R}$ and $u = \exp(i \arg(x, y))$, which gives $0 \leqslant \|x\|^2 - 2t|(x, y)| + t^2 \|y\|^2$.

The minimum of this quadratic expression w.r.t. t is $\|x\|^2 - |(x, y)|^2 / \|y\|^2$, which must be non-negative.

Furthermore, this is zero iff $x - \alpha y = 0$ for some $\alpha \in \mathbb{C}$.

(iv) By (iii),

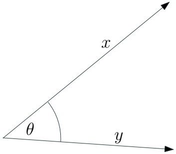
$$\|x + y\|^2 \leqslant \|x\|^2 + 2\mathrm{Re}\{(x, y)\} + \|y\|^2 \leqslant \|x\|^2 + 2|(x, y)| + \|y\|^2 \leqslant \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \qquad \square$$

Inner Product Spaces as Normed Spaces (cont.)

Applications of Cauchy-Schwarz inequality

Angle between vectors

$$\cos\theta := \frac{(x,y)}{\|x\|\,\|y\|}.$$



Probability
Let V be an inner product space of zero mean real random variables x with $\mathrm{E}\{x^2\} < \infty$, and inner product $(x,y) := \mathrm{E}\{xy\} = \mathrm{cov}(x,y)$. Then the Cauchy-Schwarz inequality implies

$$|\mathrm{cov}(x,y)|^2 = |(x,y)|^2 \leqslant \|x\|^2 \|y\|^2 = \mathrm{var}(x)\mathrm{var}(y).$$

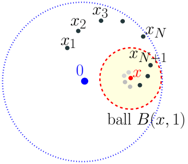
Exercise: Prove that the operations in ℓ_2 are well defined.

Applications of Cauchy-Schwarz inequality (cont.)

Theorem. In an inner product space V, the inner product is a continuous function, *i.e.*, for every sequences $(x_n),(y_n)$ s.t. $x_n \to x$ and $y_n \to y$, we have $(x_n, y_n) \to (x, y)$.

Proof. By Cauchy-Schwarz,

$$\begin{aligned} |(x,y)-(x_n,y_n)| &= |(x,y)-(x_n,y)+(x_n,y)-(x_n,y_n)| \\ &\leqslant |(x-x_n,y)|+|(x_n,y-y_n)| \\ &\leqslant \|y\|\|x-x_n\|+\|x_n\|\|y-y_n\|. \end{aligned}$$



Since (x_n) is convergent, it is also *bounded* (*i.e.*, there is an $M > 0$ s.t. $\|x_n\| \leqslant M$ for all $n \in \mathbb{N}$). Indeed, since there is an $N \in \mathbb{N}$ s.t. $\|x_n - x\| < 1$ for $n > N$, so $\|x_n\| = \|x + x_n - x\| \leqslant \|x\| + \|x_n - x\| < \|x\| + 1$, we can take $M = \max\{\|x_1\|, \dots, \|x_N\|, \|x\| + 1\}$.

Then, given $\varepsilon > 0$, there is an $N_0 \in \mathbb{N}$ s.t. for $n > N_0$, $\|x_n - x\| < \varepsilon/(2\|y\|)$ and $\|y_n - y\| < \varepsilon/(2M)$, so $|(x,y)-(x_n,y_n)| \leqslant \|y\|[\varepsilon/(2\|y\|)] + M[\varepsilon/(2M)] = \varepsilon$. □

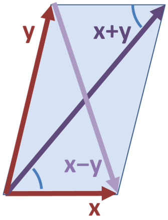
Inner Product Spaces as Normed Spaces (cont.)

Theorem (***Parallelogram Law***)
Let x, y be elements of an inner product space. Then,

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Proof. As $\|x \pm y\|^2 = \|x\|^2 \pm (x,y) \pm (y,x) + \|y\|^2$, the result follows by adding these expressions. $\square$

(*See bonus slides for converse result!*)



Theorem (***Polarization Identity***)
Let x, y be elements of an inner product space. Then,

$$(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) = \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|^2, \quad \text{(complex case)}$$

$$= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right). \quad \text{(real case)}$$

Proof. *Exercise!*

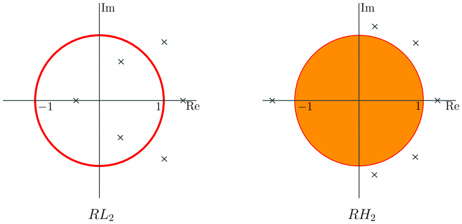
A more interesting example for system theory

RL_2: space of *rational functions*, analytic on *unit circle* $\partial\mathbb{D} := \{z \in \mathbb{C} : |z| = 1\}$ with usual addition and scalar multiplication, and inner product

$$(f,g) := \frac{1}{2\pi i}\oint_{\partial\mathbb{D}} f(z)\overline{g(z)}\frac{dz}{z} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\omega})\overline{g(e^{i\omega})}d\omega.$$

RH_2: subspace of RL_2, of functions analytic on *closed unit disc* $\overline{\mathbb{D}}$, where $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$.

In engineering terms:
RL_2 consists of rational functions without poles on $\partial\mathbb{D}$ (can be *stable* or *unstable*), and RH_2 only has functions with poles outside $\overline{\mathbb{D}}$ (*stable*).



Inner Product Spaces as Normed Spaces (cont.)

A more interesting example for system theory (cont.)

Exercise: Prove that RL_2 is an inner product space.

Cauchy integral formula simplifies calculations of inner products in RL_2: For $h \in RL_2$,

$$\frac{1}{2\pi i} \oint_{\partial \mathbb{D}} h(z) dz = \sum_{\substack{z_j = \text{pole of} \\ h \text{ in } \mathbb{D}}} \text{Res}_{z=z_j}[h(z)].$$

Example: $f(z) = \dfrac{1}{z-a}$, $g(z) = \dfrac{1}{z-b}$ $(|a| < 1,\ 0 < |b| < 1)$, thus

$$\begin{aligned}
(f,g) &= \frac{1}{2\pi i} \oint_{\partial \mathbb{D}} \frac{1}{z-a} \frac{1}{\overline{z}-b} \frac{dz}{z} = -\frac{1}{2\pi i b} \oint_{\partial \mathbb{D}} \frac{1}{z-a} \frac{1}{z-1/b} dz \qquad &&(\text{since } z\overline{z} = 1 \text{ in } \partial \mathbb{D}) \\
&= -\frac{1}{b} \text{Res}_{z=a} \left(\frac{h(z)}{z-a} \right) \quad \text{where } h(z) = \frac{1}{z-1/b} &&(h \text{ is analytic at } z = a) \\
&= -\frac{1}{b} h(a) = -\frac{1}{b} \frac{1}{a-1/b} = \frac{1}{1-ab}.
\end{aligned}$$

Next Topic

Normed Spaces

Outline

Bonus: Converse of the Parallelogram Law

The parallelogram law can be used to show that a given norm does not come from an inner product. However, when it holds, the norm can be used to derive an inner product!

Idea: Use the polarization identity! (consider the real case for simplicity)

$$(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right).$$

Let us check the properties of an inner product:

1. $(y,x) = \frac{1}{4}\left(\|y+x\|^2 - \|y-x\|^2\right) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) = (x,y).$
4. $(x,x) = \frac{1}{4}\left(\|x+x\|^2 - \|x-x\|^2\right) = \|x\|^2 > 0$ if $x \neq 0$.
3. Decompose $(x+y,z)$ in two different ways:

$$\begin{aligned}(x+y,z) &= \frac{1}{4}\left(\|x+y+z\|^2 - \|x+y-z\|^2\right) \\ &= \frac{1}{4}\left(\|x+y+z\|^2 + \|x-y+z\|^2 - \|x+y-z\|^2 - \|x-y+z\|^2\right) \\ &= \frac{1}{4}\left(\|x+y+z\|^2 + \|x-y-z\|^2 - \|x+y-z\|^2 - \|x-y-z\|^2\right).\end{aligned}$$

Bonus: Converse of the Parallelogram Law (cont.)

Applying the parallelogram law yields:

$$\begin{aligned}(x+y,z) &= \frac{1}{4}\left(2\|x+z\|^2 + 2\|y\|^2 - 2\|x\|^2 - 2\|y-z\|^2\right) \\ &= \frac{1}{4}\left(2\|x\|^2 + 2\|y+z\|^2 - 2\|y\|^2 - 2\|x-z\|^2\right).\end{aligned}$$

Averaging these expressions and applying the polarization identity gives

$$(x+y,z) = \frac{1}{4}\left(\|x+z\|^2 - \|y-z\|^2 + \|y+z\|^2 - \|x-z\|^2\right) = (x,z)+(y,z).$$

2. From the polarization identity and Property 3,

$$\begin{aligned}(-x,y) &= \frac{1}{4}\left(\|-x+y\|^2 - \|-x-y\|^2\right) = -\frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) = -(x,y), \\ (0,y) &= (x-x,y) = (x,y)+(-x,y) = (x,y)-(x,y) = 0, \\ ([n+1]x,y) &= (nx,y)+(x,y),\end{aligned}$$

so by induction on $n \in \mathbb{N}$ and the 1st expression,$(nx,y) = n(x,y)$ for all $n \in \mathbb{Z}$. Also, if $m,n \in \mathbb{Z}\setminus\{0\}$, $n([m/n]x,y) = (mx,y) = m(x,y)$, so $([m/n]x,y) = [m/n](x,y)$, thus $(\lambda x,y) = \lambda(x,y)$ for all $\lambda \in \mathbb{Q}$. Since norms are continuous (because $\big|\|x\| - \|y\|\big| \leqslant \|x-y\|$ from the triangle inequality), this last expression can be extended to all $\lambda \in \mathbb{R}$. □