From Euclid to Hilbert (1/2): Hilbert Spaces and Projection Operators

Mathematical Foundations of Signal Processing

Dr. Matthieu Simeoni

September 28, 2020

EPFL

Table of contents

A Unifying Framework for Signal Processing

Signals come in *various forms*:

$f:\{0,\cdots,N\}\to\mathbb{R}$ Vectors, compactly-supported or periodic sequences,

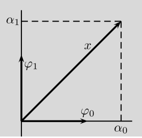
$f:\mathbb{Z}\to\mathbb{R}$ Generic sequences supported on $\mathbb{Z}$,

$f:[0,T)\to\mathbb{R}$ Compact-support or periodic functions,

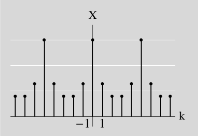
$f:\mathbb{R}\to\mathbb{R}$ Generic functions.

With the right abstraction, they can all be seen as vectors in some Hilbert spaces, which can be manipulated geometrically. Many advantages:

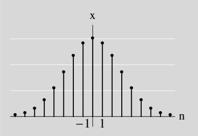
- Leverages “real world” geometric intuition.
- Unified understanding of fundamental signal processing concepts.
- Can go *faster* and *farther*!



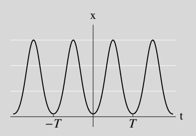
(a) Vector



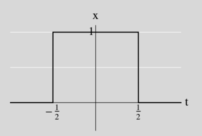
(b) Periodic sequence



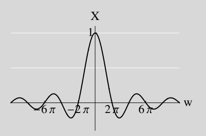
(c) Sequence



(d) Periodic function



(e) Compactly-supported function



(f) Infinitely-supported function

Objectives and Reading Material

In this lesson we will:

1. Develop the basic geometric intuition central to Hilbert spaces,
2. Discuss linear operators –with a particular focus on projections– which generalise finite-dimensional matrices.

Our aim is to extend Euclidean geometric insights to abstract signals. You will have to invest some effort to see signals as vectors in Hilbert spaces, but believe us, the effort is well placed!

Reading Material

- Chapter 2, "From Euclid to Hilbert", of [1], Sections 2.1 to 2.4 (in particular 2.3.3 and 2.4)
- Appendix 2.B, "Elements of Linear Algebra", of [1]. This is a review of basic concepts of linear algebra.[1]
- **Download link:** `http://www.fourierandwavelets.org/FSP_v1.1_2014.pdf`

[1] You should be familiar with most of these concepts, but do not hesitate to refresh your memory as we will assume these notions to be mastered!

Vector spaces

Vectors in $\mathbb{R}^N$ can be *added* or *scaled* while staying in $\mathbb{R}^N$. This can serve as an inspiration for defining *abstract vector spaces*:

Definition: (Vector Space)

A vector space over a field of scalars $\mathbb{F}$ (think $\mathbb{R}$ or $\mathbb{C}$) is a set of vectors V, together with operations of vector addition $+ : V \times V \to V$ and scalar multiplication $\cdot : \mathbb{F} \times V \to V$.
For any $x, y, z \in V$ and $\alpha, \beta \in \mathbb{F}$, these operations must moreover satisfy the following properties:

`Commutativity:` $x + y = y + x$.

`Associativity:` $(x + y) + z = x + (y + z)$ and $(\alpha\beta)x = \alpha(\beta x)$.

`Distributivity:` $\alpha(x + y) = \alpha x + \alpha y$ and $(\alpha + \beta)x = \alpha x + \beta x$.

`Additive identity:` There exists a null $0 \in V$, such that $x + 0 = 0 + x = x$, $\forall x \in V$.

`Additive inverse:` $\forall x \in V$, there exists a unique element $-x$ in V, such that $x + (-x) = (-x) + x = 0$.

`Multiplicative identity:` For every x in V, $1 \cdot x = x$.

Common Vector spaces

Examples of vector spaces

- **$\mathbb{C}$-valued finite-dimensional vectors:** $\mathbb{C}^N = \{[x_0 \quad x_1 \quad \dots \quad x_{N-1}] \mid x_n \in \mathbb{C},\ n \in \{0, 1, \dots, N-1\}\}$, where vector addition and scalar multiplication are defined *component-wise*.
- **$\mathbb{C}$-valued sequences over $\mathbb{Z}$:** $\mathbb{C}^{\mathbb{Z}} = \left\{x = \left[\dots \quad x_{-1} \quad \boxed{x_0} \quad x_1 \quad \dots\right] \,\middle|\, x_n \in \mathbb{C},\ n \in \mathbb{Z}\right\}$, where vector addition and scalar multiplication are defined *component-wise*.
- **$\mathbb{C}$-valued functions over $\mathbb{R}$:** $\mathbb{C}^{\mathbb{R}} = \{x \mid x(t) \in \mathbb{C},\ t \in \mathbb{R}\}$, where vector addition and scalar multiplication are defined *point-wise*.
- **$\mathbb{C}$-valued $N \times M$ rectangular matrices:** $\mathbb{C}^{N\times M} = \left\{\begin{bmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NN} \end{bmatrix} \,\middle|\, x_{nm} \in \mathbb{C},\ n \leq N,\ m \leq M\right\}$, where vector addition and scalar multiplication are defined *element-wise*.

Subspace & Span

A subspace is a set of vectors *closed* under vector addition and scalar multiplication:

Definition: (Subspace)

$S \subseteq V$ is a subspace if: $\forall x, y \in S, \quad \alpha, \beta \in \mathbb{F}, \qquad \alpha x + \beta y \in S.$

Examples: the subspace of *even/odd* functions, the subspace of *symmetric* matrices.
Non example: the set of *positive* functions (not closed w.r.t. addition or multiplication).

It is possible to generate (*span*) subspaces from a collection of vectors:

Definition: (Span)

The span of (potentially infinite) set of vectors $S \subset V$ consists in all finite linear combinations of vectors in S:

$$\text{span}(S) = \left\{ \sum_{k=0}^{N-1} \alpha_k \varphi_k \,\middle|\, \alpha_k \in \mathbb{F}, \varphi_k \in S \text{ and } N \in \mathbb{N} \right\}.$$

Linear Independence & Dimension

Many different sets can have the same span. It can be of interest to find the smallest spanning set. This leads to the dimension of a vector space, which depends on the concept of linear independence:

Definition: (Linear Independence)

$S=\{\varphi_k\}_{k=0}^{N-1}\subset V$ is said to be linerarly independent if:[2]

$$\sum_{k=0}^{N-1}\alpha_k\varphi_k=0 \Longrightarrow \alpha_k=0,\quad \forall k=0,\dots,N-1. \tag{1}$$

If (1) does not hold, S is said linearly dependent.

A vector space V is said to have dimension N when it contains a linearly independent set with N elements and every set with $N+1$ or more elements is linearly dependent. If no such finite N exists, V is infinite-dimensional.

[2] If S is infinite, we need every finite subset to be linearly independent.

Extending Euclidean Geometric Notions

We wish to extend to abstract vector spaces some useful Euclidean *geometric notions*:

Norm The norm of a vector $\boldsymbol{x} \in \mathbb{R}^2$ can be interpreted as its *length*:

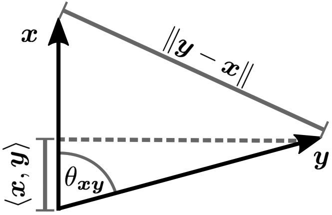
$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2}.$$

Distance The distance between two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ is the *norm of their difference*:

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

Orientation The relative orientation –or *angle*– between two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ is obtained via the inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + x_2 y_2$:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \cos\theta_{\boldsymbol{xy}} \Rightarrow \theta_{\boldsymbol{xy}} = \arccos\left(\frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|}\right).$$



Inner products

Inner products help us formalise the geometric notions of *orientation* and *orthogonality*. They measure the linear resemblance between two vectors.

Definition: (Inner product)

An inner product for V is a map $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$ satisfying, for any $x, y, z \in V$ and $\alpha \in \mathbb{F}$,

1. Distributivity: $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
2. Linearity in the 1^{st} argument : $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$
3. Hermitian symmetry: $\langle x, y\rangle = \overline{\langle y, x\rangle}$
4. Positive definiteness: $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ iff $x = 0$.

A vector space equipped with an inner product is called an inner product space or pre-Hilbert space.

Note that from Items 2 and 3 we get:[3] $\langle x, \alpha y\rangle = \overline{\alpha}\langle x, y\rangle$.

[3] We say that the inner product is antilinear w.r.t. its 2^{nd} argument.

Standard Inner Products

Standard Inner Products for Common Vector Spaces

$$\mathbb{C}^N: \quad \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{N} x_i \overline{y_i} = \boldsymbol{y}^H \boldsymbol{x}.$$

$$\mathbb{C}^{N \times M}: \quad \langle \boldsymbol{X}, \boldsymbol{Y} \rangle = \operatorname{tr}\left(\boldsymbol{X}\boldsymbol{Y}^H\right) = \sum_{i=1}^{N} \sum_{j=1}^{M} X_{i,j} \overline{Y_{i,j}} = \operatorname{vec}(\boldsymbol{Y})^H \operatorname{vec}(\boldsymbol{X}).$$

$$\mathbb{C}^{\mathbb{Z}}: \quad \langle x, y \rangle = \sum_{i \in \mathbb{Z}} x_i \overline{y_i}. \ \triangle$$

$$\mathbb{C}^{\mathbb{R}}: \quad \langle x, y \rangle = \int_{\mathbb{R}} x(t) \overline{y(t)} \, dt. \ \triangle$$

⚠ Note that inner products must be *finite*. Hence, for the last two inner products to be valid, the infinite sum/integral must converge! This restricts the set of sequences/functions on which we can operate.

Orthogonality

An inner product endows a space with geometric properties that arise from angles, such as perpendicularity and relative orientation. An inner product being *zero* has special significance.

Definition: (Orthogonality)

1. Vectors x and y are said to be orthogonal when $\langle x, y\rangle = 0$, written as $x \perp y$.
2. A set $S \subset V$ is called orthogonal when $x \perp y\ \forall x, y \in S,\ x \neq y$.
3. A set $S \subset V$ is called orthonormal when it is orthogonal and $\langle x, x\rangle = 1\ \forall x \in S$.
4. A vector x is said to be orthogonal to a set $S \subset V$ when $x \perp s\ \forall s \in S$, written as $x \perp S$.
5. Two sets $S_0 \subset V$ and $S_1 \subset V$ are orthogonal –written as $S_0 \perp S_1$– if $\forall s_0 \in S_0$ we have $s_0 \perp S_1$.
6. Given a subspace $S \subset V$, the orthogonal complement of S is the subspace $S^\perp = \{x \in V \mid x \perp S\}$.

Note that vectors in orthonormal set $\{\varphi_k\}_{k\in\mathcal{K}}$ are linearly independent since $0 = \sum_{k\in\mathcal{K}} \alpha_k \varphi_k$ implies:

$$0 = \langle 0, \varphi_i\rangle = \left\langle \sum_{k\in\mathcal{K}} \alpha_k \varphi_k, \varphi_i \right\rangle = \sum_{k\in\mathcal{K}} \alpha_k \langle \varphi_k, \varphi_i\rangle = \sum_{k\in\mathcal{K}} \alpha_k \delta_{i-k} = \alpha_i, \quad \forall i \in \mathcal{K}.$$

Example of Orthogonal Functions on $[-1, 1]$

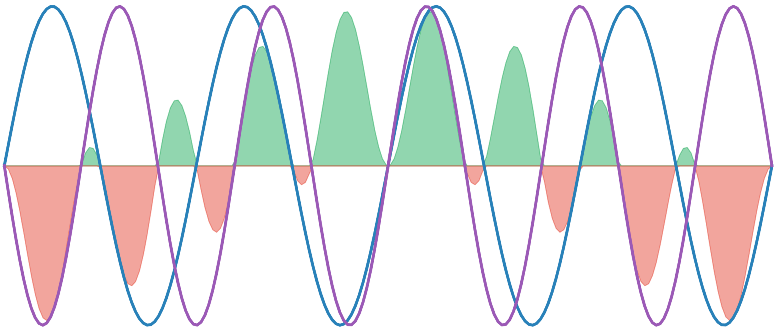


Figure: $x(t) = \sin(4\pi t)$, $x(t) = \sin(5\pi t)$, $\langle x, y\rangle = \int_{-1}^{1} \sin(4\pi t)\sin(5\pi t)\,dt = 0$

Norms

A norm is a function that assigns a length, or size, to a vector (analogously to the *magnitude* of a scalar).

Definition: (Norm)

A norm on V is a function $\|\cdot\| : V \to \mathbb{R}$ satisfying, for all $x, y \in V$ and $\alpha \in \mathbb{F}$:

1. Positive definiteness: $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$,
2. Positive scalability: $\|\alpha x\| = |\alpha| \|x\|$,
3. Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$ with equality iff $y = \alpha x$.

A vector space equipped with a norm becomes a normed vector space.[4] A normed vector space is also a metric space, since the norm can be used to define the *induced* metric[5] (or distance):

$$d(x, y) = \|x - y\|, \quad \forall x, y \in V.$$

[4] As with the inner product, we must exercise caution and choose the subspace for which the norm is finite.

[5] [1, Exercise 2.13] gives the axioms that a metric must satisfy and explores metrics that are not induced by norms.

Norms Induced by Inner Products

Any inner product induces a norm: $\|x\| = \sqrt{\langle x, x\rangle}, \quad \forall x \in V.$

Norms induced by Standard Inner Products

$$\mathbb{C}^N: \quad \|\boldsymbol{x}\|_2 = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle} = \sqrt{\sum_{i=1}^N |x_i|^2} = \sqrt{\boldsymbol{x}^H \boldsymbol{x}} \text{ } (\ell_2\text{-norm}).$$

$$\mathbb{C}^{N\times M}: \quad \|\boldsymbol{X}\|_F = \sqrt{\langle \boldsymbol{X}, \boldsymbol{X}\rangle} = \sqrt{\operatorname{tr}\left(\boldsymbol{X}\boldsymbol{X}^H\right)} = \sqrt{\sum_{i=1}^N \sum_{j=1}^M |X_{i,j}|^2} = \sqrt{\operatorname{vec}(\boldsymbol{X})^H \operatorname{vec}(\boldsymbol{X})} \text{ (Froebenius norm)}.$$

$$\mathbb{C}^{\mathbb{Z}}: \quad \|x\|_2 = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{i\in\mathbb{Z}} |x_i|^2} \text{ ⚠ } (\ell_2\text{-norm})$$

$$\mathbb{C}^{\mathbb{R}}: \quad \|x\|_2 = \sqrt{\langle x, x\rangle} = \sqrt{\int_{\mathbb{R}} |x(t)|^2 dt} \text{ ⚠ } (L_2\text{-norm})$$

⚠ Note that norms must be *finite*. Hence, for the last two norms to be valid, the infinite sum/integral must converge! This restricts the set of sequences/functions on which we can operate.

Properties of Norms Induced by Inner Products

Norms induced by inner products verify the following properties:

Pythagorean Theorem

Let V be an inner product space with induced norm $\|\cdot\|$. Then, we have

$$x \perp y \implies \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

By induction, one can derive a more general form of the Pythagorean theorem:

$$\{x_k\}_{k\in\mathcal{K}} \text{ orthogonal} \implies \left\|\sum_{k\in\mathcal{K}} x_k\right\|^2 = \sum_{k\in\mathcal{K}} \|x_k\|^2.$$

Cauchy-Schwartz Inequality[a]

Let V be an inner product space with induced norm $\|\cdot\|$. Then, we have

$$|\langle x, y\rangle| \leq \|x\|\,\|y\|,$$

with equality when $y = \alpha x$ for some $\alpha \in \mathbb{F}$.

The Cauchy-Schwartz inequality can be used to define the angle between two vectors:

$$\theta_{xy} = \arccos\left(\frac{\langle x, y\rangle}{\|x\|\,\|y\|}\right).$$

[a] See [1, Exercise 2.11] for a proof.

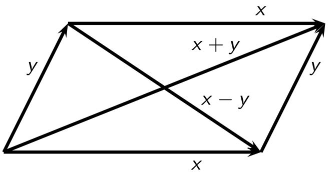
Not all Norms are Induced by Inner Products

A *necessary* and *sufficient* condition for a norm to be induced by an inner product is the following:

Proposition: (Induced Norms and Parallelogram Law)

Let $(V, \|\cdot\|)$ be some normed vector space. The norm is induced by an inner product on V iif the parallelogram law holds for any vectors $x, y \in V$:

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$



Notice that the forward direction[6] is easily obtained from properties of the inner product:

$$\begin{cases} \|x+y\|^2 = \|x\|^2 + \|y\|^2 + \langle x,y\rangle + \langle y,x\rangle \\ \|x-y\|^2 = \|x\|^2 + \|y\|^2 - \langle x,y\rangle - \langle y,x\rangle \end{cases} \Rightarrow \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

$$\|x+y\|^2 = \langle x+y, x+y\rangle$$

$$= \underbrace{\langle x, x\rangle}_{\|x\|^2} + \underbrace{\langle y, y\rangle}_{\|y\|^2} + \langle x, y\rangle + \langle y, x\rangle$$

[6] See [1, Exercise 2.11] for the backward direction.

Norms Not Induced by Inner Products

Examples of Norms Not induced by Inner Products

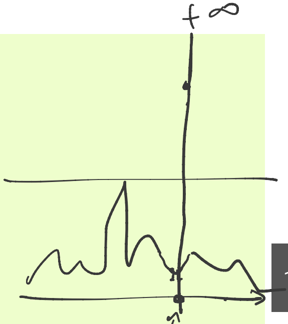
$\mathbb{C}^N$: $\|\boldsymbol{x}\|_p = \sqrt[p]{\sum_{i=1}^N |x_i|^p}$, $p \in [1,\infty)\setminus\{2\}$, $\|\boldsymbol{x}\|_\infty = \max_{i=1,\dots,N} |x_i|$ (ℓ_p-norm)

$\mathbb{C}^{N\times M}$: $\|\boldsymbol{X}\|_p = \sqrt[p]{\sum_{i=1}^N \sum_{j=1}^M |X_{i,j}|^p} = \|\operatorname{vec}(\boldsymbol{X})\|_p$, $p \in [1,\infty)\setminus\{2\}$, ($\ell_p$-norm)

$\|\boldsymbol{X}\|_{p,q} = \sup_{\|\boldsymbol{y}\|_q \leq 1} \|\boldsymbol{X}\boldsymbol{y}\|_p$ ($\ell_{p,q}$ operator norm)

$\mathbb{C}^{\mathbb{Z}}$: $\|x\|_p = \sqrt[p]{\sum_{i\in\mathbb{Z}} |x_i|^p}$, $p \in [1,\infty)\setminus\{2\}$, $\|x\|_\infty = \sup_{i\in\mathbb{Z}} |x_i|$ ⚠ (ℓ_p-norm)

$\mathbb{C}^{\mathbb{R}}$: $\|x\|_p = \sqrt[p]{\int_{\mathbb{R}} |x(t)|^p dt}$, $p \in [1,\infty)\setminus\{2\}$, $\|x\|_\infty = \underline{\text{ess}} \sup_{t\in\mathbb{R}} |x(t)|$ ⚠ (L_p-norm)



$\int f(t)\,dt$

⚠ Again, for the last two norms to be valid, the infinite sum/integral must converge and the supremum/essential supremum[7] must exist! This restricts the set of sequences/functions on which we can operate.

[7] The essential supremum is the supremum of the function *almost everywhere* –i.e. except potentially on a set of measure 0.

$\mathscr{L}^p$ Spaces

Definition: ($\mathscr{L}^p$ Space)

Let $\mathbb{K}$ denote $\mathbb{R}$ or $\mathbb{Z}$. Then $\mathscr{L}^p(\mathbb{K})$ is defined as the subspace of vectors of $\mathbb{C}^{\mathbb{K}}$ with finite L_p-norm:

$$\mathscr{L}^p(\mathbb{K}) = \left\{x \in \mathbb{C}^{\mathbb{K}} : \|x\|_p < +\infty\right\}, \qquad p \in [1, +\infty].$$

***Note:** For $\mathbb{K} = \mathbb{Z}$, the lower case notation $\ell^p(\mathbb{Z})$ is preferred.*

- $\mathscr{L}^p(\mathbb{K})$ is a normed vector space with the norm $\|\cdot\|_p$.
- $\mathscr{L}^2(\mathbb{K})$ is an inner product space with norm $\|\cdot\|_2$ *induced* by the standard inner product on $\mathbb{C}^{\mathbb{K}}$.
- Hölder's inequality generalises the Cauchy-Schwartz inequality for conjugates $p, q \in [1, +\infty)$ $(1/p + 1/q = 1)$:

$$\left|\langle x, y\rangle\right| \leq \underbrace{\|xy\|_1 \leq \|x\|_p \|y\|_q}_{\text{Hölder's inequality}} < +\infty, \quad \forall x, y \in \mathscr{L}^p(\mathbb{K}) \times \mathscr{L}^q(\mathbb{K}).$$

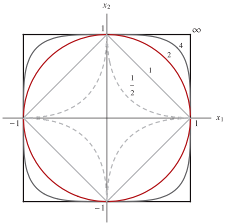


Figure: Unit balls of the ℓ_p norms in $\mathbb{R}^2$.

Convergence in Normed Vector Spaces

When working in *infinite-dimensional* vector spaces, we will often manipulate infinite summations of vectors. To make sense of such objects, we need a notion of convergence.[8] The convergence of a sequence of vectors is assessed via a *metric*, which we assume induced by a norm.

Definition: (Convergence in Normed Vector Spaces)

A sequence of vectors $(x_n)_{n\in\mathbb{N}} \in V^{\mathbb{N}}$ in a normed vector space V is said to converge to $v \in V$ if $\lim_{k\to\infty} \|v - x_k\| = 0$.

⚠ *Whether or not a sequence of vectors converges depends on the norm chosen on V!*

Convergence in Different Norms

The sequence $\left(x_k(t) = \chi_{[0,1/k]}(t)\right)_{k\in\mathbb{N}}$, converges to the null sequence $v = 0$ for the L_p-norm, when $p \in [1, +\infty)$, but does not converge for the L_∞-norm.

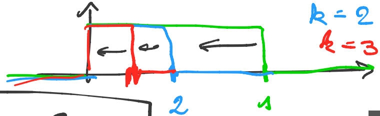
[8] See [1, Appendix 2.A.2] for a review of convergence for sequences/series of numbers or functions.

Example: Convergence in Different Norms

$x_k(t) = \chi_{[0,1/k]}(t)$, $k \in \mathbb{N} \setminus \{0\}$

$$x_k(t) = \begin{cases} 1 & \text{if } t \in [0, \frac{1}{k}] \\ 0 & \text{otherwise} \end{cases}$$

$\{x_k\}_{k \in \{1, \dots, +\infty\}} \to v = 0$



* ℓ^p-norm:

$$\lim_{k \to +\infty} \|x_k - v\|_p = \sqrt[p]{\int_0^{1/k} (\overset{1}{x_k(t)} - \overset{0}{v(t)})^p \, dt}$$

$p > 1$

$$= \sqrt[p]{\int_0^{1/k} dt} = \frac{1}{k^{1/p}} \to 0 \quad \checkmark$$

CV w.r.t ℓ_p-norm

+ ℓ^∞-norm: $\lim_{k \to +\infty} \|x_k - v\|_\infty = 1 \not\to 0$

$x_k \not\to v$ w.r.t. ∞-norm

Example: Convergence in Different Norms ✗

$x_k[n] = \frac{1}{k^\alpha}$ if $n \in \{1, 2, \cdots, k\}$, 0 otherwise, $\quad k \in \mathbb{N}, \alpha \in (0,1)$.

$\alpha \in (0,1)$

$n \in \mathbb{N}$

$$x_{k,n} = \begin{cases} 1/k^{\alpha}, & n \in \{1, \dots, k\} \\ 0 & \text{otherwise} \end{cases}$$

$$\lim_{k \to \infty} \| x_k - v \|_p = \left(\sum_{n=1}^{k} \left(\frac{1}{k^\alpha} - 0 \right)^p \right)^{1/p}$$

$$= \left(\frac{k}{k^{\alpha p}} \right)^{1/p} = k^{\frac{\alpha(p-1)}{p}}$$

$$= 0$$

$$\alpha(p-1) > 0$$

$$\Rightarrow p > 1/\alpha$$

$$\| x_k - v \|_\infty = \sup_{n \in \mathbb{N}} |x_{k,n} - 0|$$

$$= \frac{1}{k^\alpha} \to 0$$

Closed Subspace

A subspace containing the limits of all its convergent sequences is said to be closed:

Definition: (Closed Subspace)

A subspace S of a normed vector space V is called closed when it contains all limits of sequences of vectors in S. The closure $\overline{S}$ is the set of all limit points of convergent sequences in S.[9]

The closure of a set is by definition always closed. Subspaces of finite-dimensional normed vector spaces are also always closed, but this is not true in infinite dimension. In particular, the span of an infinite set of vectors *may not be closed*.

The closure of the span of an infinite set of vectors is the set of all convergent infinite linear combinations:

$$\overline{\operatorname{span}}\left(\{\varphi_k\}_{k\in\mathcal{K}}\right) = \left\{ \sum_{k\in\mathcal{K}} \alpha_k \varphi_k \,\middle|\, (\alpha_k)_{k\in\mathcal{K}} \in \mathbb{F}^{\mathcal{K}} \text{ and the sum converges} \right\}.$$

[9] We have hence in particular $S \subset \overline{S}$ since the constant sequences belong to S.

Example: Span Needs Not Be Closed

$$\ell_2(\mathbb{Z}) = \{ (x_n)_{n\in\mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}} : \sum_{n\in\mathbb{Z}} |x_n|^2 < +\infty \}$$

$$- \{ s_k \}_{k\in\mathbb{Z}} \qquad s_k = (\cdots 0, 0, 1, 0 \cdots)$$

kth position

$$\mathrm{span}\{s_k\} = \{ \sum_{j=1}^{n} a_j s_j \mid n \in \mathbb{N} \}$$

$$x_n = \frac{1}{n^2} \in \ell_2(\mathbb{Z}) \qquad v_k[n] = \begin{cases} x_k & n \in \{0, \ldots, k\} \\ 0 & \text{otherwise} \end{cases}$$

$$\lim_{k\to+\infty} \| x - v_k \|_2^2 = \sum_{n=k+1}^{+\infty} x_k^2 \Leftrightarrow 0 \longrightarrow 0$$

Example: Span Needs Not Be Closed

Completeness

Completeness of a space is the property that ensures that any sequence that intuitively ought to converge –i.e. Cauchy sequences– indeed does converge to a limit in the same space.

Definition: (Cauchy Sequence of Vectors)

A sequence of vectors $(x_n)_{n\in\mathbb{N}}$ in a normed vector space is called a Cauchy sequence if

$$\forall \varepsilon > 0, \exists K_\varepsilon > 0: \quad \|x_k - x_m\| < \varepsilon \qquad \forall k, m > K_\varepsilon.$$

Since the elements of a Cauchy sequence eventually stay arbitrarily close to each other, it makes intuitive sense that the sequence *should* converge. $\mathbb{R}$ is complete but $\mathbb{Q}$ is not:

$\mathbb{Q}$ is not complete

From the *Taylor expansion* of e^x in zero we have that: $\sum_{n=0}^{+\infty} 1/n! = e$. Since $\left(\sum_{n=0}^{N} 1/n!\right)_{N\in\mathbb{N}} \in \mathbb{Q}^{\mathbb{N}}$ and $e \notin \mathbb{Q}$, the set $\mathbb{Q}$ is not complete.

Example: $(\mathscr{C}([-1,1]), \|\cdot\|_2)$ is not Complete

$(\mathscr{C}([-1,1]), \|\cdot\|_2)$ is not complete. Indeed, we know from *Fourier analysis* that:

$$\lim_{N\to+\infty} \left\| \underbrace{\operatorname{sgn}(\sin(\pi t))}_{\text{Square wave}} - \sum_{n=0}^{N} \frac{\sin(\pi n t)}{n} \right\|_2 = 0.$$

The sequence of *continuous* functions

$$\left(\sum_{n=0}^{N} \frac{\sin(\pi n t)}{n} \right)_{N\in\mathbb{N}} \in \mathscr{C}([-1,1])^{\mathbb{N}}$$

has hence for limit the square wave which is discontinuous, showing that $\mathscr{C}([-1,1])$ is not complete w.r.t. $\|\cdot\|_2$. Note that it is however complete w.r.t. $\|\cdot\|_\infty$.[10]

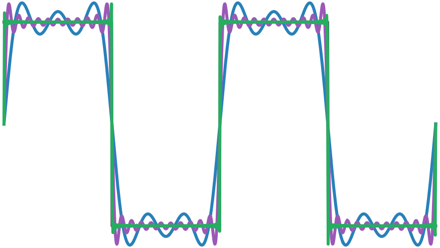


Figure: Approximation of the square wave with its truncated Fourier series $\sum_{n=0}^{N} \frac{\sin(\pi n t)}{n}$ for $N = 2, 10, 150$.

[10] This is because the limit of a uniformly convergent sequence of continuous functions is continuous.

Hilbert and Banach Spaces

Hilbert and Banach Spaces

Complete inner product vector spaces are called Hilbert spaces.[a] Complete normed vector spaces are called Banach spaces.

- Finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$ are Banach.
- All $\mathscr{L}^p$ spaces are Banach. In particular, $\mathscr{L}^2$ is Hilbert.
- $\mathscr{C}([a,b])$ is Banach with $\|\cdot\|_\infty$. It is not Hilbert w.r.t. $\|\cdot\|_2$.

[a] Hence the alternative appellation "pre-Hilbert space" sometimes used to denote a non-complete inner product space.

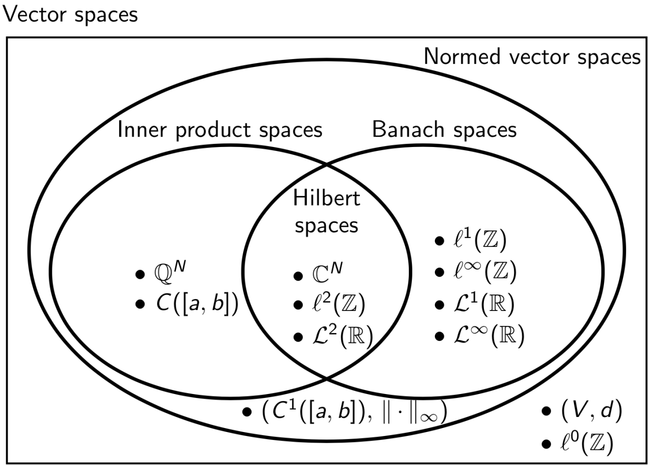


Figure: Classification of standard vector spaces.

Linear Operators

Linear operators generalise finite-dimensional matrices to abstract vector spaces:

Definition: (Linear Operator)

A map $A : H_0 \to H_1$ is a linear operator when $\forall\ x, y \in H_0, \alpha \in \mathbb{C}$ the following hold:

$$\begin{aligned} \texttt{Additivity:} \quad & A(x+y) = Ax + Ay, \\ \texttt{Scalability:} \quad & A(\alpha x) = \alpha(Ax). \end{aligned}$$

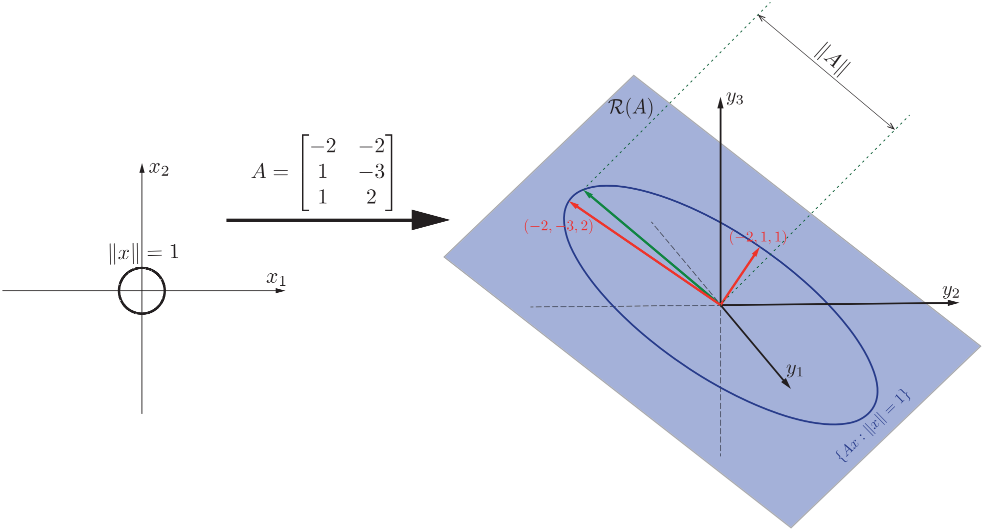
Just like in finite dimension, we can define the null space and range of an operator as, respectively:

$$\mathcal{N}(A) = \{x \in H_0 |\ Ax = 0\} = A^{-1}(\{0\}), \qquad \mathscr{R}(A) = \{Ax \in H_1 |\ x \in H_0\} = A(H_0).$$

Definition: (Operator Norm & Bounded Operator)

The operator norm of A, denoted by $\|A\|$, is defined as $\|A\| = \sup_{\|x\|=1} \|Ax\|$. A linear operator is called bounded when its operator norm is finite.

Linear Operators: illustration



Adjoint of an Operator

The adjoint generalises the Hermitian transpose for finite-dimensional matrices.

Definition: (Adjoint)

The linear operator $A^*: H_1 \to H_0$ is called the adjoint of the linear operator $A: H_0 \to H_1$ when

$$\langle Ax, y\rangle_{H_1} = \boxed{\langle x, A^* y\rangle_{H_0}}, \qquad \text{for every } x \text{ in } H_0 \text{ and } y \text{ in } H_1.$$

When $A = A^*$, the operator A is called self-adjoint or Hermitian.

The angle between Ax and y in H_1 is the *same* as the angle between A^*y and x in H_0.

$\alpha^* \to \overline{\alpha}$

$$\alpha : \begin{cases} \mathbb{C} \to \mathbb{C} \\ x \mapsto \alpha x \end{cases}$$

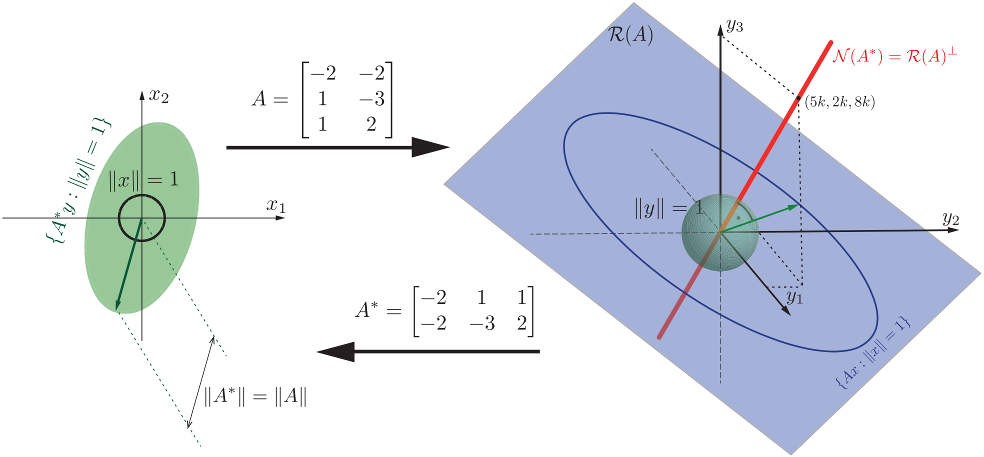
Examples of Adjoints

`Adjoint of a scalar` $\alpha \in \mathbb{C}$: We have $\boxed{\langle \alpha x, y\rangle} = \alpha\langle x, y\rangle = \langle x, \overline{\alpha} y\rangle \ \forall x, y \in H$. Hence, $\alpha^* = \overline{\alpha}$.

`Adjoint of a matrix` $\boldsymbol{A} \in \mathbb{C}^{N\times M}$: $\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{m=1}^{M}\left(\sum_{n=1}^{N} A_{mn}x_n\right)\overline{y_m} = \sum_{n=1}^{N} x_n \overline{\left(\sum_{m=1}^{M} \overline{A_{mn}} y_m\right)} = \boxed{\langle \boldsymbol{x}, \boldsymbol{A}^H \boldsymbol{y}\rangle}$ for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{C}^N \times \mathbb{C}^M$. Hence we get $\boldsymbol{A}^* = \boldsymbol{A}^H$.

$$\sum_{m=1}^{M} \overline{A}_{mn}\, y_m \qquad A^H = A^*$$

Adjoint Operator: illustration



Example: Local Averaging and its Adjoint

$$(Ax)_n = \int_{n-1/2}^{n+1/2} x(t)\,dt, \quad n \in \mathbb{Z},\ x \in \mathcal{L}^2(\mathbb{R}).$$

$$\int x(t) z(t)\,dt$$

$$A : \begin{cases} \mathcal{L}^2(\mathbb{R}) \to \mathbb{C}^{\mathbb{Z}} \\ x \mapsto y_n = \int_{n-1/2}^{n+1/2} x(t)\,dt \end{cases}$$

Linear operator

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2$$

$$\langle Ax, y \rangle_{\mathbb{R}^{\mathbb{Z}}} \overset{?}{=} \langle x, A^* y \rangle_{\mathcal{L}^2(\mathbb{R})}$$

$\in \mathcal{L}^2(\mathbb{R})$ $\quad \mathbb{R}^{\mathbb{Z}}$

$$A^* : \mathbb{C}^{\mathbb{Z}} \to \mathcal{L}^2(\mathbb{R})$$

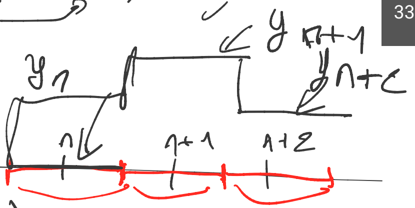
$$= \sum_{n \in \mathbb{Z}} (Ax)_n\, y_n$$

$$= \sum_{n \in \mathbb{Z}} \int_{n-1/2}^{n+1/2} x(t)\,dt\; y_n$$

$\mathbb{R}$

$$= \int_{\mathbb{R}} x(t) \left(\sum_{n \in \mathbb{Z}} y_n\, \chi_{[n-1/2,\, n+1/2]}(t) \right) dt$$

$\mathbb{R}$



Example: Local Averaging and its Adjoint

$$= \langle x, \underline{A^*} y \rangle$$

$$\underbrace{(A^* y)(t) = \sum_{n \in \mathbb{Z}} y_n \chi_{[n-1/2,\, n+1/2]}(t)}$$

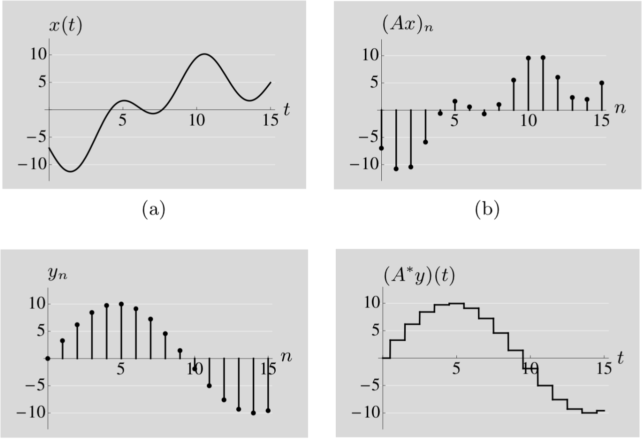
Example: Local Averaging and its Adjoint

(a) We start with a function x in $\mathscr{L}^2(\mathbb{R})$.

(b) The local averaging operator A gives a sequence in $\ell^2(\mathbb{Z})$.

(c) y is an arbitrary sequence in $\ell^2(\mathbb{Z})$.

(d) The adjoint A^* is a linear operator from $\ell^2(\mathbb{Z})$ to $\mathscr{L}^2(\mathbb{R})$ that uniquely preserves geometry in that $\langle Ax, y\rangle_{\ell^2(\mathbb{Z})} = \langle x, A^* y\rangle_{\mathscr{L}^2(\mathbb{R})}$. The adjoint of local averaging is to form a piecewise-constant function.



(a) (b) (c) (d)

Properties of the Adjoint

Theorem: (Adjoint properties)

Let $A : H_0 \longrightarrow H_1$ be a bounded linear operator. Then,

1. A^* exists and is unique,
2. $(A^*)^* = A$,
3. AA^* and A^*A are self-adjoint,
4. $\|A^*\| = \|A\|$,
5. If A is invertible, $(A^{-1})^* = (A^*)^{-1}$,
6. $B : H_0 \longrightarrow H_1$ bounded, $(A+B)^* = A^* + B^*$,
7. $B : H_1 \longrightarrow H_2$ bounded, $(BA)^* = A^* B^*$,
8. $\mathscr{R}(A)^\perp = \mathscr{N}(A^*)$ and $\overline{\mathscr{R}(A)} = \mathscr{N}(A^*)^\perp$.[11]

[11] A and A^* can be interchanged in these two relations. For the second relation, we must take the closure of the range since in infinite dimension the range is not necessarily a closed subspace while the nullspace is always a closed subspace. Note that this will rarely be a concern in practice since most of the operators we will work with will have closed ranges.

Proof Sketch: points 2, 3

2. $(A^*)^* = A$ $\quad A : H_0 \to H_1 \; / \; A^* : H_1 \to H_0$

3. $(AA^*)^* = AA^*$

$$\langle AA^* y, z \rangle_{H_1} = \langle A^* y, A^* z \rangle$$
$$= \langle y, AA^* z \rangle$$

Proof Sketch: points 5, 6 ,7

$x, z \in H_0$

5: $(A^{-1})^* = (A^*)^{-1}$

$\langle x, z \rangle = \langle AA^{-1}x, z \rangle$

$= \langle A^{-1}x, A^* z \rangle$

$= \langle x, (A^{-1})^* A^* z \rangle$

$\Rightarrow \langle x, (I - (A^{-1})^* A^*) z \rangle = 0$

$\Rightarrow (I - (A^{-1})^* A^*) z = 0 \quad \forall z \in H_0$

$\Downarrow$

$I = (A^{-1})^* A^*$

$(A^{-1})^* = (A^*)^{-1}$

6: $(A+B)^* = A^* + B^*$

$\langle (A+B)x, y \rangle$

$= \langle Ax, y \rangle + \langle Bx, y \rangle$

$= \langle x, A^* y \rangle + \langle x, B^* y \rangle$

$= \langle x, (A^* + B^*) y \rangle \Rightarrow (A+B)^* = A^* + B^*$

7. $\langle ABx, y \rangle = \langle Bx, A^* y \rangle = \langle x, B^* A^* y \rangle$

$(AB)^* = B^* A^*$

Proof Sketch: point 8

$$\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$$

α $\quad \mathcal{R}(A)^\perp \subseteq \mathcal{N}(A^*)$:

$$\forall x \in \mathcal{R}(A)^\perp \Rightarrow x \in \mathcal{N}(A^*)$$

$$x \perp s \in \mathcal{R}(A) = \{Ay, y \in$$

$$\langle x, Ay \rangle = 0 \quad \forall y$$

$$\langle A^* x, y \rangle = 0 \quad \forall y \in H_0 \Rightarrow A^* x \perp H_0$$

$$\Rightarrow A^* x = 0$$

$$\in \mathcal{N}(A^*) = \{x \in H_1 : A^* x$$

$$S \subset H : \quad S^\perp = \{x \in H : x \perp S\}$$

$$x \in S^\perp \qquad y \in S : \langle x, y \rangle = 0$$

Generalised Inverse & Pseudoinverse

Definition: (Generalised Inverse & Pseudoinverse)

Let $A : H_0 \to H_1$ be a linear operator. Consider a bounded linear operator $A^\ddagger : H_1 \to H_0$ as well as the Penrose conditions:

1. $AA^\ddagger A = A$,
2. $A^\ddagger AA^\ddagger = A^\ddagger$,
3. $(AA^\ddagger)^* = AA^\ddagger$,
4. $(A^\ddagger A)^* = A^\ddagger A$.

Then, $A^\ddagger$ is called: a generalised inverse if it satisfies 1, a reflexive generalised inverse if it satisfies 1 & 2, the pseudoinverse[12] denoted by $A^\dagger$ if it satisfies 1 to 4.

Generalised inverses are not necessarily unique. The pseudoinverse of an operator A with closed range exists and is unique [2, Section 2].

[12]The pseudoinverse is also sometimes called the Moore-Penrose inverse after the pioneering works by Moore and Penrose.

Inverse

Definition: (Inverse)

A linear operator $A: H_0 \to H_1$ is said invertible if there exists a bounded linear operator $B: H_1 \to H_0$ such that:

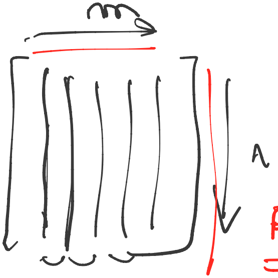
1. $BAx = x, \quad \forall x \in H_0,$
2. $ABy = y, \quad \forall y \in H_1.$

In which case, B is unique and is called the inverse of A, denoted by A^{-1}. Moreover, B is called a left inverse if it satisfies 1 only, and a right inverse if it satisfies 2 only.

It is easy to see that when A is invertible, then the pseudoinverse and the inverse coincide. A left (respectively right) inverse is moreover also a generalised inverse.

Example: Pseudoinverse of Matrices with Full Column Ranks

$A \in \mathbb{R}^{n \times m}$

$A =$ 

$A^+A = (A^HA)^{-1}A^HA = I$

$\dagger = H$

Pseudoinverse = left inverse

$\rightarrow A^HA$ 

$(ABC)^H = C^HB^HA^H$

$A^+ = (A^HA)^{-1}A^H$

1. $AA^+A = A\underbrace{(A^HA)^{-1}A^HA}_{=I}$

2. $A^+AA^+ = (A^HA)^{-1}\underbrace{A^HA(A^HA)^{-1}}A^H = (A^HA)^{-1}\overset{I}{}A^H = A^+$

3. $(AA^+)^* = AA^+$

$\left(A(A^HA)^{-1}A^H\right)^H = A(A^HA)^{-1}A^H = AA^+$

Example: Pseudoinverse of Matrices with Full Row Ranks

$$\begin{array}{c} m \\ A = \begin{bmatrix} \text{———} \\ \text{———} \\ \text{———} \end{bmatrix} \end{array} \leftarrow n \qquad (AA^H)^{-1}$$

$$A^+ = A^H(AA^H)^{-1}$$

pseudoinverse

and

a right inverse

$$AA^+ = AA^H(AA^H)^{-1} = I$$

Unitary Operator

Unitary operators are operators that preserve the geometry (angles and distances) when mapping one Hilbert space to another.

Definition: (Unitary Operator)

A bounded linear operator $A : H_0 \longrightarrow H_1$ is unitary when:

1. A is invertible,
2. A preserves inner products: $\langle Ax, Ay \rangle_{H_1} = \langle x, y \rangle_{H_0}$, for every $x, y \in H_0$.

Preservation of inner products implies preservation of norms since $\|Ax\|^2 = \langle Ax, Ax \rangle = \langle x, x \rangle = \|x\|^2 \ \forall x \in H_0$.

Theorem: (Characterisation of Unitary Operators)

A bounded linear operator A is unitary iff its inverse is its adjoint: $A^{-1} = A^*$.

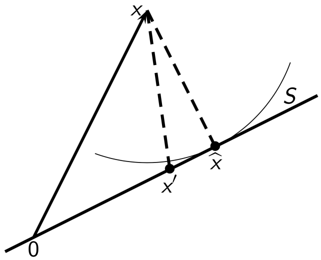
Best Approximation Problem

Let S be some subspace of a Hilbert space H. For a given $x \in H$, we define the best approximation problem[13] as:

$$\hat{x} = \arg\min_{s \in S} \|x - s\|. \tag{2}$$

Equation (2) tries to find an element of the subspace S that best approximates –understand that is *the closest to*– the vector $x \in H$.

Our intuition from Euclidean geometry tells us that the solution to (2) is unique and such that the residual $x - \hat{x} \perp S$. We now generalise this result to abstract Hilbert spaces.



[13] When $\|\cdot\|$ is the canonical 2 norm, the best approximation problem is also called the least-squares problem.

Projection Theorem

Theorem: (Projection Theorem)

Let S be a closed subspace of Hilbert space H and let $x \in H$.

1. Existence: There exists $\hat{x} \in S$ such that $\|x - \hat{x}\| \leq \|x - s\|$ for all $s \in S$,
2. Orthogonality: $x - \hat{x} \perp S$ is necessary and sufficient to determine $\hat{x}$,
3. Uniqueness: $\hat{x}$ is unique,
4. Linearity: $\hat{x} = Px$ where P is a linear operator,
5. Idempotency: $P(Px) = Px$ for all $x \in H$,
6. Self-adjointness: $P = P^*$.

The effect of the operator P that arises in the projection theorem is to move the input vector x in a direction orthogonal to the subspace S until S is reached at $\hat{x}$. We will see that P has the defining properties of an orthogonal projection operator.

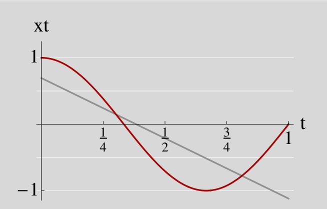
Example: Polynomial Approximation

Example: Polynomial Approximation

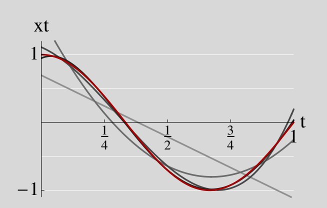
We want to approximate $x(t) = \cos(3\pi t/2) \in \mathcal{L}^2([0,1])$ by a degree 1 polynomial $s(t) = a_0 + a_1 t \in S = \mathrm{span}\{1, t\}$.
Using point 2 of the projection theorem we get that the approximand $\hat{x} \in S$ verifies:

$$\begin{cases} \langle x - \hat{x}, 1\rangle & = \displaystyle\int_0^1 \cos(3\pi t/2) - a_0 - a_1 t \, dt = -\frac{2}{3\pi} - a_0 - \frac{a_1}{2} = 0, \\ \langle x - \hat{x}, t\rangle & = \displaystyle\int_0^1 \cos(3\pi t/2)\, t - a_0 t - a_1 t^2 \, dt = -\frac{4+6\pi}{9\pi^2} - \frac{a_0}{2} - \frac{a_1}{3} = 0, \end{cases}$$

$$\Longleftrightarrow \begin{cases} a_0 & = \dfrac{8+4\pi}{3\pi^2}, \\ a_1 & = -\dfrac{16+12\pi}{3\pi^2}, \end{cases}$$ and hence $\hat{x}(t) = \frac{8+4\pi}{3\pi^2} - \frac{16+12\pi}{3\pi^2} t.$



(a) Approximation of x(t) by a degree 1 polynomial.



(b) Approximation of x(t) by a degree $K \in \{1,2,3,4\}$ polynomial.

Projection Operator

The operator P that arises from solving the best approximation problem is an orthogonal projection operator:

Definition: (Projection Operator)

- An idempotent[14] operator P is an operator such that $P^2 = P$.
- A projection operator is a bounded linear operator that is idempotent.
- An orthogonal projection operator is a projection operator that is self-adjoint.
- An oblique projection operator is a projection operator that is not self-adjoint.

Theorem: (Characterisation of Projection Operators)

P is a bounded orthogonal projection operator $\iff \langle x - Px, Py \rangle = 0,\ \forall x, y \in H$.

[14] An operator is idempotent when applying it twice is no different than applying it once.

Proof Sketch: Characterisation of Projection Operators

Projection via Generalised Inverses

Theorem: (Projection via Generalised Inverses)

Consider a bounded linear operator $A : H_0 \to H_1$ with closed range. Denote moreover by $A^\ddagger : H_1 \to H_0$ some reflexive generalised inverse of A and by $A^\dagger$ its pseudoinverse.[15] Then,

1. $AA^\ddagger : H_1 \to H_1$ is a projection operator onto $\mathscr{R}(A)$,
2. $A^\ddagger A : H_0 \to H_0$ is a projection operator onto $\mathscr{R}\left(A^\ddagger\right)$,
3. $AA^\dagger : H_1 \to H_1$ is an orthogonal projection operator onto $\mathscr{R}(A)$,
4. $A^\dagger A : H_0 \to H_0$ is an orthogonal projection operator onto $\mathscr{R}(A^*)$.

We can deduce [1, Theorems 2.29 and 2.30] as corollaries of the above result (check it!):

- If $A^\ddagger$ is a left inverse to A, then $AA^\ddagger$ is a projection operator onto $\mathscr{R}(A)$,
- If A^*A is invertible, then $A(A^*A)^{-1}A^*$ is an orthogonal projection operator onto $\mathscr{R}(A)$,
- If AA^* is invertible, then $A^*(AA^*)^{-1}A$ is a orthogonal projection operator onto $\mathscr{R}(A^*)$.

[15] Which exists since A has closed range.

Proof Sketch: Points 1 & 3

Example: Projection onto a Subspace of $\mathbb{R}^3$

Let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Since B is a left inverse of A, we know from the previous theorem that $P = AB$ is a projection operator:

$$P = AB = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix}.$$

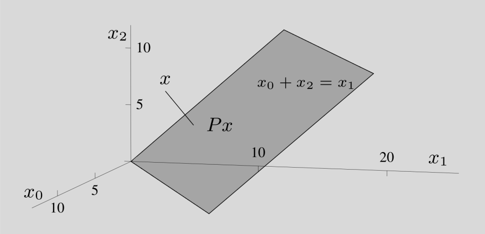


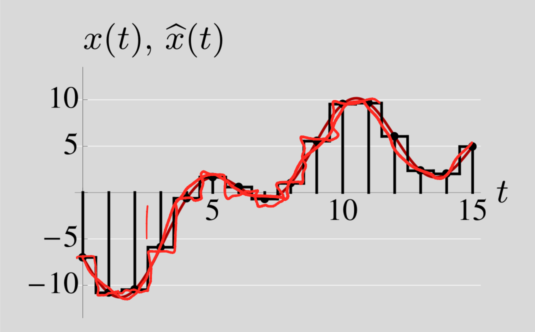
Figure: The two-dimensional range of the oblique projection operator P is the plane $x_0 + x_2 = x_1$.

It is easy to verify that $P^2 = P$. The two-dimensional range of this projection operator is the *set of three-tuples with middle component equal to the sum of the first and last*.
Note that $P \neq P^*$, so the projection is oblique.

Example: Projection onto Piecewise-Constant Functions

Example: Projection onto Piecewise-Constant Functions

$$\int_{n-1/2}^{n+1/2} x(t)\,dt$$



$$AA^* = I$$

$$A^* A x = \hat{x}$$

Figure: Given a function $x \in \mathscr{L}^2(\mathbb{R})$, the function in the subspace of piecewise-constant functions $A^* \ell^2(\mathbb{Z})$ that is closest to x in $\mathscr{L}_2$ norm is the one obtained by replacing $x(t)$, $t \in [n-1/2, n+1/2)$, by its local average $\int_{n-1/2}^{n+1/2} x(t)\,dt$.

Projections and Direct Sums

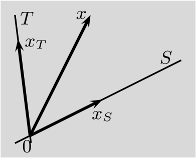
Projection operators generate direct sum decompositions.

Definition: (Direct Sum)

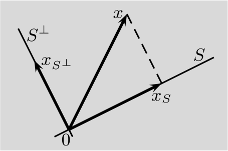
A vector space V is a direct sum decomposition of subspace S and T (written $V = S \oplus T$) when any nonzero vector $x \in V$ can be written uniquely as $x = x_S + x_T$, where $x_S \in S$ and $x_T \in T$.

Theorem: (Direct-sum Decomposition from Projection)

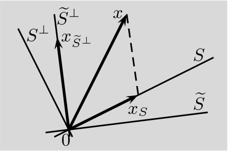
- A projection P on H generates direct sum decomposition $H = S \oplus T$, where $S = \mathcal{R}(P)$ and $T = \mathcal{N}(P)$.
- If S, T are closed subspaces s.t. $H = S \oplus T$, then there exists a projection P on H s.t. $S = \mathcal{R}(P)$ and $T = \mathcal{N}(P)$.



(a) Decomposition



(b) Orthogonal projection



(c) Oblique projection

References I

[1] Martin Vetterli, Jelena Kovačević, and Vivek K Goyal.
***Foundations of signal processing*.**
Cambridge University Press, 2014.

[2] Ole Christensen.
Operators with closed range, pseudo-inverses, and perturbation of frames for a subspace.
Canadian Mathematical Bulletin, 42(1):37–45, 1999.