AUDIOVISUAL COMMUNICATIONS LCAV

Mathematical Foundations of Signal Processing

Hilbert Spaces and Projection Operators

Benjamín Haro Béjar
Mihailo Kolundžija
Reza Parhizkar
Martin Vetterli

September 15, 2014

Outline

1. Spaces
   - Vector spaces
   - Hilbert spaces
2. Operators
   - Linear operators
   - Projection operators
3. Summary

Goal:

- Establish the basics in a Hilbert space setup through geometric intuition

Readings:

- Chapter 2, "From Euclid to Hilbert", of *Foundations of Signal Processing*, Sections 2.1 to 2.4 (in particular 2.3.3 and 2.4)

Vector Spaces

For a vector space, we need:

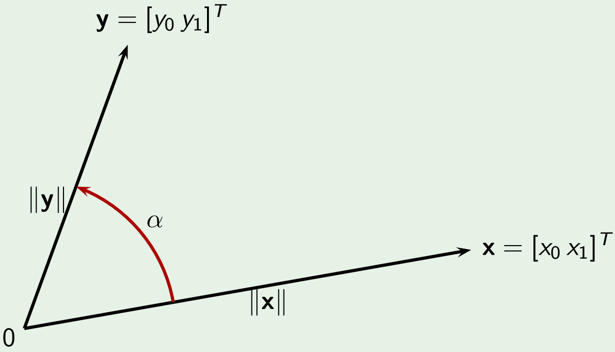
- A set of vectors V
  - These can be vectors in $\mathbb{R}^N$, functions, etc.
  - Think of geomety in $\mathbb{R}^2$ or $\mathbb{R}^3$, we will use pictures!
- A field of scalars F
  - Real or complex numbers
- Vector addition $+$
- Scalar multiplication $\cdot$

Easy case: N finite, linear algebra, matrices

Beware: N goes to infinity... convergence!

Vector spaces

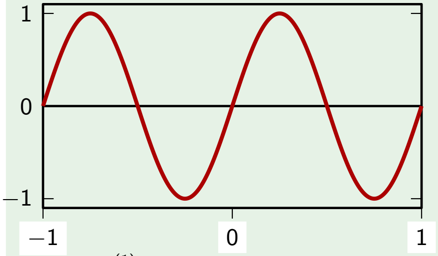
Vectors in $\mathbb{R}^2$



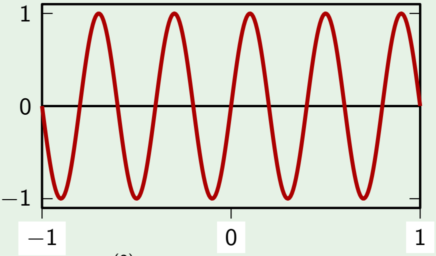
Vector spaces

Vectors can be very general objects!

Example: space of square-integrable functions over $[-1, 1]$: $\mathcal{L}^2([-1, 1])$



$\mathbf{x}^{(1)} = \sin(f_1\, t), \quad f_1 = 2\pi$



$\mathbf{x}^{(2)} = \sin(f_2\, t), \quad f_2 = 5\pi$

$$\langle \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \rangle = \int_{-1}^{1} \sin(f_1\, t) \sin(f_2\, t) dt$$

Vector spaces

Axioms

- A vector space V is defined over a field $\mathbb{F}$ (think $\mathbb{R}$ or $\mathbb{C}$) as a set with two operations
  - Vector addition: $V \times V \to V$
  - Scalar multiplication: $\mathbb{F} \times V \to V$

That satisfies the following axioms

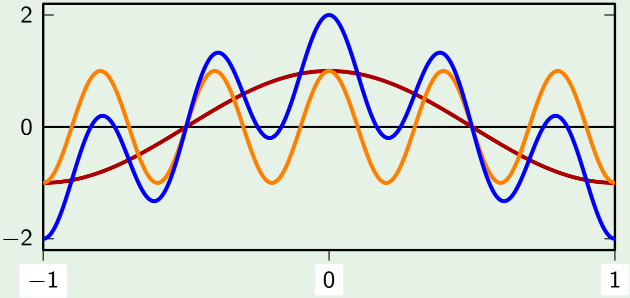
1. $x + y = y + x$
2. $(x + y) + z = x + (y + z)$
3. $\exists\, 0 \in V$ s.t. $x + 0 = x$ for all $x \in V$
4. $\alpha(x + y) = \alpha x + \alpha y$
5. $(\alpha + \beta)x = \alpha x + \beta x$
6. $(\alpha\beta)x = \alpha(\beta x)$
7. $0x = 0$ and $1x = x$

Vector spaces

Key notions

- Subspace
  - $S \subseteq V$ is a subspace when it is closed under vector addition and scalar multiplication:
    - For all x, y in $S, x + y$ is in S
    - For all x in S, α in $\mathbb{C}$ (or $\mathbb{R}$), αx is in S

Subspace of symmetric functions over $\mathcal{L}^2[-1,1]$



$\mathbf{x} = \cos(\pi t)$, $\mathbf{y} = \cos(5\pi t) \Rightarrow \mathbf{x} + \mathbf{y}$, symmetric

Vector spaces

Key notions

- Span
  - S: set of vectors (could be infinite)
  - $span(S)$ = set of all finite linear combinations of vectors in S

    $$span(S) = \left\{ \sum_{k=0}^{N-1} \alpha_k \varphi_k \mid \alpha_k \in \mathbb{C} \text{ (or } \mathbb{R}), \varphi_k \in S \text{ and } N \in \mathbb{N} \right\}$$

  - $span(S)$ is always a subspace

Vector spaces

Key notions

- Linear independence
  - $S = \{\varphi_k\}_{k=0}^{N-1}$ is linerarly independent when:

$$\text{If } \sum_{k=0}^{N-1} \alpha_k \varphi_k = 0 \text{ then } \alpha_k = 0 \text{ for all } k$$

  - If S is infinite, we need every finite subset to be linearly independent
- Dimension
  - $\text{Dim}(V) = N$ if V contains a linearly independent set with N vectors and every set with $N+1$ or more vectors is linearly dependent
  - V is infinite dimensional if no such finite N exists

Inner products

Definition (Inner product)

- Formalize the geometric notions of orientation and orthogonality
- Measure similarity between vectors
- An inner product for V is a function $\langle \cdot, \cdot \rangle \; : \; V \times V \to \mathbb{C}$ satisfying
  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 : $\langle x, y \rangle^* = \langle y, x \rangle$
  4. Positive definiteness : $\langle x, x \rangle \geq 0$; $\langle x, x \rangle = 0$ iff $x = 0$

- Note: $\langle x, \alpha y \rangle = \alpha^* \langle x, y \rangle$

Inner products

Examples

- On $\mathbb{C}^N$: $\langle x, y \rangle \;=\; \sum_{n=0}^{N-1} x_n y_n^* = y^* x$

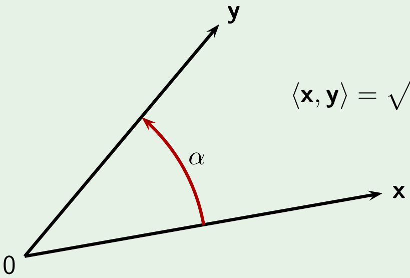
- On $\mathbb{C}^{\mathbb{Z}}$: $\langle x, y \rangle \;=\; \sum_{n \in \mathbb{Z}} x_n y_n^* = y^* x$

- On $\mathbb{C}^{\mathbb{R}}$: $\langle x, y \rangle \;=\; \int_{-\infty}^{\infty} x(t) y^*(t) \, \mathrm{d}t$

Inner products

Inner product in $\mathbb{R}^2$

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_0 y_0 + x_1 y_1$$

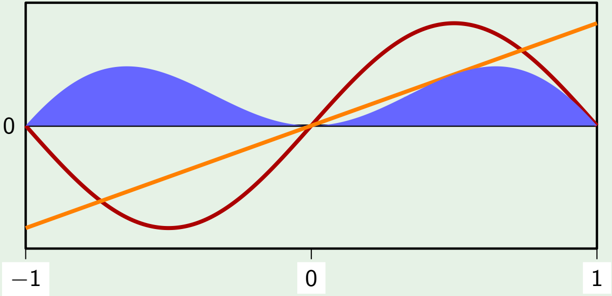


$$\langle \mathbf{x}, \mathbf{y} \rangle = \sqrt{(x_0^2 + x_1^2)(y_0^2 + y_1^2)}\,\cos\alpha$$

Inner products

Inner product in $\mathcal{L}^2[-1,1]$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \int_{-1}^{1} x(t)y(t)dt = \int_{-1}^{1} t\sin(\pi t)dt$$



$\mathbf{x} = \sin(\pi t)$, $\mathbf{y} = t$, $\langle \mathbf{x}, \mathbf{y} \rangle = 2/\pi \approx 0.6367$

Orthogonality

Let $S = \{\varphi_i\}_{i \in \mathcal{I}}$ be a set of vectors

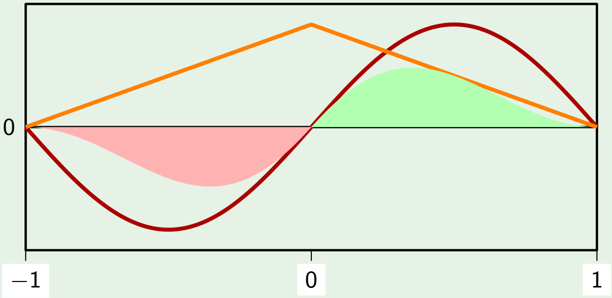
Definition (Orthogonality)

- x and y are orthogonal when $\langle x, y \rangle = 0$ written $x \perp y$
- S is orthogonal when for all $x, y \in S,\ x \neq y$ we have $x \perp y$
- S is orthonormal when it is orthogonal and for all $x \in S,\ \langle x, x \rangle = 1$
- x is orthogonal to S when $x \perp s$ for all $s \in S$, written $x \perp S$
- S_0 and S_1 are orthogonal when every $s_0 \in S_0$ is orthogonal to S_1, written $S_0 \perp S_1$

Orthogonality

Inner product in $L_2[-1, 1]$

$\mathbf{x}, \mathbf{y}$ from orthogonal subspaces

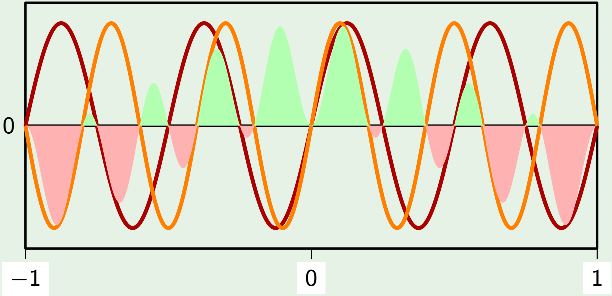


$\mathbf{x} = \sin(\pi t)$, $\mathbf{y} = 1 - |t|$; $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

Orthogonality

Inner product in $L_2[-1,1]$

$\mathbf{x}, \mathbf{y}$ from orthogonal subspaces



$\mathbf{x} = \sin(4\pi t)$, $\mathbf{y} = \sin(5\pi t)$, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

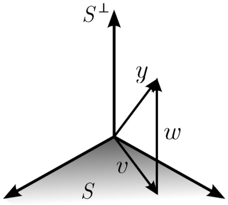
Orthogonal complement

- If S is a subspace of V, the orthogonal complement of S (in V) is the set

$$S^{\perp} \ = \ \{x \in V, \ x \perp S\}$$

- If V is closed (contains all limits) then given $y \in V$, there exists $v \in S, \ w \in S^{\perp}$ s.t.

$$y \ = \ v + w \quad , \quad V \ = \ S \oplus S^{\perp}$$



Norm

Definition (Norm)

- Measure length, size of vectors
- A norm on V is a function $\|\cdot\| \ : \ V \to \mathbb{R}$ satisfying

  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 $y = \alpha x$

- Note: We use $\|\|$ for the 2-norm. Other norms will be specified as well explicitly

Norms

Examples

- On $\mathbb{C}^N$: $\|x\| \;=\; \sqrt{\langle x, x\rangle} \;=\; \left(\sum_{n=0}^{N-1} |x_n|^2\right)^{1/2}$

- On $\mathbb{C}^\mathbb{Z}$: $\|x\| \;=\; \sqrt{\langle x, x\rangle} \;=\; \left(\sum_{n\in\mathbb{Z}} |x_n|^2\right)^{1/2}$

- On $\mathbb{C}^\mathbb{R}$: $\|x\| \;=\; \sqrt{\langle x, x\rangle} \;=\; \left(\int_{-\infty}^{\infty} |x(t)|^2 \, \mathrm{d}t\right)^{1/2}$

Distances, norms and inner products

- A norm "induces" a distance

$$d(x, y) \; = \; \|x - y\|$$

- An inner product induces a norm

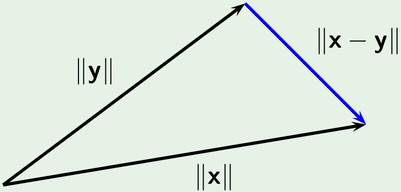
$$\|x\| \; = \; \sqrt{\langle x, x\rangle}$$

- Not all norms are induced by an inner product

Distances, norms and inner products

Norm and distance in $\mathbb{R}^2$

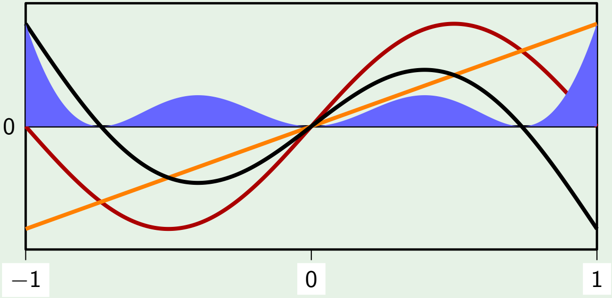
$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_0^2 + x_1^2}$$
$$\|\mathbf{y}\| = \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle} = \sqrt{y_0^2 + y_1^2}$$
$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_0 - y_0)^2 + (x_1 - y_1)^2}$$



Distances, norms and inner products

Norm and distance in $\mathcal{L}^2[-1,1]$

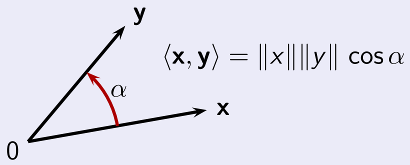
$$\|\mathbf{x}-\mathbf{y}\|^2 = \int_{-1}^{1} |x(t)-y(t)|^2 dt \text{ (MSE)}$$



$\mathbf{x} = \sin(\pi t)$; $\mathbf{y} = t$; $\mathbf{x}-\mathbf{y}$; $\|\mathbf{x}-\mathbf{y}\| = \sqrt{5/3 - 4/\pi} \approx 0.6272$

Norms induced by inner products

Properties



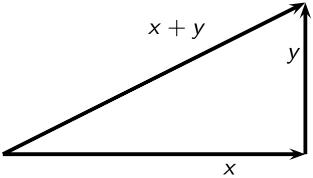
- Cauchy-Schwarz inequality

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

Norms induced by inner products

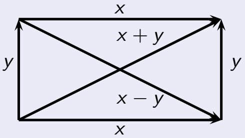
Properties

- Pythagorean theorem
  - $x \perp y \quad \Rightarrow \quad \|x+y\|^2 = \|x\|^2 + \|y\|^2$
  - $\{x_k\}_{k\in K}$ orthogonal $\quad \Rightarrow \quad \left\|\sum_{k\in K} x_k\right\|^2 = \sum_{k\in K} \|x_k\|^2$



Norms induced by inner products

Properties

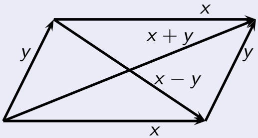


From Pythagorean theorem:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2; \qquad \|x-y\|^2 = \|x\|^2 + \|y\|^2$$

- Parallelogram Law

$$\|x+y\|^2 + \|x-y\|^2 \;=\; 2(\|x\|^2 + \|y\|^2)$$



Normed vector spaces: Standard spaces

- $\mathbb{C}^N \;:\; \langle x, y \rangle \;=\; \sum_{n=0}^{N-1} x_n\, y_n^*, \qquad \|x\| \;=\; \left(\sum_{n=0}^{N-1} |x_n|^2 \right)^{1/2}$
- $\ell^2(\mathbb{Z})$: square-summable sequences ("finite energy sequences")

$$\langle x, y \rangle \;=\; \sum_{n \in Z} x_n\, y_n^*, \qquad \|x\| \;=\; \left(\sum_{n \in \mathbb{Z}} |x_n|^2 \right)^{1/2}$$

- $\mathcal{L}^2(\mathbb{R})$: square-integrable functions ("finite energy functions")

$$\langle x, y \rangle \;=\; \int_{-\infty}^{\infty} x(t) y^*(t)\, dt, \qquad \|x\| \;=\; \left(\int_{-\infty}^{\infty} |x(t)|^2\, dt \right)^{1/2}$$

Normed vector spaces: Standard spaces

- $\mathbb{C}^N$: The p norm : $\|x\|_p = \left(\sum_{n=0}^{N-1} |x_n|^p \right)^{1/p}$, for $p \in [1, \infty)$
- $\ell^p(\mathbb{Z})$ spaces : $\|x\|_p = \left(\sum_{n \in \mathbb{Z}} |x_n|^p \right)^{1/p}$, for $p \in [1, \infty)$
- Extend p norm to ℓ^∞ norm as $\|x\|_\infty = \sup_{n \in \mathbb{Z}} |x_n|$
- $x \in \ell^p(\mathbb{Z})$ iff $\|x\|_p < \infty$
- $p = 2$: the only ℓ^p norm induced by an inner product

Normed vector spaces: Standard spaces

- $\mathcal{L}^p(\mathbb{R})$ spaces : $\|x\|_p \;=\; \left(\int_{-\infty}^{\infty} |x(t)|^p \,\mathrm{d}t\right)^{1/p}$

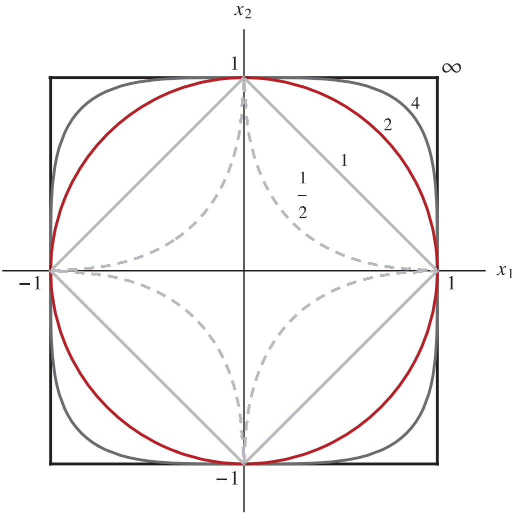
- Extend to $p = \infty$: $\mathcal{L}^\infty$ norm $\|x\|_\infty \;=\; \operatorname*{ess\,sup}_{t\in\mathbb{R}} |x(t)|$

- $x \in \mathcal{L}^p(\mathbb{R})$ iff $\|x\|_p < \infty$

- $p = 2$: the only $\mathcal{L}^p$ norm induced by an inner product

The world looks different using different norms!

Unit balls in different norms: quasinorm $\ell_{1/2}$, norms $\ell_1, \ell_2, \ell_4, \ell_\infty$



Solution of linear systems using different norms

- Consider an under-determined system of equations

$$\mathbf{x} = \mathbf{A}\alpha$$

  where $\mathbf{x}$ is $N \times 1$, $\mathbf{A}$ is $N \times M$, α is $M \times 1$ and $N < M$.
- Expansion with respect to an overcomplete set of vectors is not unique.
- Example:

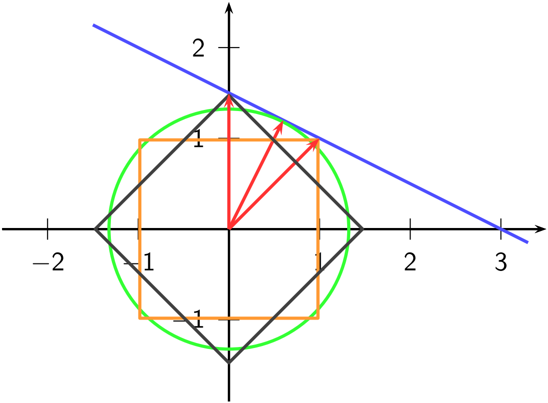
$$x = \frac{1}{5} \cdot [1\ 2] \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}$$

$$\alpha' = \alpha + \alpha^{\perp} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} x + \begin{bmatrix} 2 \\ -1 \end{bmatrix} \gamma,$$

  This is a line with slope $-1/2$ in the $[\alpha_0, \alpha_1]$ plane.

Solution of linear systems using different norms

Different norm minimizations $\|\alpha\|_p$, $p \in \{0, 1, 2\}$ give different solutions (Ex: $x = 3/5$)



and one of them is sparse!

Space of random variables

- Random variables X with finite second moment
$$\mathbb{E}[|X|^2] < \infty$$
- Inner product and norm
$$\langle X, Y \rangle = \mathbb{E}[XY^*]$$
$$\|X\| = \sqrt{\mathbb{E}[|X|^2]}$$
- Apply all the abstract theorems to random variables.

$C^p([a, b])$ Spaces

- $C([a, b])$: inner product space of complex, continuous functions over interval $[a, b]$
- $C^p([a, b])$: inner product space of complex, continuous functions with p-continuous derivatives over interval $[a, b]$
- Usual inner product, usual norm

$$\langle x, y \rangle = \int_a^b x(t) y^*(t) \, \mathrm{d}t, \quad \|x\| = \left(\int_a^b |x(t)|^2 \, \mathrm{d}t \right)^{1/2}$$

- **Example**: set of polynomial functions over an interval forms a subspace of $C^p([a, b])$, for any a, b in $\mathbb{R}$ and p in $\mathbb{N}$.
- **Why:** closed under vector space operations, and polynomials are indefinitely differentiable

Hilbert spaces: Convergence

Definition

A sequence of vectors $x_0, x_1, \ldots$ in a normed vector space V is said to converge to $v \in V$ when $\lim_{k\to\infty} \|v - x_k\| = 0$, or for any $\varepsilon > 0$, there exists K_ε such that $\|v - x_k\| \; < \; \varepsilon$ for all $k > K_\varepsilon$.

- Choice of the norm in V is key

Example

For $k \in \mathbb{Z}^+$, let

$$x_k(t) \; = \; \begin{cases} 1, & \text{for } t \in [0, 1/k]; \\ 0, & \text{otherwise.} \end{cases}$$

$v(t) = 0$ for all t. Then for $p \in [1, \infty)$,

$$\|v - x_k\|_p \; = \; \left(\int_{-\infty}^{\infty} |v(t) - x_k(t)|^p \, dt \right)^{1/p} \; = \; \left(\frac{1}{k} \right)^{1/p} \; \xrightarrow{k\to\infty} \; 0,$$

For $p = \infty$: $\|v - x_k\|_\infty \; = \; 1 \quad$ for all k

Hilbert spaces: Completeness

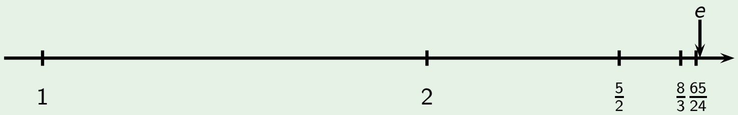
Definitions

- A sequence $\{x_n\}$ is a Cauchy sequence in a normed space when for any $\varepsilon > 0$, there exists k_ε such that $\|x_k - x_m\| < \varepsilon$ for all $k, m > k_\varepsilon$
- A normed vector space V is complete if every Cauchy sequence converges in V
- A complete normed vector space is called a Banach space
- A complete inner product space is called a Hilbert space

Hilbert spaces

Examples

- $\mathbb{Q}$ is not a complete space
  - $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} \longrightarrow \frac{\pi^2}{6} \quad \in \mathbb{R}, \notin \mathbb{Q}$
  - $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n!} \longrightarrow e \quad \in \mathbb{R}, \notin \mathbb{Q}$



- $\mathbb{R}$ is a complete space

Hilbert spaces

Examples

- All finite dimensional spaces are complete
- $\ell^p(\mathbb{Z})$ and $\mathcal{L}^p(\mathbb{R})$ are complete
  - $\ell^2(\mathbb{Z})$ and $\mathcal{L}^2(\mathbb{R})$ are Hilbert spaces
- $C^q([a, b])$, functions on $[a, b]$ with q continuous derivatives, are not complete except for $q = 0$ under $\mathcal{L}^\infty$ norm
- Vector space of random variables as already defined is a Hilbert space

Gibbs phenomenon

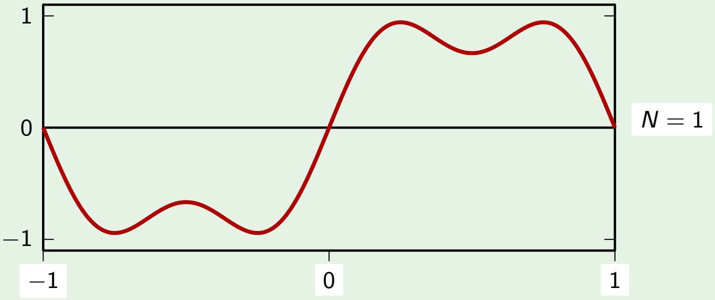
Approximating a square wave with partial sums of the Fourier series

$$\sum_{k=0}^{N} \mathbf{x}^{(2k+1)}, \qquad \mathbf{x}^{(n)} = \sin(\pi\, n\, t)/n, \quad t \in [-1, 1]$$

Gibbs phenomenon

Approximating a square wave with partial sums of the Fourier series

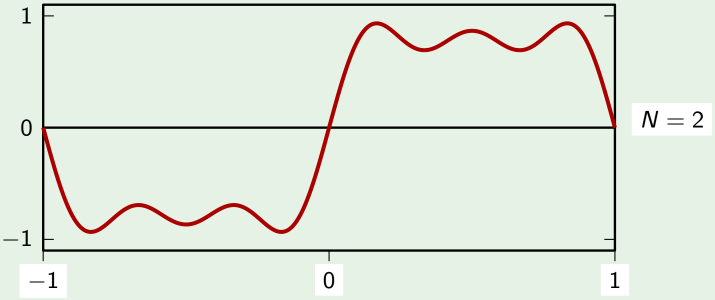
$$\sum_{k=0}^{N} \mathbf{x}^{(2k+1)}, \qquad \mathbf{x}^{(n)} = \sin(\pi\, n\, t)/n, \quad t \in [-1, 1]$$



Gibbs phenomenon

Approximating a square wave with partial sums of the Fourier series

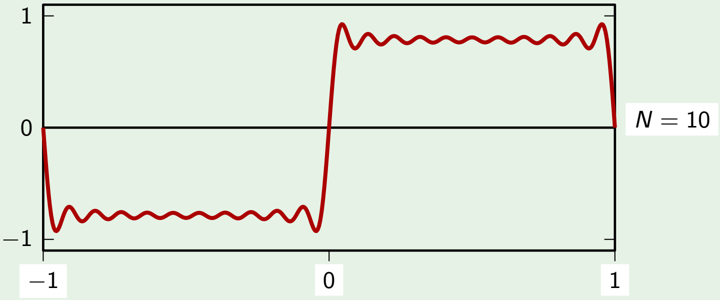
$$\sum_{k=0}^{N} \mathbf{x}^{(2k+1)}, \qquad \mathbf{x}^{(n)} = \sin(\pi\, n\, t)/n, \quad t \in [-1, 1]$$



Gibbs phenomenon

Approximating a square wave with partial sums of the Fourier series

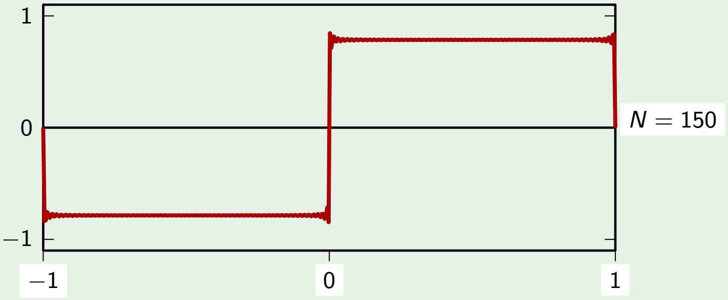
$$\sum_{k=0}^{N} \mathbf{x}^{(2k+1)}, \qquad \mathbf{x}^{(n)} = \sin(\pi\, n\, t)/n, \quad t \in [-1, 1]$$



Gibbs phenomenon

Approximating a square wave with partial sums of the Fourier series

$$\sum_{k=0}^{N} \mathbf{x}^{(2k+1)}, \qquad \mathbf{x}^{(n)} = \sin(\pi\, n\, t)/n, \quad t \in [-1, 1]$$



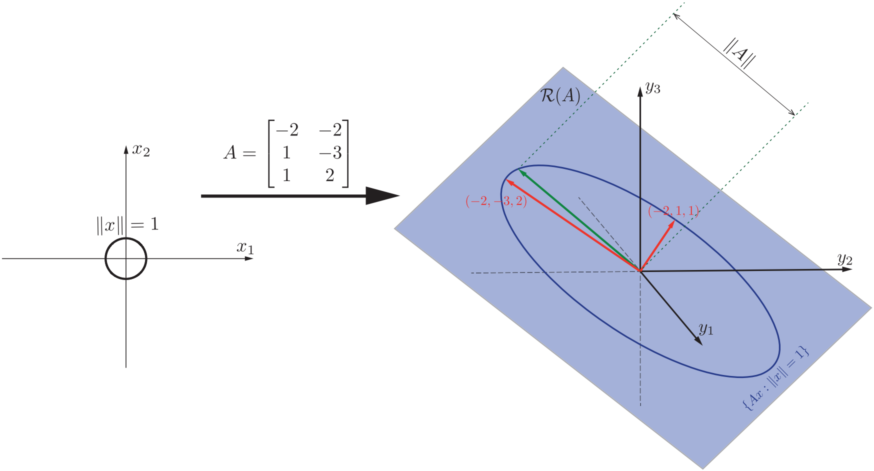
Linear operators

Linear operators generalize matrices

Definitions

- $A : H_0 \to H_1$ is a linear operator when for all $x, y \in H_0, \alpha \in \mathbb{C}$:
  1. Additivity: $A(x + y) = Ax + Ay$
  2. Scalability: $A(\alpha x) = \alpha(Ax)$
- Null space (subspace of H_0): $\mathcal{N}(A) = \{x \in H_0, Ax = 0\}$
- Range space (subspace of H_1): $\mathcal{R}(A) = \{Ax \in H_1, x \in H_0\}$
- Operator norm: $\|A\| = \sup_{\|x\|=1} \|Ax\|$
- A is bounded when: $\|A\| < \infty$
- Inverse: Bounded $B : H_1 \to H_0$ inverse of bounded A if and only if:
  $BAx = x$, for every $x \in H_0$
  $ABy = y$, for every $y \in H_1$

Linear operators: Illustration



- $\mathcal{R}(A)$ is the plane $5y_1 + 2y_2 + 8y_3 = 0$ since $(-2, 1, 1) \times (-2, -3, 2) = (5, 2, 8)$, where $\times$ denotes the cross-product

Adjoint operators

Adjoint generalizes Hermitian transposition of matrices

Definition (Adjoint and self-adjoint operators)

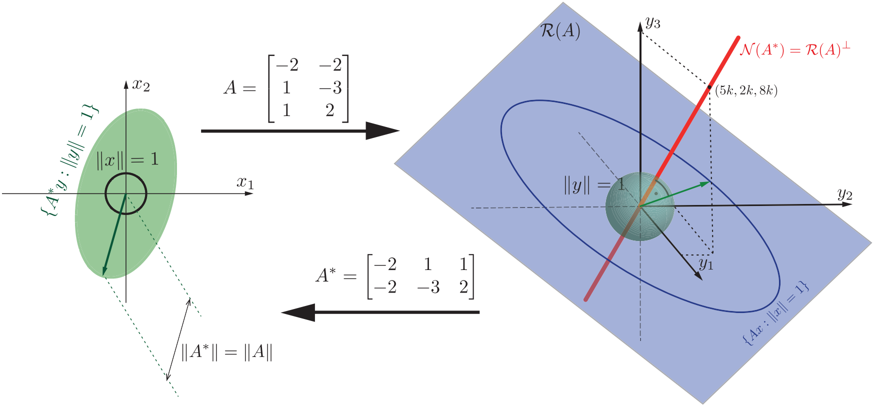
- $A^* : H_1 \to H_0$ is the adjoint of $A : H_0 \to H_1$ when

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

- If $A = A^*, A$ is self-adjoint or Hermitian

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

Adjoint operator: Illustration



- $\mathcal{N}(A^*)$ is the line $\frac{y_1}{5} = \frac{y_2}{2} = \frac{y_3}{8}$, since again $(-2, 1, 1) \times (-2, -3, 2) = (5, 2, 8)$

Adjoint operators

Theorem (Adjoint properties)

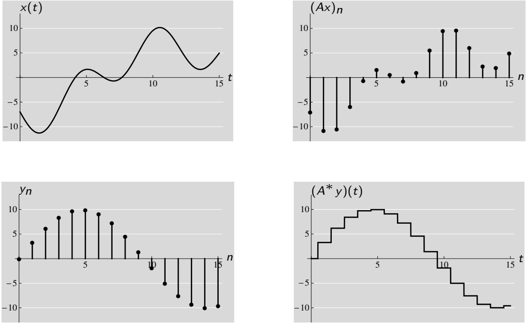
Let $A : H_0 \longrightarrow H_1$ *be a bounded linear operator*

1. A^* *exists and is unique*
2. $(A^*)^* = A$
3. AA^* *and* A^*A *are self-adjoint*
4. $\|A^*\| = \|A\|$
5. *If* A *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,* $(BA)^* = A^*B^*$

Adjoint operators: Local averaging

$$A : \mathcal{L}^2(\mathbb{R}) \to \ell^2(\mathbb{Z}) \qquad (Ax)_k = \int_{k-1/2}^{k+1/2} x(t)\,dt$$

$$\begin{aligned}
\langle Ax, y\rangle_{\ell^2} &= \sum_{n\in\mathbb{Z}} (Ax)_n\, y_n^* = \sum_{n\in\mathbb{Z}} \left(\int_{n-1/2}^{n+1/2} x(t)\,dt\right) y_n^* = \sum_{n\in\mathbb{Z}} \int_{n-1/2}^{n+1/2} x(t) y_n^*\,dt \\
&= \sum_{n\in\mathbb{Z}} \int_{n-1/2}^{n+1/2} x(t)(A^* y)^*(t)\,dt = \int_{-\infty}^{\infty} x(t)(A^* y)^*(t)\,dt = \langle x, A^* y\rangle_{\mathcal{L}^2}
\end{aligned}$$



Unitary operators

Definition (Unitary operators)

- 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 $x, y \in H_0$

- If A is unitary, then $\|Ax\|^2 = \|x\|^2$
- A is unitary if and only if $A^{-1} = A^*$

Projection operators

Definition (Projection, orthogonal projection, oblique projection)

- P is idempotent when $P^2 = P$
- A projection operator is a bounded linear operator that is idempotent
- An orthogonal projection operator is a self-adjoint projection operator
- An oblique projection operator is not self adjoint

Theorem

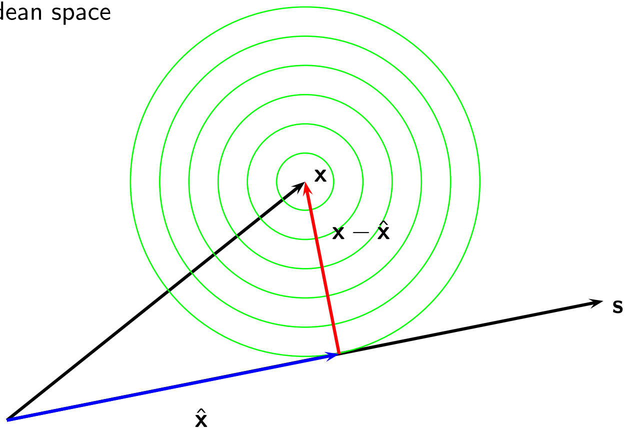
- *A bounded linear operator P on H satisfies $\langle x - Px, Py \rangle = 0$ for all $x, y \in H$ iff P is an orthogonal projection operator*

Theorem

- *If $A : H_0 \to H_1$, $B : H_1 \to H_0$ bounded and A is a left inverse of B, then BA is a projection operator. If $B = A^*$ then, $BA = A^* A$ is an orthogonal projection*

Best approximation: Euclidean geometry

- x is a point in Euclidean space
- S is a line in Euclidean space



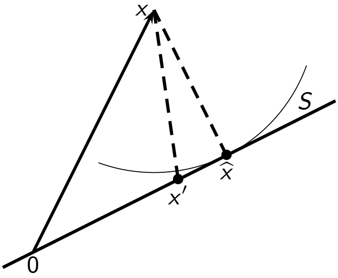
- Nearest point problem: Find $\widehat{x} \in S$ that is closest to x
- Solution uniquely determined by $x - \widehat{x} \perp S$
  - Circle must touch S in one point, radius $\perp$ tangent

Best approximation: Hilbert space geometry

- S closed subspace of a Hilbert space
- Best approximation problem:

Find $\widehat{x} \in S$ that is closest to x

$$\widehat{x} = \operatorname*{argmin}_{s \in S} \|x - s\|$$



Best approximation by orthogonal projection

Theorem (Projection theorem)

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

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

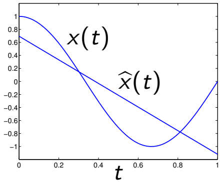
All "nearest vector in a subspace" problems in Hilbert spaces are the same

Example 1: Least-square polynomial approximation

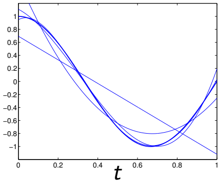
- Consider: $x(t) = \cos(\frac{3}{2}\pi t) \in \mathcal{L}^2([0,1])$
- Find the degree-1 polynomial closest to x (in $\mathcal{L}^2$ norm)
- Solution: Use orthogonality

$$0 \;=\; \langle x(t) - \widehat{x}(t),\, 1\rangle_t \;=\; \int_0^1 \left(\cos(\tfrac{3}{2}\pi t) - (a_0 + a_1 t)\right)\, 1\, dt \;=\; -\frac{2}{3\pi} - a_0 - \frac{1}{2}a_1$$

$$0 \;=\; \langle x(t) - \widehat{x}(t),\, t\rangle_t \;=\; \int_0^1 \left(\cos(\tfrac{3}{2}\pi t) - (a_0 + a_1 t)\right)\, t\, dt \;=\; -\frac{4+6\pi}{9\pi^2} - \frac{1}{2}a_0 - \frac{1}{3}a_1$$



Approx. with degree 1 polynomial



Approx. with higher degree polynomials

Example 2: MMSE estimate

- Consider: Real-valued random variable x
- Find the constant c that minimizes $\mathrm{E}\big[\,(\mathrm{x}-c)^2\,\big]$
- Note:
  - Expected square is a Hilbert space norm
  - Constants are a closed subspace in vector space of random variables
- Solution: Use orthogonality
  - c determined uniquely by $\mathrm{E}[\,(\mathrm{x}-c)\,\alpha c\,] \;=\; 0$ for all $\alpha \in \mathbb{R}$
  - $c \;=\; \mathrm{E}[\,\mathrm{x}\,]$
- Alternative:
  - Expand into quadratic function of c and minimize with calculus
  - Not too difficult, but lacks insight

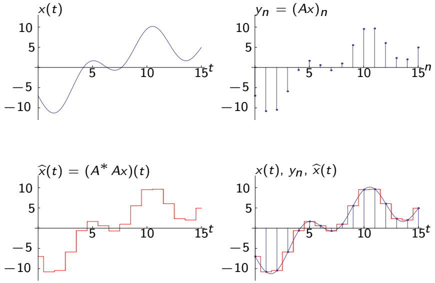
Example 3: Best piecewise-constant approximation

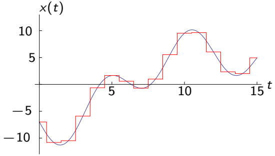
- Local averaging

$$A : \mathcal{L}^2(\mathbb{R}) \to \ell^2(\mathbb{Z}) \qquad\qquad (Ax)_k = \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} x(t)dt$$

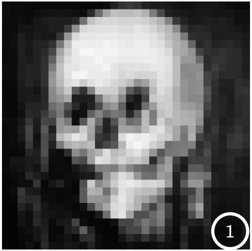
  has adjoint $A^* : \ell^2(\mathbb{Z}) \to \mathcal{L}^2(\mathbb{R})$ that produces staircase function

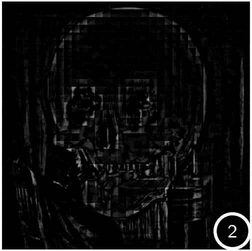
- AA^* is identity, so A^*A is orthogonal projection

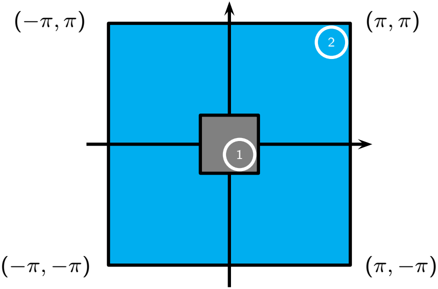




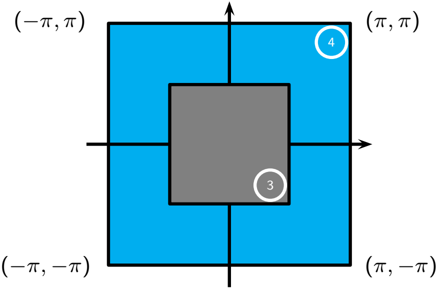
Example 4: Approximations of “All is vanity” image—Haar



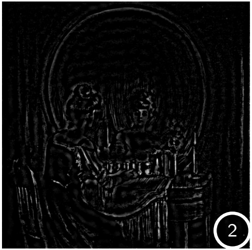


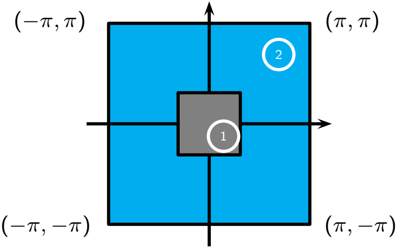


Example 4: Approximations of "All is vanity" image—Haar



Example 4: Approximations of “All is vanity” image—sinc

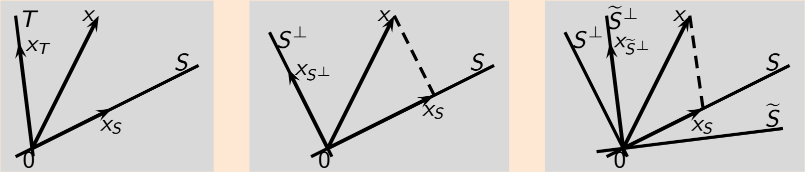




Projection and direct sums

Theorem

- *P projection on H, $S = \mathcal{R}(P)$, $T = \mathcal{N}(P)$. Then $H = S \oplus T$*



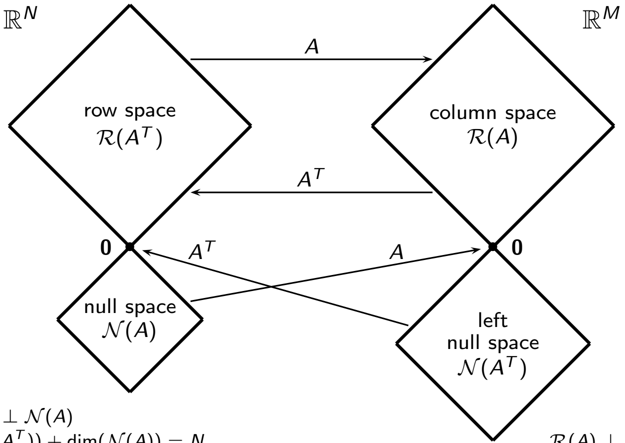
(a) Decomposition *(b) Orthogonal projection $T = S^\perp$* *(c) Oblique projection $T = \widetilde{S}^\perp$*

- *If S, T closed subspaces s.t. $H = S \oplus T$ then there exists projection P on H s.t. $S = \mathcal{R}(P)$ and $T = \mathcal{N}(P)$*

Summary

- Geometry is key to gain intuition and understanding
- Vector spaces, subspaces
- Norms, inner products
- Hilbert spaces
- Linear operators, adjoints
- Projections

As an exercise...



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

$\dim(\mathcal{R}(A^T)) + \dim(\mathcal{N}(A)) = N$

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

$\dim(\mathcal{R}(A)) + \dim(\mathcal{N}(A^T)) = M$