COM-202: Signal Processing

Chapter 2: Signal Processing and Vector Space

Overview:

- signal processing as geometry
- vectors and vector space
  - inner product space
  - bases
- subspace approximations

Introduction to vector space and linear algebra

You (most likely) already know and use the mechanics:

- "standard" (Euclidean) vectors, dot product
- matrices, determinants, transposition, matrix-vector multiplication
- change of basis, linear independence

What we can do for you:

- show you the practical (engineering) use of all that
- show you how general and useful this is
- help you start thinking in infinitely dimensional space

What is linear algebra

What is algebra

algebra is branch of mathematics in which arithmetical operations and formal manipulations are applied to abstract symbols rather than specific numbers

[Britannica]

What is linear algebra

Linear algebra:

- focuses on linear operations (addition and scaling)
- operates on *multidimensional* entities called vectors and matrices
- defines a “world” called *vector space* that has an intuitive geometrical structure

Key ideas behind linear algebra

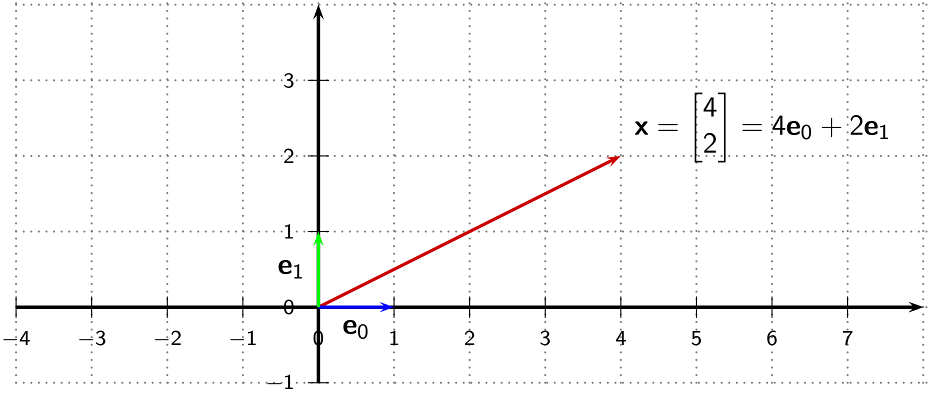
- a vector can be "anything"
- any vector can be expressed as a linear combination of fundamental building blocks (aka *basis vectors*)
- basis vectors define a coordinate space
- the "amount" of each basis vector in a given vector is the *scalar* (numerical) value of the corresponding *coordinates* in the space

and therefore:

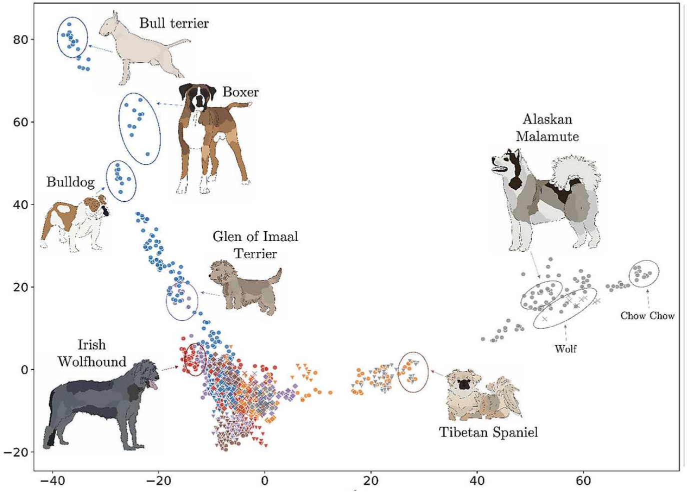
- coordinates are just tuples of numbers ($\begin{bmatrix}2 & 3.5 & 2 & -0.8\end{bmatrix}^T$)
- we can use standard linear algebra tools independently of the "true" nature of vectors

Example

$$\mathbb{R}^2: \qquad \mathbf{x} = \begin{bmatrix} x_0 & x_1 \end{bmatrix}^T$$



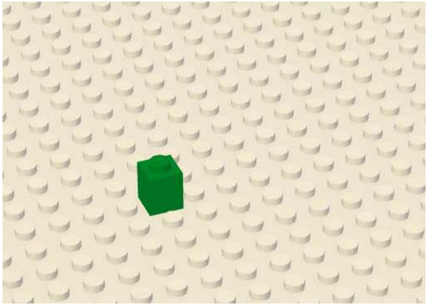
Example



Bull terrier
Boxer
Bulldog
Glen of Imaal Terrier
Irish Wolfhound
Alaskan Malamute
Chow Chow
Wolf
Tibetan Spaniel
80
60
40
20
0
−20
−40
−20
0
20
40
60

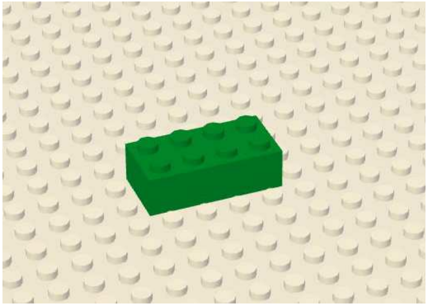
Analogy: LEGO

basic building block: $\mathbf{g} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$



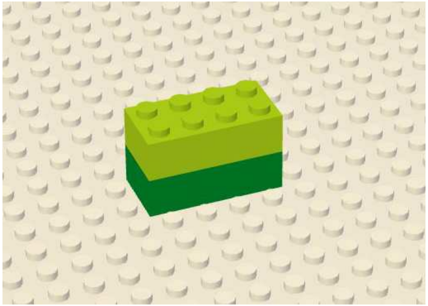
Analogy: LEGO

$$\text{reshaping: } \mathbf{x} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{g}$$



Analogy: LEGO

combining: $\mathbf{y} = \mathbf{x} \oplus \mathbf{x}$



Why use vector spaces in SP?

all types of signals can be represented as vectors
in a suitable vector space

The unifying framework

- space of continuous-time bandlimited signals
  (more on this later in the course!)
- space of discrete-time signals of length N
- space of discrete-time periodic signals of period N
- space of discrete-time, finite-energy signals
- etc.

Main advantages

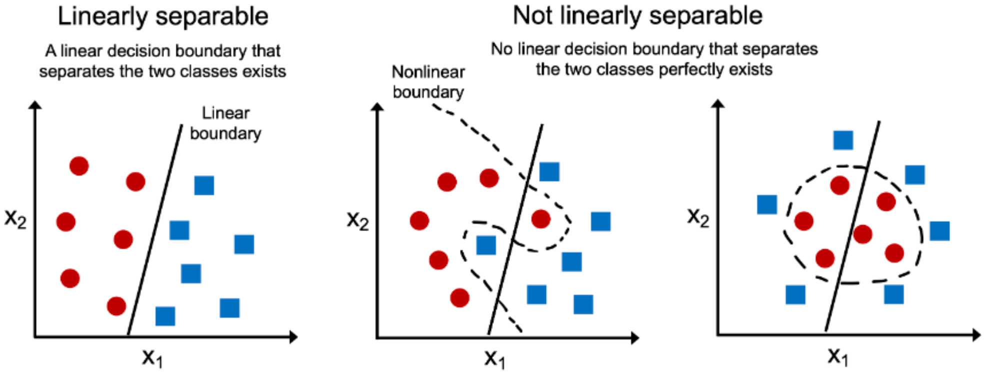
Math will be easier and signal properties will automatically apply to all signal spaces:

- we can select the "easiest" space to prove a result
- the Fourier Transform will be easy to understand
- and so will the proof of the sampling theorem

Vector space is the closest thing to an algorithmic framework for SP:

- most vectors can be thought of as simple arrays
- approximation and compression algorithms are simple applications of linear algebra
- low-dimensional intuition is super helpful

Are you into data science?



vector space

Let's talk about Vector Spaces...

Some spaces should be very familiar:

- $\mathbb{R}^2, \mathbb{R}^3$: Euclidean space, geometry
- $\mathbb{R}^N, \mathbb{C}^N$: linear algebra

Others perhaps not so much...

- $\ell_2(\mathbb{Z})$: space of square-summable infinite sequences
- $L_2([a, b])$: space of square-integrable *functions* over an interval

yes, vectors can be functions!

The space of square-summable infinite sequences $\ell_2(\mathbb{Z})$

$$\mathbf{x} \in \ell_2(\mathbb{Z}) \iff \|\mathbf{x}\|^2 = \sum_{n \in \mathbb{Z}} \big|x[n]\big|^2 < \infty$$

this is exactly the space of finite energy infinite signals

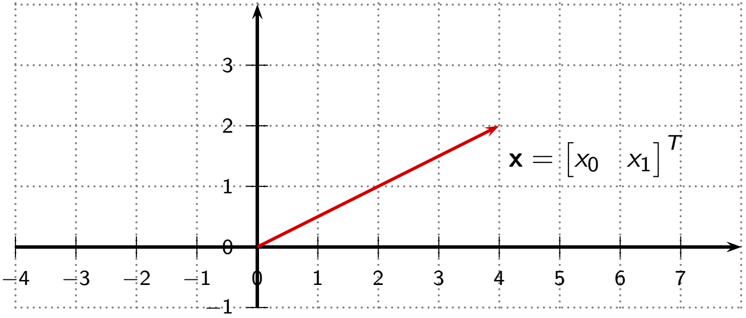
The space of square-integrable functions over an interval $L_2([a, b])$

$$\mathbf{x} \in L_2([a, b]) \iff \int_a^b |x(t)|^2 dt < \infty$$

Graphical representation of a vector

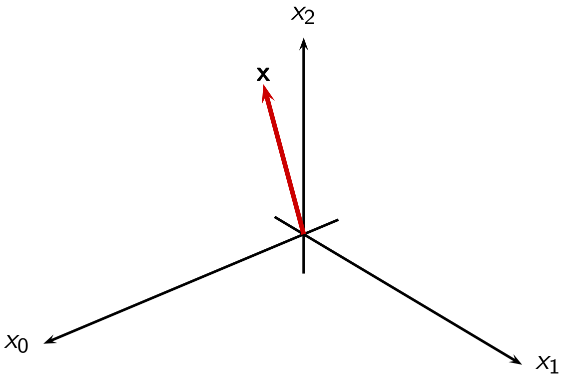
Some spaces can be shown graphically...

$$\mathbb{R}^2: \qquad \mathbf{x} = \begin{bmatrix} x_0 & x_1 \end{bmatrix}^T$$



Some spaces can be shown graphically...

$$\mathbb{R}^3: \qquad \mathbf{x} = \begin{bmatrix} x_0 & x_1 & x_2 \end{bmatrix}^T$$

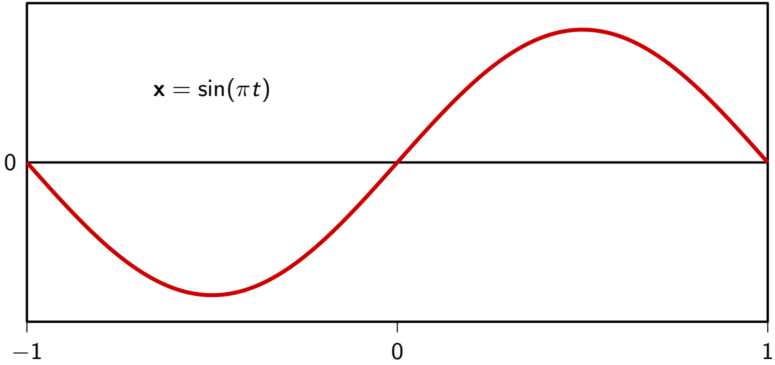


...but most spaces cannot

$$\mathbb{R}^N \text{ for } N > 3: \qquad \mathbf{x} = \begin{bmatrix} x_0 & x_1 & \dots & x_{N-1} \end{bmatrix}^T$$

Some spaces can be shown graphically...

$$L_2([-1,1]): \qquad \mathbf{x} = x(t) \in \mathbb{R}, \quad t \in [-1,1]$$



...but most spaces cannot

$f : \mathbb{C} \to \mathbb{C}$, analytic

The three take-home lessons today

- vectors represent a wide variety of different objects
- vector space is defined by properties that do not depend on the nature of the vectors
- once you are in vector space, you can always use the same tools

Analogy for programmers: OOP

```
class Polygon:
    def __init__(self, num_sides, side_len=1, x=0, y=0):
        self.num_sides = num_sides
        self.side_len = side_len
        self.center = [x, y]

    def resize(self, factor):
        self.side_len *= factor

    def translate(self, x, y):
        self.center[0] += x
        self.center[1] += y

    def plot(self):
        ...
```

Analogy for programmers: OOP

```
class Triangle(Polygon):
    def __init__(self):
        super().__init__(3)

    ...


class Square(Polygon):
    def __init__(self):
        super().__init__(4)

    ...
```

Vector space: operational definition

Ingredients:

- a set of vectors V
- a set of scalars (say $\mathbb{C}$)

To have a vector space we need *at least* to be able to:

- resize vectors (scalar multiplication)
- combine vectors together (vector addition)

Vector space: the axioms

For $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha, \beta \in \mathbb{C}$:

- $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
- $\exists 0 \in V \quad | \quad \mathbf{x} + 0 = 0 + \mathbf{x} = \mathbf{x}$
- $\forall \mathbf{x} \in V \ \exists (-\mathbf{x}) \quad | \quad \mathbf{x} + (-\mathbf{x}) = 0$
- $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{y} + \alpha\mathbf{x}$
- $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$
- $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$ (field and scalar multiplications play well together)
- $1 \cdot \mathbf{x} = \mathbf{x} \qquad 1 \in \mathbb{C}$

Example: $\mathbb{R}^N$

$$\mathbf{x} = \begin{bmatrix} x_0 & x_1 & \dots & x_{N-1} \end{bmatrix}^T$$

$$\mathbf{y} = \begin{bmatrix} y_0 & y_1 & \dots & y_{N-1} \end{bmatrix}^T$$

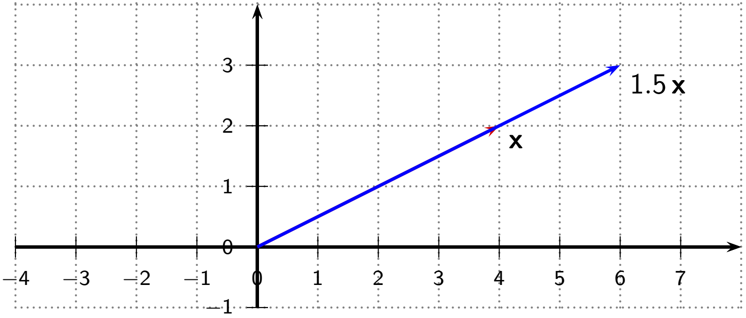
These definition of scalar multiplication and vector addition fulfill the axioms:

$$\alpha\mathbf{x} = \begin{bmatrix} \alpha x_0 & \alpha x_1 & \dots & \alpha x_{N-1} \end{bmatrix}^T$$

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_0 + y_0 & x_1 + y_1 & \dots & x_{N-1} + y_{N-1} \end{bmatrix}^T$$

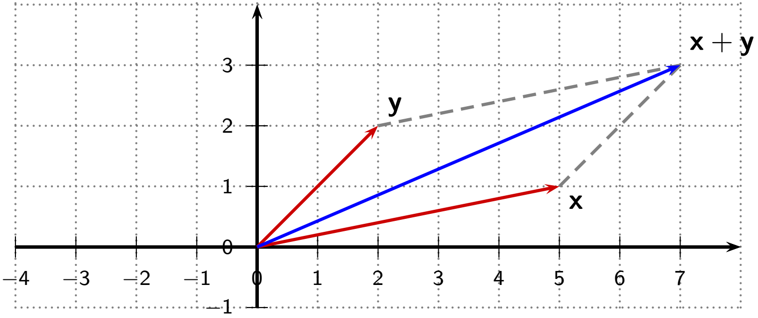
Scalar multiplication in $\mathbb{R}^2$

$$\alpha \mathbf{x} = \begin{bmatrix} \alpha x_0 & \alpha x_1 \end{bmatrix}^T$$



Vector addition in $\mathbb{R}^2$

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_0 + y_0 & x_1 + y_1 \end{bmatrix}^T$$



Example: $L_2[-1, 1]$

Vectors are now functions over $[-1, 1]$

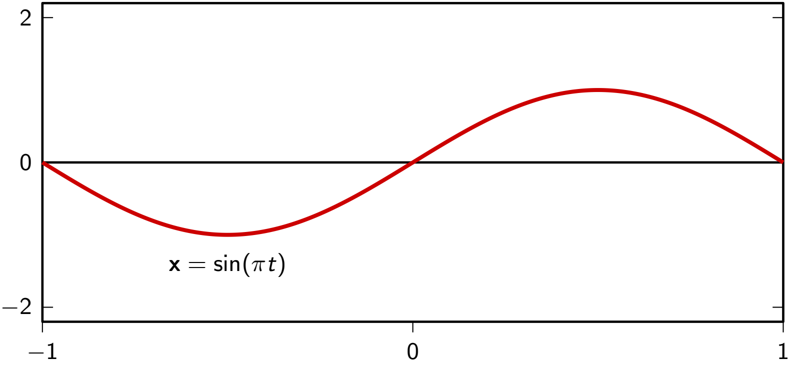
$$\mathbf{x} = x(t), \quad t \in [-1, 1]$$
$$\mathbf{y} = y(t), \quad t \in [-1, 1]$$

These definition of scalar multiplication and vector addition fulfill the axioms:

$$\alpha \mathbf{x} = \alpha\, x(t)$$
$$\mathbf{x} + \mathbf{y} = x(t) + y(t)$$

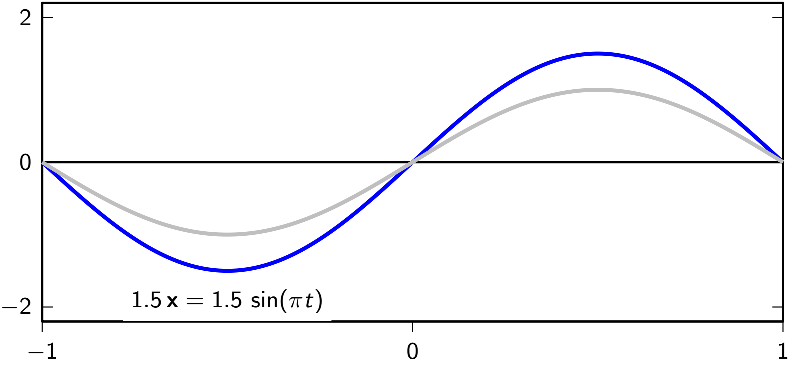
Scalar multiplication in $L_2[-1,1]$

$$\alpha \mathbf{x} = \alpha\, x(t)$$



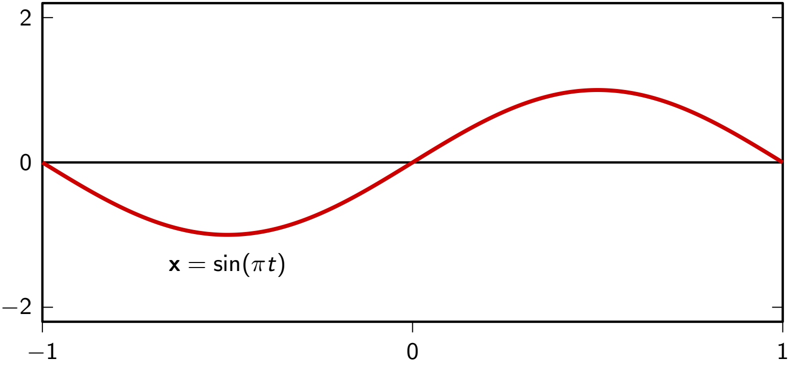
Scalar multiplication in $L_2[-1,1]$

$$\alpha \mathbf{x} = \alpha\, x(t)$$



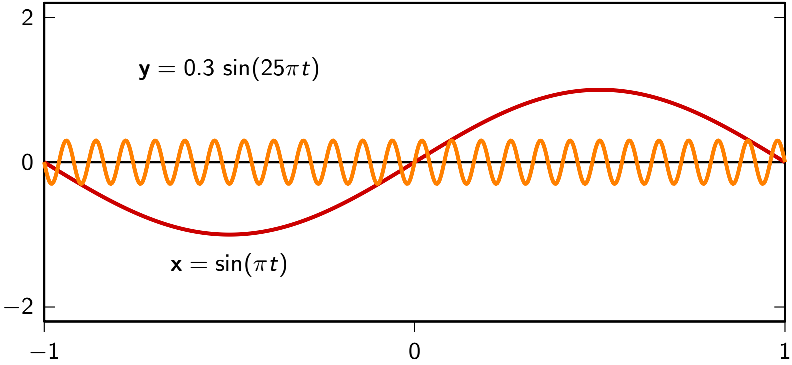
Addition in $L_2[-1, 1]$

$$\mathbf{x} + \mathbf{y} = x(t) + y(t)$$



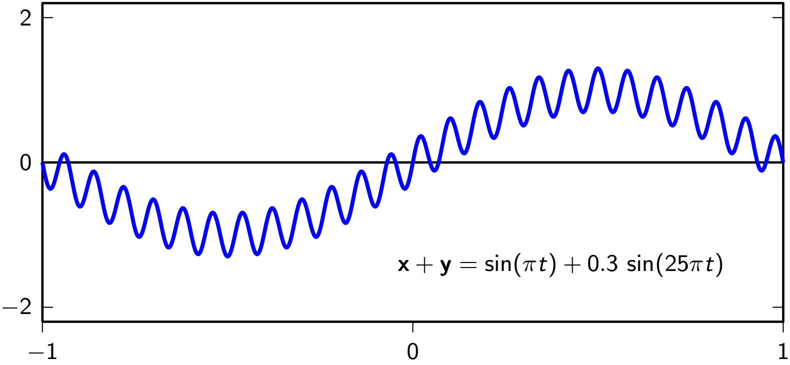
Addition in $L_2[-1, 1]$

$$\mathbf{x} + \mathbf{y} = x(t) + y(t)$$



Addition in $L_2[-1,1]$

$$\mathbf{x} + \mathbf{y} = x(t) + y(t)$$



Vector space is kinda dull: we need something more

So far:

- the set of vectors V
- a set of scalars (say $\mathbb{C}$)
- scalar multiplication
- vector addition

We need something to measure and compare:
inner product (aka dot product)

inner product space

Inner product

$$\langle \cdot, \cdot \rangle \; : \; V \times V \to \mathbb{C}$$

- measure of similarity between vectors
- inner product is zero? vectors are *orthogonal* (maximally different)

Inner product space: the axioms

Given $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha \in \mathbb{C}$:

- $\langle \mathbf{x}, \alpha(\mathbf{y} + \mathbf{z}) \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle + \alpha \langle \mathbf{x}, \mathbf{z} \rangle$ (linearity in the second argument)
- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^*$ (conjugate symmetry)
- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ (positive-semidefinte)
- $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

The *norm* (aka the "length" of a vector):

- definition $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$
- we can take the root because of axiom 3
- axiom 2 is necessary to ensure axiom 3 when the scalar field is $\mathbb{C}$

Word of the day: sesquilinear

From the axioms we can prove this:

$$\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$$

$$\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha^* \langle \mathbf{x}, \mathbf{y} \rangle$$

- usually, linearity is in the first argument
- our choice will make more sense later when we study the Fourier transform

Inner product in $\mathbb{C}^N$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=0}^{N-1} x^*[n]y[n]$$

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

$$\mathbf{x} = \begin{bmatrix} x_0 & x_1 \end{bmatrix}^T$$

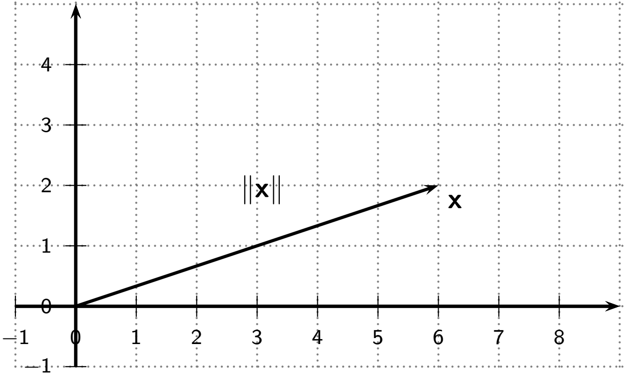
$$\mathbf{y} = \begin{bmatrix} y_0 & y_1 \end{bmatrix}^T$$

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

- not immediate to see why it's a measure of similarity
- easier to start with the norm

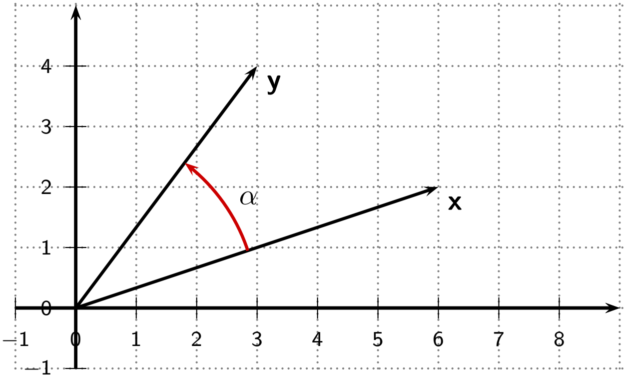
The norm in $\mathbb{R}^2$ (i.e. Pythagoras' theorem)

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_0^2 + x_1^2 = \|\mathbf{x}\|^2$$



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

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



Inner product as a measure of similarity

If two vectors have unit norm, $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$, then $\langle \mathbf{x}, \mathbf{y} \rangle = \cos \alpha$

Important cases for unit-norm vectors (easy to see in $\mathbb{R}^2$ but valid for all types of vector space!):

- $-1 \leq \langle \mathbf{x}, \mathbf{y} \rangle \leq 1$
- when $\langle \mathbf{x}, \mathbf{y} \rangle = 1$ vectors are aligned (maximally similar)
- when $\langle \mathbf{x}, \mathbf{y} \rangle = -1$ vectors are pointing in opposite directions (still maximally similar, but with a "flip")

Orthogonality

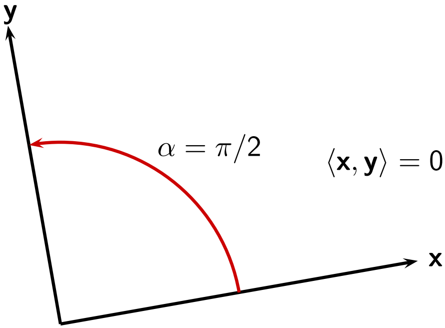
Orthogonality is one of the most powerful concepts in inner-product space:

- two nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ are called orthogonal if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$
- orthogonal vectors are maximally different (i.e. they have nothing in common)

- orthogonality is geometrically intuitive in low-dimensional Euclidean space
- "orthogonal" is Greek for "right angle"
- intuition helps us with the abstract concept of orthogonality

Orthogonality in $\mathbb{R}^2$

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



$$\langle \mathbf{x}, \mathbf{y} \rangle = 0$$

Inner product and orthogonality in functional spaces

Back to $L_2[-1,1]$; definition of inner product:

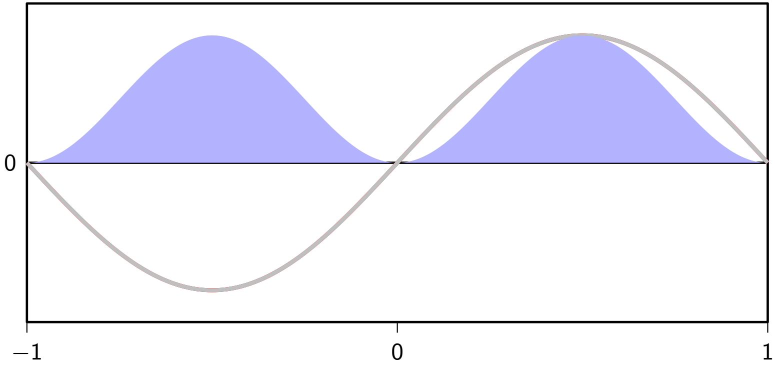
$$\mathbf{x} = x(t), \quad t \in [-1,1]$$
$$\mathbf{y} = y(t), \quad t \in [-1,1]$$

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

let's use our low-dimensional intuition to understand norm and orthogonality...

The norm in $L_2[-1,1]$

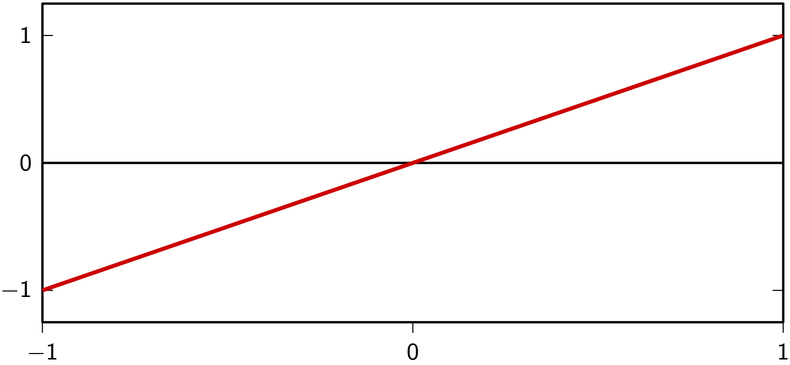
$\mathbf{x} = \sin \pi t \qquad \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2 = \int_{-1}^{1} \sin^2(\pi t) dt = 1$



if $x(t)$ was a voltage, the norm would measure the energy over the time interval $[-1, 1]$

The norm in $L_2[-1, 1]$

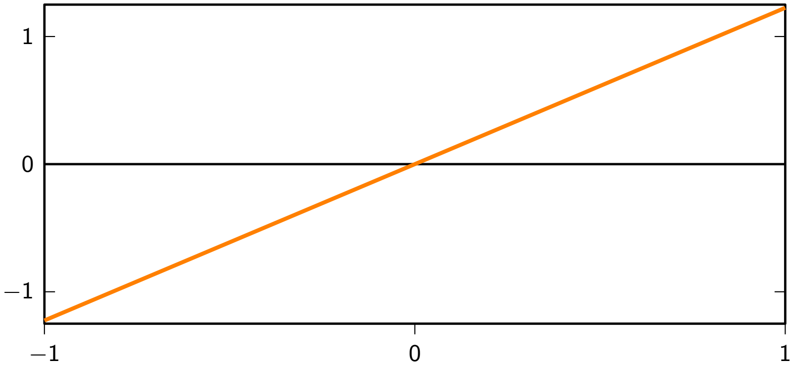
$$\mathbf{z} = t \qquad\qquad \|\mathbf{z}\|^2 = \int_{-1}^{1} t^2\, dt = 2/3$$



Normalizing a vector

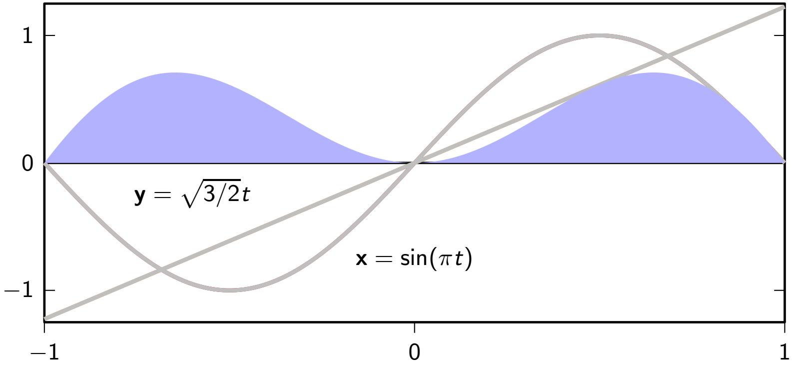
Normalization ensures that the norm is equal to one:

$$\mathbf{y} = \mathbf{z}/\|\mathbf{z}\| = \sqrt{3/2}\, t \qquad\qquad \|\mathbf{y}\|^2 = (3/2) \int_{-1}^{1} t^2 \, dt = 1$$



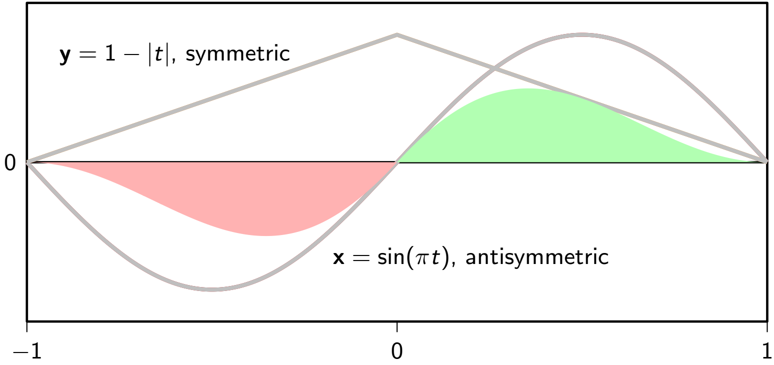
Inner product in $L_2[-1,1]$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \int_{-1}^{1} x(t)y(t)dt = \int_{-1}^{1} \sqrt{3/2}t \sin(\pi t)dt = \approx 0.78 \approx \cos(38.7^\circ)$$



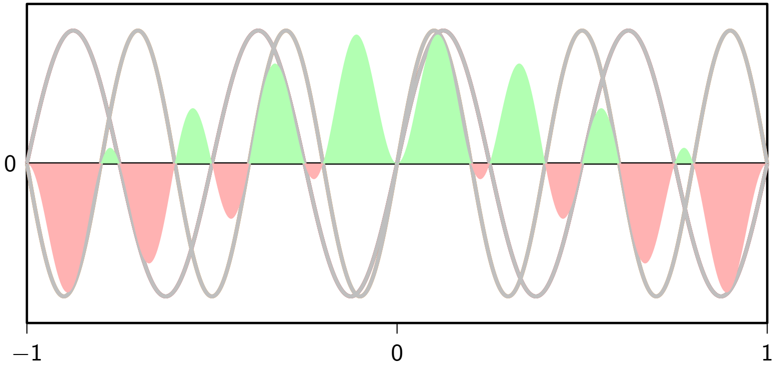
Orthogonality in $L_2[-1, 1]$

$\mathbf{x}, \mathbf{y}$ from orthogonal subspaces: $\langle \mathbf{x}, \mathbf{y} \rangle = 0$



Orthogonality in $L_2[-1, 1]$

sinusoids with frequencies integer multiples of a fundamental



$\mathbf{y} = \sin(5\pi t)$

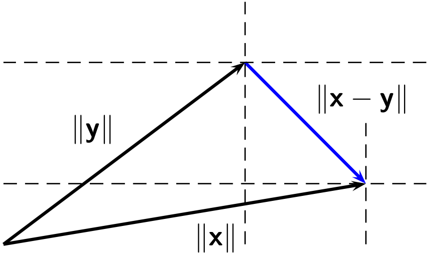
$\mathbf{x} = \sin(4\pi t)$

Norm vs Distance

- inner product defines a norm: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$
- norm defines a distance: $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$

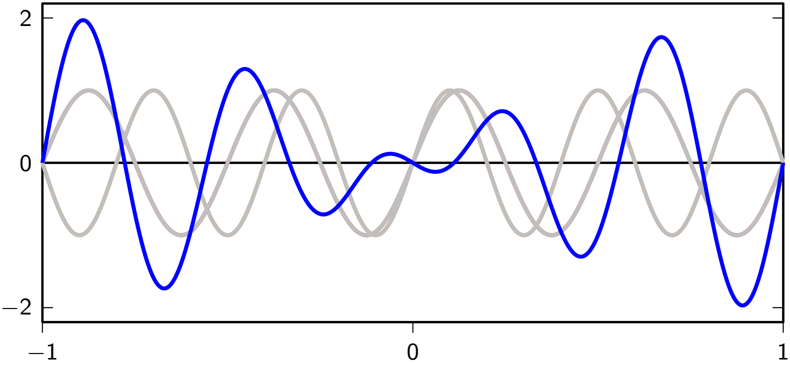
Norm and distance in $\mathbb{R}^2$

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



Distance in $L_2[-1,1]$: the Mean Square Error

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



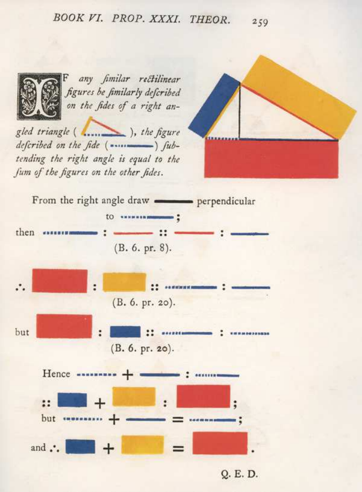
$\mathbf{y} = \sin(5\pi t)$

$\mathbf{x} = \sin(4\pi t)$

A familiar result

Pythagorean theorem:

$$\|\mathbf{x}+\mathbf{y}\|^2 = \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle$$
$$= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \text{ for } \mathbf{x} \perp \mathbf{y}$$



From Euclid's elements by Oliver Byrne (1810 - 1880)

Caution: dangerous turns ahead

We want to make sure that the inner product is always well defined. Trouble spots:

- inner product of infinite-length vectors
- inner product of arbitrary functions

We will need to “restrict” the elements of a vector space

$\ell_2(\mathbb{Z})$ vs $\mathbb{C}^\infty$

For infinite-length vectors the inner product is an infinite sum that may explode

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=-\infty}^{\infty} x^*[n]y[n]$$

To ensure this doesn't happen we use $\ell_2(\mathbb{Z})$:

- vectors in $\ell_2(\mathbb{Z})$ are infinite-dimensional tuples
- $\mathbf{x} \in \ell_2(\mathbb{Z}) \Leftrightarrow \|\mathbf{x}\|^2 = \sum_{n=-\infty}^{\infty} |x_n|^2 < \infty$
- elements of $\ell_2(\mathbb{Z})$ are known as *finite-energy sequences*

$\ell_2(\mathbb{Z})$

Exercise: consider the infinite-length vectors $\mathbf{x}$ and $\mathbf{y}$ whose elements are

$$x_n = \begin{cases} 0 & n \leq 0 \\ 1/\sqrt{n} & n > 0 \end{cases}$$

$$y_n = \begin{cases} 0 & n \leq 0 \\ 1/n & n > 0 \end{cases}$$

and show that $\mathbf{x} \notin \ell_2(\mathbb{Z})$ but $\mathbf{y} \in \ell_2(\mathbb{Z})$

$L_2([a, b])$

Similarly, to ensure that $\langle \mathbf{x}, \mathbf{y} \rangle = \int_a^b x^*(t) y(t) < \infty$, we define $L_2([a, b])$ as the space of square-integrable functions over $[a, b]$:

$$\mathbf{x} = x(t) \in L_2([a, b]) \Leftrightarrow \|\mathbf{x}\|^2 = \int_a^b |x(t)|^2 < \infty$$

Example: $\mathbf{x} = 1/t \notin L_2([0, 1])$

bases

Vector space and bases: providing structure

Key ideas:

- a basis is a coordinate system for a vector space
- every vector can be expressed as a linear combination of basis elements
- a vector space has infinitely many bases and we can move from one to the other
- certain special bases allow us to see more clearly the information contained in vectors

Basis

Given a vector space V, a basis is a set of vectors $\{\mathbf{w}_k\}_{k=0,1,\ldots,N-1} \in V$ such that any $\mathbf{x} \in V$ can be expressed as a linear combination

$$\mathbf{x} = \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k$$

for a unique set of scalars $\alpha_0, \ldots, \alpha_{N-1}$.

The canonical $\mathbb{R}^2$ basis

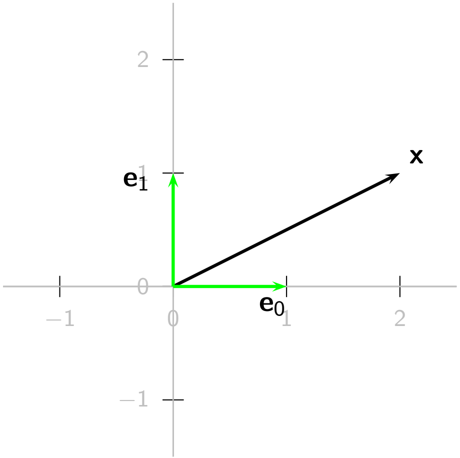
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{e}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{e}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The canonical $\mathbb{R}^2$ basis

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = 2\mathbf{e}_0 + \mathbf{e}_1$$



Another $\mathbb{R}^2$ basis

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

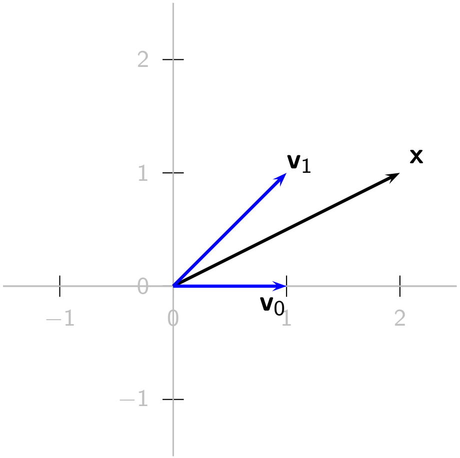
$$\mathbf{v}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\alpha_1 = x_1 - x_2, \quad \alpha_2 = x_2$$

Another $\mathbb{R}^2$ basis

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = \mathbf{v}_0 + \mathbf{v}_1$$



Not all pairs of vectors form a basis for $\mathbb{R}^2$...

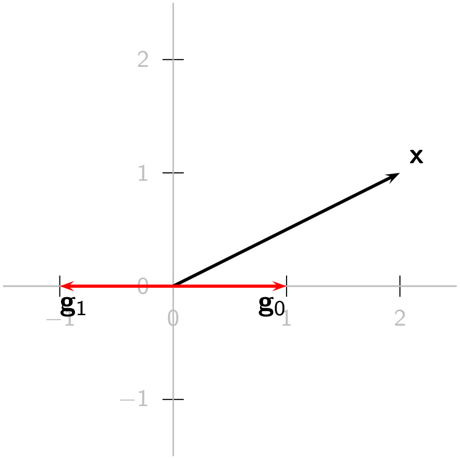
$$\mathbf{g}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{g}_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq \alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} \text{ whenever } x_2 \neq 0$$

“Parallel” vectors are not a basis for $\mathbb{R}^2$

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\mathbf{x} \neq \alpha_1 \mathbf{g}_0 + \alpha_2 \mathbf{g}_1$$



What about infinite-dimensional spaces?

$$\mathbf{x} = \sum_{k=-\infty}^{\infty} \alpha_k \, \mathbf{w}_k$$

A basis for $\ell_2(\mathbb{Z})$

$$\mathbf{e}_k = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

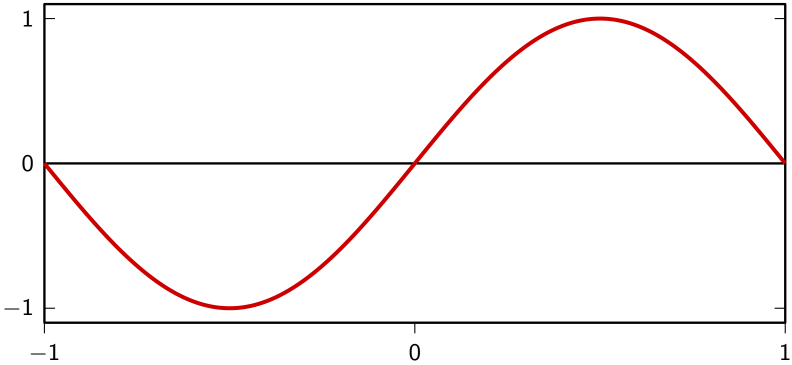
k-th position, $k \in \mathbb{Z}$

What about functional vector spaces?

$$f(t) = \sum_k \alpha_k \, h_k(t)$$

A basis for odd functions over $[-1, 1]$

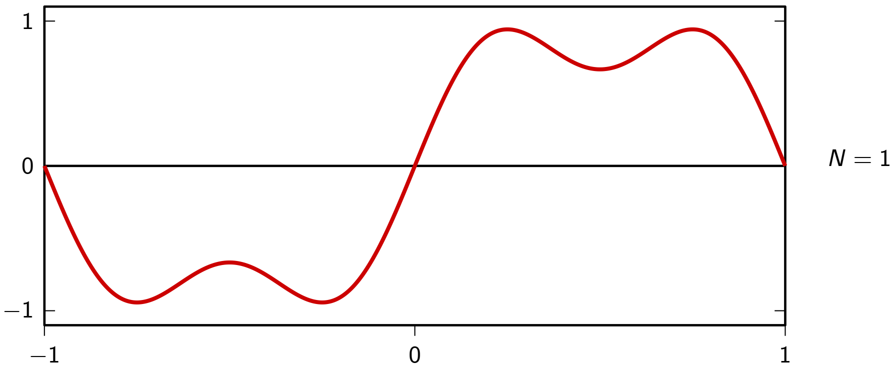
$$\sum_{k=0}^{N} \frac{1}{2k+1}\mathbf{v}_k, \qquad \mathbf{v}_k = v_k(t) = \sin(\pi(2k+1)t), \quad t \in [-1, 1]$$



$N = 0$

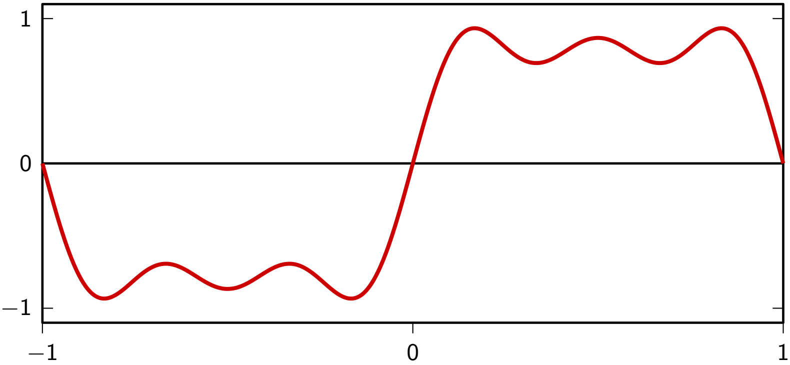
A basis for odd functions over $[-1, 1]$

$$\sum_{k=0}^{N} \frac{1}{2k+1}\mathbf{v}_k, \qquad \mathbf{v}_k = v_k(t) = \sin(\pi(2k+1)t), \quad t \in [-1, 1]$$



A basis for odd functions over $[-1, 1]$

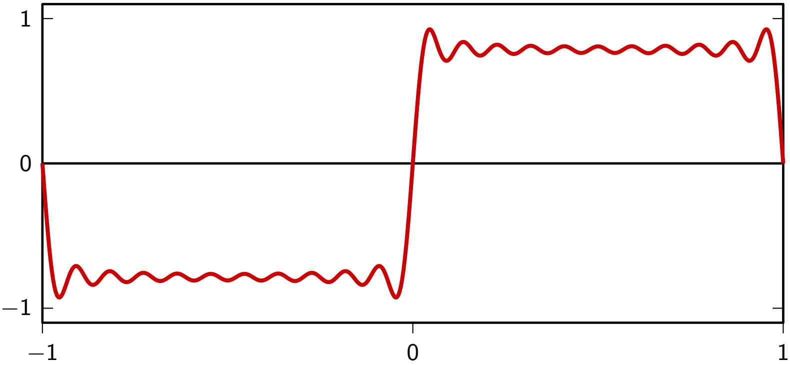
$$\sum_{k=0}^{N} \frac{1}{2k+1}\mathbf{v}_k, \qquad \mathbf{v}_k = v_k(t) = \sin(\pi(2k+1)t), \quad t \in [-1, 1]$$



$N = 2$

A basis for odd functions over $[-1, 1]$

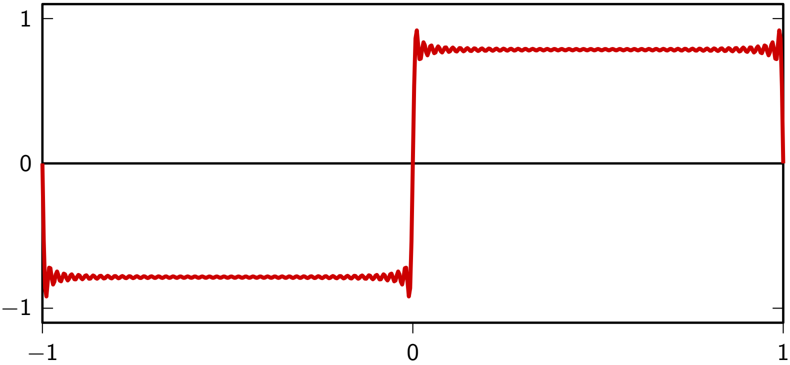
$$\sum_{k=0}^{N} \frac{1}{2k+1}\mathbf{v}_k, \qquad \mathbf{v}_k = v_k(t) = \sin(\pi(2k+1)t), \quad t \in [-1, 1]$$



$N = 10$

A basis for odd functions over $[-1, 1]$

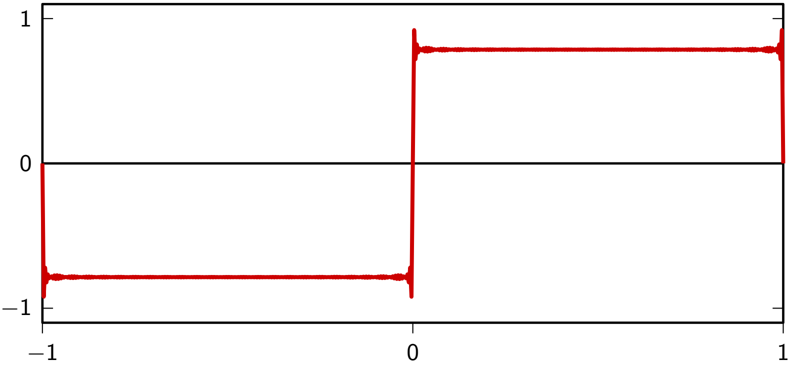
$$\sum_{k=0}^{N} \frac{1}{2k+1} \mathbf{v}_k, \qquad \mathbf{v}_k = v_k(t) = \sin(\pi(2k+1)t), \quad t \in [-1, 1]$$



$N = 50$

A basis for odd functions over $[-1, 1]$

$$\sum_{k=0}^{N} \frac{1}{2k+1}\mathbf{v}_k, \qquad \mathbf{v}_k = v_k(t) = \sin(\pi(2k+1)t), \quad t \in [-1, 1]$$



$N = 150$

Bases: formal definition

Given:

- a vector space V
- a set of N vectors from V: $W = \{\mathbf{w}_k\}_{k=0,1,\ldots,N-1}$

W is a basis for V if:

1. we can write for *all* $\mathbf{x} \in V$:

$$\mathbf{x} = \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k, \quad \alpha_k \in \mathbb{C}$$

2. the elements of W are linearly independent:

$$\sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k = 0 \quad \Longleftrightarrow \quad \alpha_k = 0, \; k = 0, 1, \ldots, N-1$$

Bases: formal definition

Linear independence implies unique representation

$$\mathbf{x} = \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k, \quad \alpha_k \in \mathbb{C}$$

and the coefficients α_k are unique

Special bases

Orthogonal basis:

$$\langle \mathbf{w}_k, \mathbf{w}_h \rangle = 0 \text{ for } k \neq h$$

Orthonormal basis:

$$\langle \mathbf{w}_k, \mathbf{w}_h \rangle = \delta[h-k] = \begin{cases} 1 & h = k \\ 0 & h \neq k \end{cases}$$

(we can always orthonormalize a basis via the Gram-Schmidt algorithm)

Basis expansion

$$\mathbf{x} = \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k$$

how do we find the coefficients?

Orthonormal bases are the best:

$$\alpha_k = \langle \mathbf{w}_k, \mathbf{x} \rangle$$

Orthogonal change of basis

Start from a vector expressed in terms of a basis $\{\mathbf{w}_k\}$

$$\mathbf{x} = \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k = \sum_{h=0}^{N-1} \beta_k \mathbf{v}_h$$

We want to express the same vector using a new orthonormal basis $\{\mathbf{v}_k\}$

$$\mathbf{x} = \sum_{h=0}^{N-1} \beta_k \mathbf{v}_h$$

Orthogonal change of basis matrix

$$\beta_h = \langle \mathbf{v}_h, \mathbf{x} \rangle = \langle \mathbf{v}_h, \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k \rangle$$

$$= \sum_{k=0}^{N-1} \alpha_k \langle \mathbf{v}_h, \mathbf{w}_k \rangle = \sum_{k=0}^{N-1} \alpha_k \, c_{h,k}.$$

$$\begin{bmatrix} \beta_0 \\ \vdots \\ \beta_{N-1} \end{bmatrix} = \begin{bmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,N-1} \\ & & \vdots & \\ c_{N-1,0} & c_{N-1,1} & \cdots & c_{N-1,N-1} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_{N-1} \end{bmatrix} = \mathbf{C}\boldsymbol{\alpha}$$

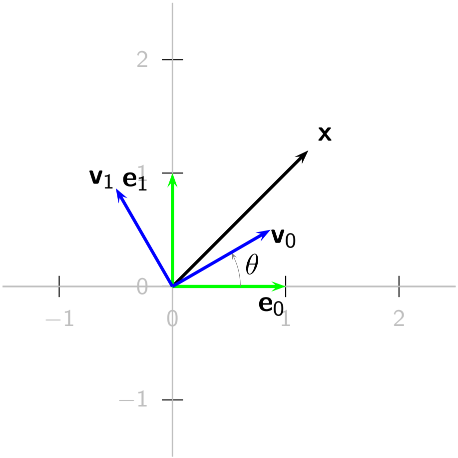
Change of basis: example

- canonical basis $E = \{\mathbf{e}_0, \mathbf{e}_1\}$
- $\mathbf{x} = \alpha_0 \mathbf{e}_0 + \alpha_1 \mathbf{e}_1$
- new basis $V = \{\mathbf{v}_0, \mathbf{v}_1\}$ with

$$\mathbf{v}_0 = \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix}^T$$

$$\mathbf{v}_1 = \begin{bmatrix} -\sin\theta & \cos\theta \end{bmatrix}^T$$

- $\mathbf{x} = \beta_0 \mathbf{v}_0 + \beta_1 \mathbf{v}_1$



Change of basis: example

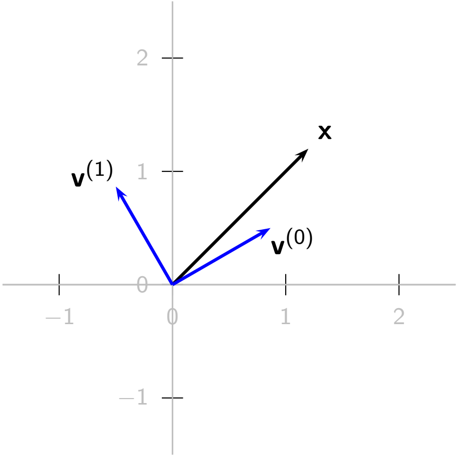
- new basis is orthonormal;
  the change of basis matrix is

$$c_{hk} = \langle \mathbf{v}_h, \mathbf{e}_k \rangle$$

- new coefficients in compact form:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \mathbf{R}\boldsymbol{\alpha}$$

- $\mathbf{R}$ is a 2D rotation matrix
- key property: $\mathbf{R}^T\mathbf{R} = \mathbf{I}$



Orthonormal change of basis matrices

Orthonormal change of basis matrices are unitary:

$$\mathbf{C}^H\mathbf{C} = \mathbf{C}\mathbf{C}^H = \mathbf{I}$$

The superscript H denotes Hermitian transposition:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^H = \begin{bmatrix} a^* & d^* & g^* \\ b^* & e^* & h^* \\ c^* & f^* & i^* \end{bmatrix}$$

Norm and energy

We saw that the square norm is a proxy for the "energy" of a vector.
If

- $\mathbf{x} = \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k$
- $\{\mathbf{w}_k\}$ is an orthonormal basis

then

$$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \sum_{k=0}^{N-1} |\alpha_k|^2 = \|\boldsymbol{\alpha}\|^2$$

$$= \boldsymbol{\alpha}^H \boldsymbol{\alpha} = \begin{bmatrix} \alpha_0^* & \ldots & \alpha_{N-1}^* \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_{N-1} \end{bmatrix}$$

Parseval's Theorem (conservation of energy)

If:

- $\mathbf{x} = \sum_{k=0}^{N-1} \alpha_k \mathbf{w}_k = \sum_{k=0}^{N-1} \beta_k \mathbf{v}_k \qquad (\{\mathbf{w}_k\}, \{\mathbf{v}_k\}$ orthonormal)
- $\boldsymbol{\beta} = \mathbf{C}\boldsymbol{\alpha}$

then:

$$\begin{aligned}
\|\mathbf{x}\|^2 &= \|\boldsymbol{\beta}\|^2 \\
&= \boldsymbol{\beta}^H \boldsymbol{\beta} \\
&= (\mathbf{C}\boldsymbol{\alpha})^H (\mathbf{C}\boldsymbol{\alpha}) \\
&= \boldsymbol{\alpha}^H \mathbf{C}^H \mathbf{C} \boldsymbol{\alpha} \\
&= \boldsymbol{\alpha}^H \boldsymbol{\alpha} \\
&= \|\boldsymbol{\alpha}\|^2
\end{aligned}$$

Conservation of energy: example

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \mathbf{R}\boldsymbol{\alpha}$$

- square norm in canonical basis: $\|\mathbf{x}\|^2 = \alpha_0^2 + \alpha_1^2$
- square norm in rotated basis: $\|\mathbf{x}\|^2 = \beta_0^2 + \beta_1^2$
- let's verify Parseval:

$$\begin{aligned} \beta_0^2 + \beta_1^2 &= \boldsymbol{\beta}^T \boldsymbol{\beta} \\ &= (\mathbf{R}\boldsymbol{\alpha})^T (\mathbf{R}\boldsymbol{\alpha}) \\ &= \boldsymbol{\alpha}^T (\mathbf{R}^T \mathbf{R}) \boldsymbol{\alpha} \\ &= \boldsymbol{\alpha}^T \boldsymbol{\alpha} = \alpha_0^2 + \alpha_1^2 \end{aligned}$$

Finite basis example: Fourier basis

Claim: the set of N signals in $\mathbb{C}^N$

$$w_k[n] = e^{j\frac{2\pi}{N}nk}, \qquad n, k = 0, 1, \ldots, N-1$$

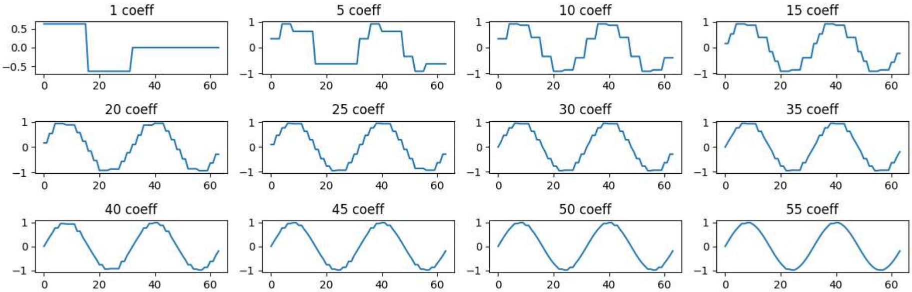
is an orthogonal basis in $\mathbb{C}^N$.

More on this next week!

Finite basis example: Haar basis

- orthonormal
- represents the signal information robustly
  resilient against data loss
- easy to compute

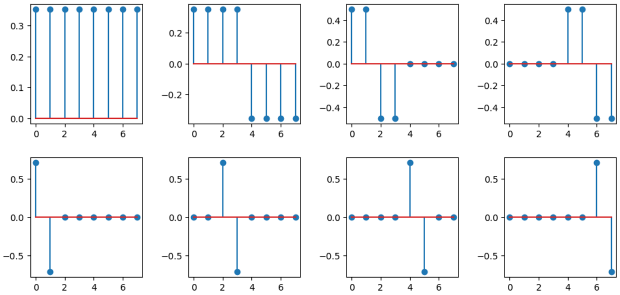
Haar basis: robust representation



Haar basis: construction

- the coefficient for the first basis vector encodes the average value of the data
- the coefficient for the second basis vector encodes the difference between the averages of the first half and the second half of the data
- every subsequent coefficient encodes a difference between the averages of alternating sets of data points

Haar basis $N = 8$



See the first python lab for code examples

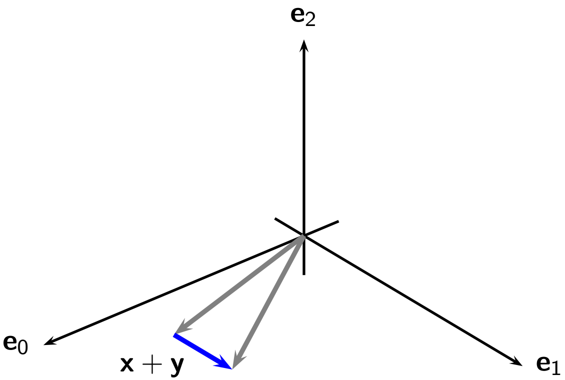
subspace approximations

Vector subspace

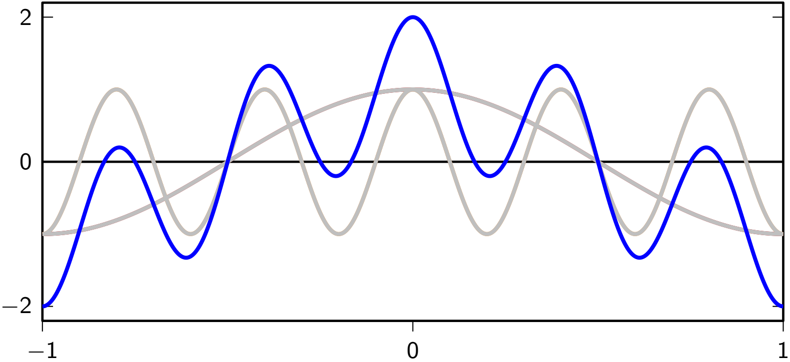
a subset of vectors *closed* under addition and scalar multiplication

Example in Euclidean Space

intuition: $\mathbb{R}^2 \subset \mathbb{R}^3$



Subspace of symmetric functions over $L_2[-1, 1]$



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

Subspaces have their own basis

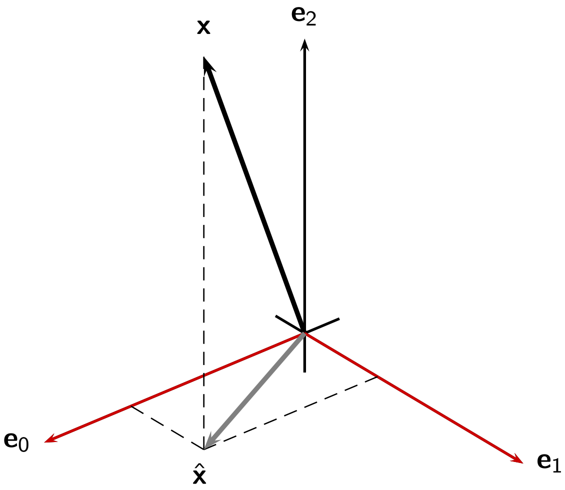
$$\mathbf{e}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{e}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

basis vector for the plane in $\mathbb{R}^3$

Approximation

Problem:

- vector $\mathbf{x} \in V$
- subspace $S \subseteq V$
- approximate $\mathbf{x}$ with $\hat{\mathbf{x}} \in S$



Least-Squares Approximation

- $\{\mathbf{s}_k\}_{k=0,1,\ldots,K-1}$ orthonormal basis for $S \subseteq V$
- orthogonal projection:

$$\hat{\mathbf{x}} = \sum_{k=0}^{K-1} \langle \mathbf{s}_k, \mathbf{x} \rangle \, \mathbf{s}_k$$

orthogonal projection is the "best" approximation over S

Least-Squares Approximation: definition of “best”

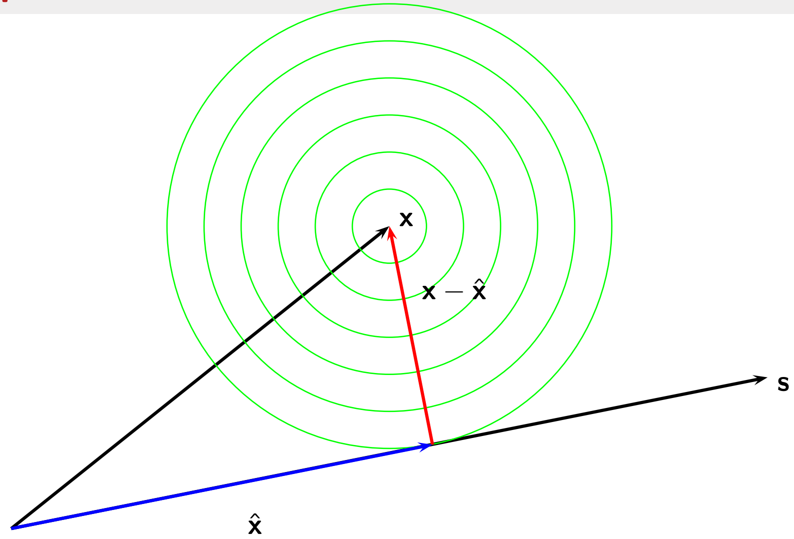
- orthogonal projection has minimum-norm error (minimum-energy error):

$$\arg\min_{\mathbf{y}\in S} \|\mathbf{x} - \mathbf{y}\| = \hat{\mathbf{x}}$$

- error is orthogonal to approximation (everything has been extracted):

$$\langle \mathbf{x} - \hat{\mathbf{x}}, \hat{\mathbf{x}} \rangle = 0$$

Least Squares Approximation



Example: polynomial approximation

$P_N[-1,1]$: subspace of polynomials of degree N over the interval $[-1,1]$

- $P_N[-1,1] \subset L_2[-1,1]$
- $\mathbf{p} = a_0 + a_1 t + \ldots + a_N t^N$
- clearly $P_N[-1,1]$ is closed under addtion and scalar multiplication
- a self-evident, naive basis: $\mathbf{s}_k = t^k, \quad k = 0, 1, \ldots, N$
- naive basis is not orthonormal

Example: polynomial approximation

goal: approximate $\mathbf{x} = \sin t \in L_2[-1,1]$ over $P_2[-1,1]$

- build orthonormal basis from naive basis
- project $\mathbf{x}$ over the orthonormal basis
- compute approximation error
- compare to well-known Taylor approximation

Building an orthonormal basis

Gram-Schmidt orthonormalization procedure:

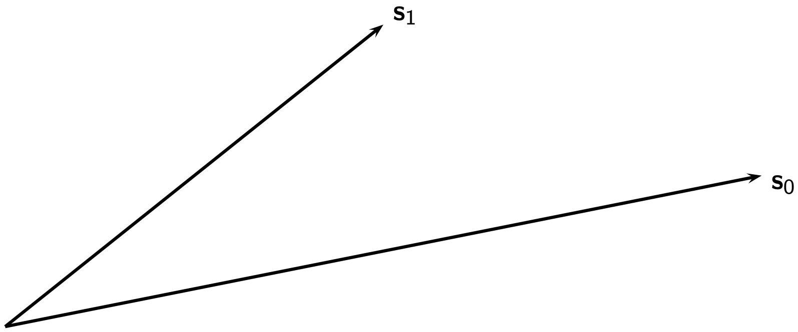
$$\{\mathbf{s}_k\} \longrightarrow \{\mathbf{u}_k\}$$

original set → orthonormal set

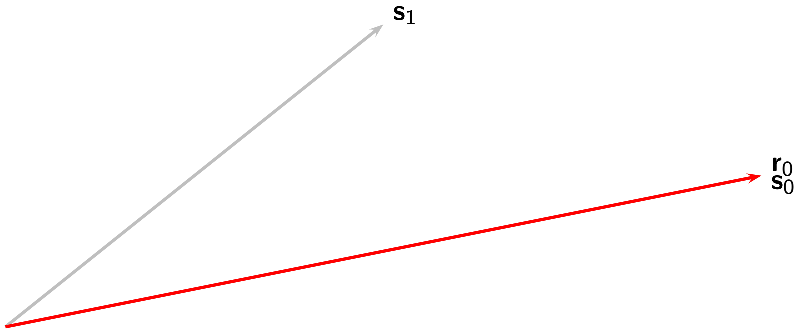
Algorithmic procedure: at each step k

1. $\mathbf{r}_k = \mathbf{s}_k - \sum_{n=0}^{k-1} \langle \mathbf{u}_n, \mathbf{s}_k \rangle \, \mathbf{u}_n$
2. $\mathbf{u}_k = \mathbf{r}_k / \|\mathbf{r}_k\|$

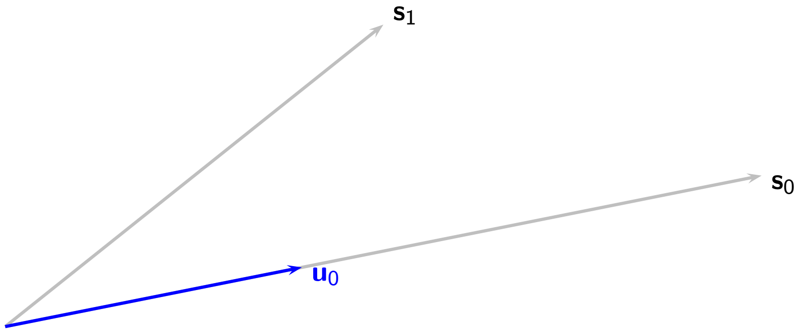
Building an orthonormal basis



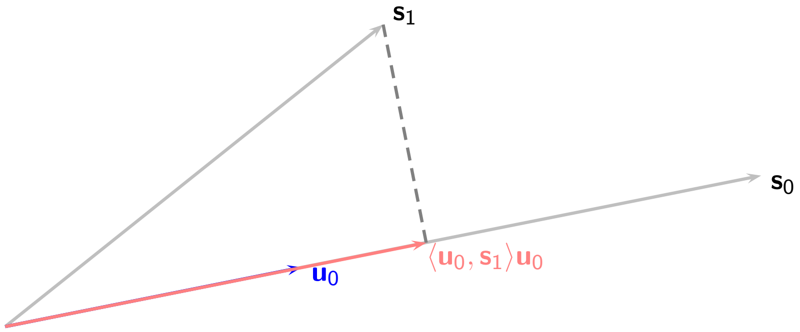
Building an orthonormal basis



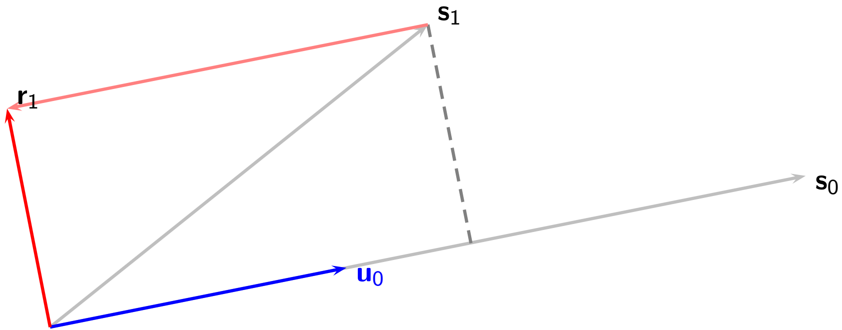
Building an orthonormal basis



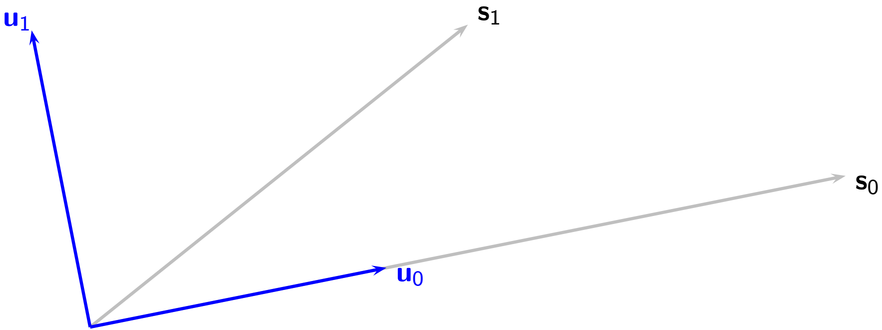
Building an orthonormal basis



Building an orthonormal basis



Building an orthonormal basis



Building an orthonormal basis

Gram-Schmidt orthonormalization of the naive basis: $\{\mathbf{s}_k\} \rightarrow \{\mathbf{u}_k\}$

- $\mathbf{s}_0 = 1$
  - $\mathbf{r}_0 = \mathbf{s}_0 = 1$
  - $\|\mathbf{r}_0\|^2 = 2$
  - $\mathbf{u}_0 = \mathbf{r}_0 / \|\mathbf{r}_0\| = \sqrt{1/2}$
- $\mathbf{s}_1 = t$
  - $\langle \mathbf{u}_0, \mathbf{s}_1 \rangle = \int_{-1}^{1} t/\sqrt{2} = 0$
  - $\mathbf{r}_1 = \mathbf{s}_1 = t$
  - $\|\mathbf{r}_1\|^2 = 2/3$
  - $\mathbf{u}_1 = \sqrt{3/2}\, t$
- $\mathbf{s}_2 = t^2$
  - $\langle \mathbf{u}_0, \mathbf{s}_2 \rangle = \int_{-1}^{1} t^2/\sqrt{2} = 2/3\sqrt{2}$
  - $\langle \mathbf{u}_1, \mathbf{s}_2 \rangle = \int_{-1}^{1} t^3/\sqrt{2} = 0$
  - $\mathbf{r}_2 = \mathbf{s}_2 - (2/3\sqrt{2})\mathbf{u}_0 = t^2 - 1/3$
  - $\|\mathbf{r}_2\|^2 = 8/45$
  - $\mathbf{u}_2 = \sqrt{5/8}(3t^2 - 1)$

Legendre polynomials

The Gram-Schmidt algorithm leads to an orthonormal basis for $P_N([-1,1])$

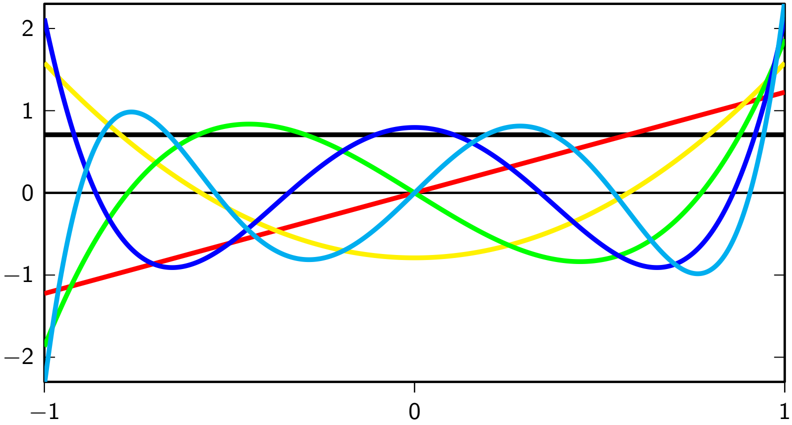
$$\mathbf{u}_0 = \sqrt{1/2}$$

$$\mathbf{u}_1 = \sqrt{3/2}\, t$$

$$\mathbf{u}_2 = \sqrt{5/8}(3t^2 - 1)$$

$$\mathbf{u}_3 = \dots$$

Legendre Polynomials



Orthogonal projection over $P_2[-1,1]$

$$\alpha_k = \langle \mathbf{u}_k, \mathbf{x} \rangle = \int_{-1}^{1} u_k(t) \, \sin t \, dt$$

- $\alpha_0 = \langle \sqrt{1/2}, \sin t \rangle = 0$
- $\alpha_1 = \langle \sqrt{3/2}\, t, \sin t \rangle \approx 0.7377$
- $\alpha_2 = \langle \sqrt{5/8}(3t^2 - 1), \sin t \rangle = 0$

Approximation

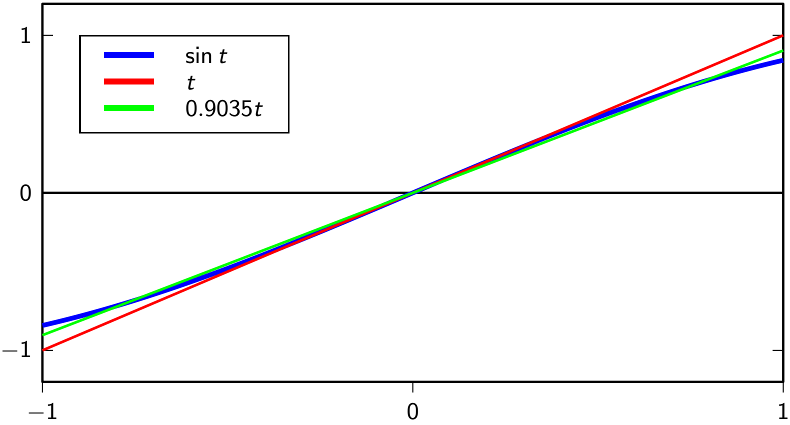
Using the orthogonal projection over $P_2[-1,1]$:

$$\sin t \to \alpha_1 \mathbf{u}_1 \approx 0.9035\, t$$

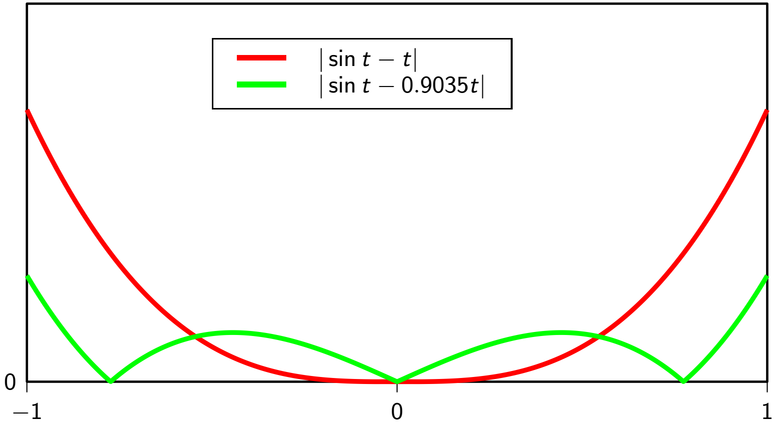
Using Taylor's series:

$$\sin t \approx t$$

Sine approximation



Approximation error



Error norm

Orthogonal projection over $P_2[-1, 1]$:

$$\|\sin t - \alpha_1 \mathbf{u}_1\| \approx 0.0337$$

Taylor series:

$$\|\sin t - t\| \approx 0.0857$$

Hilbert space

Hilbert Space – the ingredients:

1. a vector space: $H(V, \mathbb{C})$
2. an inner product: $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$
3. completeness

Completeness

limiting operations must yield vector space elements

Example of an *incomplete* space: the set of rational numbers

$$x_n = \sum_{k=0}^{n} \frac{1}{k!} \in \mathbb{Q} \qquad \text{but} \qquad \lim_{n\to\infty} x_n = e \notin \mathbb{Q}$$

Signals in Hilbert Space

Why did we do all this?

- finite-length and periodic signals live in $\mathbb{C}^N$
- infinite-length signals live in $\ell_2(\mathbb{Z})$

- different bases are different observation tools for signals
- subspace projections are useful in filtering and compression