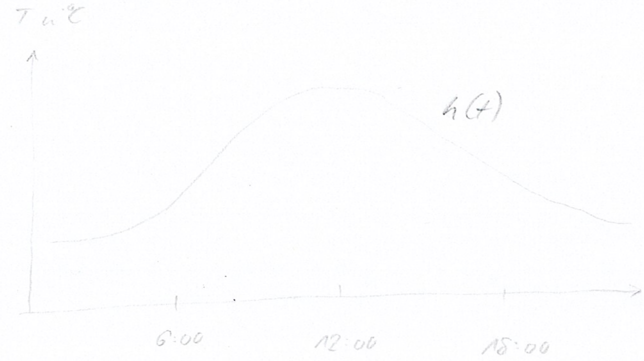
11.4.2011

Function Spaces

- So far: vector (or linear) spaces were familiar Euclidean spaces $\mathbb{R}^n$

  → Many more objects (collections of objects) have the same essential structure

- Ex



We used vectors to describe the velocity of particles

Now we would like to describe the temperature evolution over a day

→ Arguably simplest description of $h(t)$: polynomial

- $h(t) = mt + n$
- $h(t) = at^2 + bt + c$

Why does one consider linear, quadratic models, e.g. for linear regression?

↳ Two (or three) parameters to describe entire family of curves

$$h(t) = \sum_{i=0}^{n} a_i t^i$$

$$= \underbrace{a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n}_{}$$

h is fully specified by the coefficients a_i w.r.t. fixed t^i

→ $h(t)$ is a weighted linear combination of the t^i

→ This is very similar to how a vector $v \in \mathbb{R}^n$ is written as a linear combination of basis vectors, i.e.

$$v = v_1 \vec{e}_1 + v_2 e_2$$

Scalar weights

basis vectors

This suggests the hypothesis:

The space of functions that can be written as

$$\underline{f(x) = \sum_{i=1}^{n} a_i x^i} \quad , a_i < \infty$$

is a vector space.

Let us recall the final definition of a vector space that we developed in the last lecture based on its original applications: the description of velocities of point particles

<u>Def</u>: A <u>linear</u> or <u>vector</u> space V over $\mathbb{R}$ (or $\mathbb{C}$) is a set together with addition

$$+ : V \times V \to V$$

and scalar multiplication

$$\cdot : \mathbb{R} \times V \to V$$

that satisfies the following properties:

i) Associativity of addition: $u + (v+w) = (u+v) + w$

ii) Commutativity of addition: $u + v = v + u$

iii) Identity element of addition: $\exists! \, 0 \in V$ s.t. $v + 0 = v$

iv) Inverse element: $\forall v \in V$: there exists $-v \in V$ s.t. $v + (-v) = 0$

Compatibility between $(\mathbb{R}, \cdot)$ and vector space

v) Associativity of scalar mult: $a(bv) = (ab)v$

vi) Identity of scalar mult: $1 \cdot v = v$

vii) Distributivity of scalar mult. $a(u+v) = au + av$

$(a+b)u = au + bu$

→ If the above hypothesis is correct then the space of polynomials has to satisfy the properties of a vector space

↳ But first we need a definition of addition and scalar multiplication

0.i) Addition

$$f(x) + g(x) = \left(\sum_{i=0}^{n} f_i x^i\right) + \left(\sum_{j=0}^{n} g_j x^j\right)$$

$$= \left(f_0 + f_1 x + f_2 x^2 + \dots\right)$$

$$+ \left(g_0 + g_1 x + g_2 x^2 + \dots\right)$$

$$= \underbrace{\sum_{i=1}^{n} (f_i + g_i) x^i}_{\text{This is again a polynomial of degree } n}$$

0.ii) Scalar multiplication:

$$a f(x) = a \sum_{i=1}^{n} f_i x^i$$

$$= \underbrace{\sum_{i=1}^{n} (a f_i) x^i}_{\text{Again a polynomial of degree } n}$$

With this we can verify the properties of a vector space:

i) Associativity of addition

↳ Follows from standard rules of arithmetic

ii) Commutativity of addition

↳ Follows from standard rules of arithmetic

iii) Identity element: $\vec{0} = \sum_{i=1}^{n} 0 x^i$

iv) Inverse element: $-f(x) = \sum_{i=1}^{n} (-f_i) x^i = -\sum_{i=1}^{n} f_i x^i$

v) Associativity of scalar multiplication

↳ Follows from standard rules of arithmetic

vi) Identity of scalar mult: $1 f(x) = f(x)$

vii) Distributivity: $a(f(x) + g(x)) = a f(x) + a g(x)$

$(a+b) f(x) = a f(x) + b f(x)$

↳ Follows from standard rules of arithmetic

→ The space of polynomials of degree n,

$$\mathcal{P}^n = \{ f(x) = \sum_{i=1}^{n} a_i x^i \}$$

forms a vector space. The space is n-dimensional

We constructed the space with the help of a basis, the so called monomial basis

$$\{ x^i \}_{i=1}^{n}$$

Analogous to $\mathbb{R}^n$, an element in $\mathcal{P}^n$ is fully characterized by the coefficients f_i, i.e

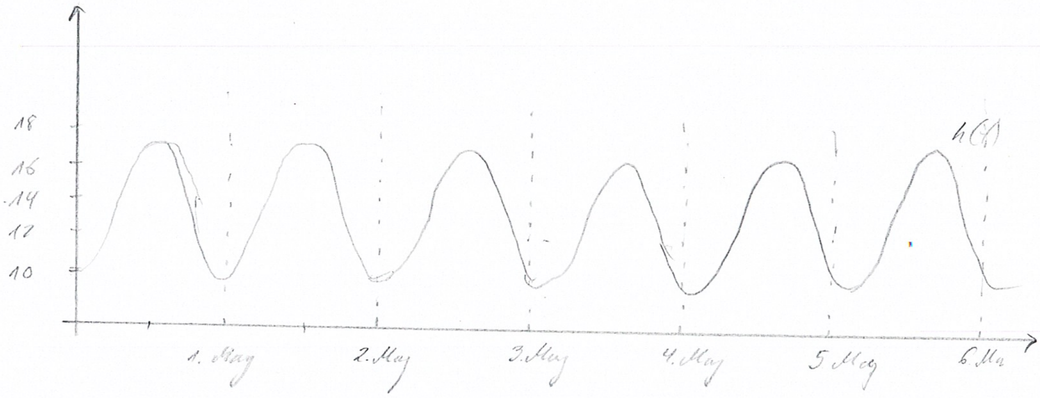
$$f(x) = f_0 + f_1 x + f_2 x^2 + \dots \mathrel{\hat{=}} \begin{pmatrix} f_0 \\ \vdots \\ f_n \end{pmatrix}$$

We can represent a continuous function on a computer (up to floating point errors)!

→ Crucial: $\mathcal{P}^n$ is finite dimensional and (our) computers are finite machines

→ Is the example of polynomials special or have more classes of functions a vector space structure?

Second example: let us consider temperature over several days:



→ Natural to describe data using harmonics

$$\sin(2\pi m x), \quad \cos(2\pi m x)$$

↳ m controls the frequency of oscillations

- $m=2$: two oscillations in $[0,1]$
- $m=3$: three oscillations in $[0,1]$
- ...

→ What frequency m should we use?

↳ Data will also not be perfectly harmonic and we would like to be able to describe - and measure - how significant this deviation is.

↳ Polynomial representation: let coefficients "decide" which parts match signal

→ This is the vector space principle!

"We only" choose a maximum capacity.

⇒ Given the previous examples, let us consider representations of the form

$$h(t) = h_0 + \sum_{m=1}^{M} h_m \cos(2\pi m t) + \sum_{n=1}^{N} h_n \sin(2\pi n t)$$

weights

"vectors"

→ Hypothesis: Functions that admit above representation form a vector space, $h_i < \infty$

Let us see if this hypothesis holds true. Again, we first need to define "addition" and "scalar multiplication"

- Addition:

$$f(x) + g(x) = \left(f_0 + \sum_m f_m \cos(2\pi m x) + \sum_n f_n \sin(2\pi n x)\right)$$
$$+ \left(g_0 + \sum_m g_m \cos(2\pi m x) + \sum_n g_n \sin(2\pi n x)\right)$$
$$= (f_0 + g_0) + \sum_m (f_m + g_m) \cos(2\pi m x) + \sum (f_n + g_n) \sin(2\pi n x)$$

- Scalar multiplication:

$$a f(x) = a \cdot \left(f_0 + \sum_n \ldots + \sum_n \ldots\right)$$
$$= (a f_0) + \sum_n (a f_m) \cos(2\pi m x) + \sum_n (a f_n) \sin(2\pi n x)$$

Verification of vector space properties

i) Associativity of addition: $u + (v + w) = (u + v) + w$

↳ Standard rules of arithmetic

ii) Commutativity of addition

↳ Standard rules of arithmetic

⋮

⇒ The space of functions of the form:

$$f(x) = f_0 + \sum_{m=1}^{N} f_m \cos(2\pi m x) + \sum_{n=1}^{N} f_n \sin(2\pi n x), \quad N < \infty$$

forms a vector space

↳ Again, a function is completely specified by the coefficients f_n ⇒ amenable to a computer implementation

↳ 1_x, $\cos(2\pi m x)$, $\sin(2\pi m x)$ are our basis "vectors" (functions)

→ Our temperature function $h(t)$ is represented using the coefficients

$$h = (h_0, h_{m1}, h_{n1})^T$$

↳ How do we get the coefficients actually?

↳ How good is this representation?

↳ For example, we could also have used polynomials

→ How do we get the coefficients in $\mathbb{R}^n$?

↳ Model for vector space

$$\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2$$

$$v_i = v \circ \vec{e}_i = \sum_{j=1}^{2} v_j e_j^i$$

→ How to generalize the dot product?

- Coefficient representation is pointwise scalar multiplication for each $\underbrace{\text{"degree of freedom"}}$

Basis functions vary from point to point

→ Do pointwise multiplication at every point

→ Then the sum has to become an integral

⇒ Hypotheses: A continuous generalization of the dot product is

$$\langle f, g \rangle = \int_0^T f(x)\, g(x)\, dx$$

→ Has to at least satisfy properties of inner or dot product:

<u>Def</u>. An inner product on a vector space V over $\mathbb{R}$ is a bilinear mapping

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

satisfying ($u, v \in V$, $a, b \in \mathbb{R}$)

i.) Linearity: $\langle u+v, w \rangle = \langle u, v \rangle + \langle u, w \rangle$
$\langle au, bv \rangle = ab \langle u, v \rangle$

ii) Symmetry $\langle u, v \rangle = \langle v, u \rangle$

iii) Positive definiteness $\langle u, u \rangle \geq 0$
$\langle u, u \rangle = 0 \Rightarrow u = 0$

Verification of properties for candidate:

i) Linearity

↳ Follows immediately from linearity of integral

ii) Symmetry

↳ Follows immediately from definition

iii) Positive definiteness:

$$\langle f, f \rangle = \int_0^T (f(x))^2 \, dx \geq 0$$

$\Rightarrow f(x) = 0$ for integral to vanish

<u>Def</u>. The <u>L_2 inner product</u> over $X \subset \mathbb{R}^n$ is

$$\langle f, g \rangle = \int_X f(x) \, g(x) \, dx$$

for $f, g : X \to \mathbb{R}$.

↳ This provides a generalization of the dot product to continuous functions. Does this also yield the coefficients for the harmonic representation of $h(x)$?

- Basic validity check:
  - $h(x) = \sin/\cos(2\pi k x)$ then only the corresponding coefficient should be non-trivial

$$h_0 = \int_0^1 h(x)\,dx$$
$$= \int_0^1 \cos(2\pi k x)\,dx$$
$$= 0$$

$$h_m = \int_0^1 \cos(2\pi k x)\cos(2\pi m x)\,dx \; :$$

$$\cos(\alpha)\cos(\beta) = \frac{1}{2}\cos(\alpha-\beta) + \cos(\alpha+\beta)$$

$$= \frac{1}{2}\left(\int_0^1 \cos(2\pi x(k-m))\,dx + \int_0^1 \cos(2\pi x(k+m))\,dx\right)$$

a.) $k = m$: $h_m = \frac{1}{2}$ since $\cos(0) = 1$

b.) $k \neq m$ $h_m = 0$

$$h_n = \int_0^1 \cos(2\pi k x)\sin(2\pi n x)\,dx$$

$$\cos\alpha \sin\beta = \frac{1}{2}\sin(\alpha+\beta) - \sin(\alpha-\beta)$$

$$= \frac{1}{2}\left(\int_0^1 \sin(2\pi x(n+k))\,dx - \int_0^1 \sin(2\pi x(m-k))\,dx\right)$$

$$= 0$$

→ This extends by linearity to any $h(x)$ that is a linear combination of basic harmonics

↓

But what happens when $h(x)$ cannot exactly be written in this way?

↳ How large should M, N be?

↳ If $M, N = \infty$, can we represent any $h(x)$?

↳ How to compute integral numerically?