Gradient Descent in Hilbert Space

Carson James

December 6, 2021

Outline

1. Banach Spaces
   - bounded linear maps
   - Frechet differentiation
2. Calculus
   - tools
   - results
3. Hilbert Spaces
   - Riesz representation theorem
   - gradients
4. Convex Analysis
   - results for Frechet differentiable functions
5. Reproducing Kernel Hilbert Spaces
   - representer theorem
6. Applications to Gaussian Processes
   - predictive posterior

Banach Spaces

Definition

Let X, Y be a normed vector spaces and $T : X \to Y$ a linear map. Then T is said to be **bounded** if there exists $C \geq 0$ such that for each $x \in X$,

$$\|Tx\| \leq C\|x\|$$

We define

$$L(X, Y) = \{T : X \to Y : T \text{ is linear and bounded}\}$$

Definition

Let $X_1, \cdots, X_n$ and Y be a normed vector spaces and $T : \prod_{j=1}^{n} X_j \to Y$ a multilinear linear map. Then T is said to be **bounded** if there exists $C \geq 0$ such that for each $(x_j)_{j=1}^n \in \prod_{j=1}^{n} X_j$,

$$\|T(x_1, \cdots, x_n)\| \leq C\|x_1\| \cdots \|x_n\|$$

We define

$$L^n\left(\prod_{j=1}^{n} X_j, Y\right) = \{T : X \to Y : T \text{ is multilinear and bounded}\}$$

If $X_1, \cdots, X_n = X$, we write $L^n(X, Y)$ in place of $L^n(\prod_{j=1}^n X_j, Y)$.

Remark

Let X and Y be normed vector spaces. We may identify $L(X, L(X, \cdots, L(X, Y)) \cdots)$ and $L^n(X, Y)$ via the isometric isomorphism given by $\phi \mapsto \psi_\phi$ where

$$\psi_\phi(x_1, x_2, \cdots, x_n) = \phi(x_1)(x_2), \cdots, (x_n)$$

Definition

Let X be a normed vector space over $\mathbb{R}$. We define the **dual space of** X, denoted X^*, by $X^* = L(X, \mathbb{R})$. Let $T : X \to \mathbb{R}$. Then T is said to be a **bounded linear functional on** X if $T \in X^*$.

Definition

Let X be a normed vector space. Then X is said to be a **Banach space** if X is complete.

Definition

Let X, Y be a banach spaces, $A \subset X$ open, $f : A \to Y$ and $x_0 \in A$. Then f is said to be **Frechet differentiable at** x_0 if there exists $Df(x_0) \in L(X, Y)$ such that,

$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + o(\|h\|) \qquad \text{as } h \to 0$$

If f is Frechet differentiable at x_0, we define the **Frechet derivative of** f **at** x_0 to be $Df(x_0)$. We say that f is **Frechet differentiable** (or 1-**st order Frechet differentiable**) if for each $x_0 \in A$, f is Frechet differentiable at x_0.
If f is Frechet differentiable, we define the **Frechet derivative** of f, denoted $Df : A \to L(X, Y)$, by

$$x \mapsto Df(x)$$

Definition

Let X, Y be a banach spaces, $A \subset X$ open, $f : A \to Y$. We define *n*-**th** order Frechet differentiablility inductively.
Since f

Calculus

Remark

The various tools used to obtain the main calculus results are the following:

- Frechet derivative
- Hahn-Banach theorem (not introduced)
- Bochner Integral (not introduced)

Convex Analysis

Result

The various tools used to obtain the main convex analysis results are the following:

-
- *Frechet derivative*
- *Hahn-Banach theorem (not introduced)*
- *Bochner Integral (not introduced)*