Calculus on Normed Vector Spaces

Joydeep Medhi
2013MT60599

Supervisor: Amit Priyadarshi

Dept. of Mathematics
Indian Institute of Technology

Midterm Presentation 2

February 21, 2018

Overview

Aim

- To study the notion of derivatives on general normed vector spaces and do Calculus on them.
- To generalise the basic calculus of function of several variables to Normed Vector Spaces.'
- To explore the applications of these concepts.

Quick overview of previous work

- Normed Vector Spaces
- Directional and Partial Derivatives
- Differentiation and Mean Value Theorems
- Higher order Differentials and Derivatives
- Taylor Theorems and Applications (Extrema)

Hilbert Space

A vector space E over the field F together with an inner product, i.e., with a map

$$< .,. > : E \text{ x } E \to F$$

which follows the three axioms for all vectors $\in E$

- Positive-definiteness
- Linearity in first argument
- Conjugate symmetry

Then the pair $(E, <, >)$ is called an inner product space.

for $x \in E$, $\|x\|$ is defined as $\sqrt{< x, x >}$.

If E is an inner product space and complete with the norm derived from the inner product, Then E is said to be a **Hilbert Space**

The Riesz Representation Theorem

If H is a Hilbert space, then by the Riesz representation theorem, we may associate an element of H to a continuous linear form.

Before looking at the general Theorem, let us see what happens in $\mathbb{R}^n$. If l is a linear form defined on $\mathbb{R}^n$, (e_i) its standard basis and $x = \sum_{i=1}^{n} x_i e_i$ then

$$l(x) = \sum_{i=1}^{n} x_i l(e_i) = x.w$$

where $w = (l(e_1), ..., l(e_n))$.

If $\overline{w}$ is such that $l(x) = x.\overline{w}$ for all $x \in \mathbb{R}^n$, then $x.(w - \overline{w}) = 0$ for all $x \in \mathbb{R}^n$, it follows that $w - \overline{w} = 0$. Hence, the element w such that $l(x) = x.w$ for all x is **unique**.

The Riesz Representation Theorem

Theorem Let l be a continuous linear form defined on a Hilbert space H . Then there is a unique element $a \in H$ such that

$$l(x) = < x, a >$$

for all $x \in H$. In addition, $|l|_{H^*} = \|a\|$.

he Riesz Representation Theorem

If f is a real-valued mapping defined on an open subset O of a Hilbert space H and is differentiable at a point $x \in O$, then $f'(x)$ is a continuous linear form and so, from Theorem, there is a unique element $a \in H$ such that

$$f'(x)h = < h, a >$$

for all $h \in H$. We call **a** the gradient of f at x and write $\nabla f(x)$ for a. If f is differentiable on O, then we obtain a mapping ∇f from O into H to which we also give the name gradient.

Remark. If f has a second differential at a point $x \in O$, then ∇f is differentiable at x. If f is of class C^2 on O, then ∇f is of class C^1 on O.

Convex functions

Let X be a convex subset of a vector space V . We say that $f : X \to R$ is **convex** if for all $x, y \in X$ and $\lambda \in (0, 1)$ we have

$$f\left(\lambda x + (1 - \lambda)y\right) \leq \lambda f(x) + (1 - \lambda)f(y)$$

If this inequality is strict when $x \neq y$, then we say that f is **strictly convex**.

Convex Hull

Let E be a vector space and $x_1, \ldots, x_n \in E$. We say that $y \in E$ is a *convex combination* of the points $x_1, \ldots, x_n$ if there exists $\lambda_1, \ldots, \lambda_n \in [0, 1]$ with $\sum_{i=1}^{n} \lambda_i = 1$ such that $y = \sum_{i=1}^{n} \lambda_i x_i$.

If X is a nonempty subset of E, then we define *co* X , the convex hull of X , to be the set of points $y \in E$ which are convex combinations of points in X.

Defination. A subset X of a vector space E is convex if and only if

$$co\, X = X.$$

Continuity of Convex function

All convex functions are not continuous.

For example, if we define f on $[0,1]$ by

$$f(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}$$

then f is convex, but not continuous. However, f is continuous on the interior $[0,1]$.

Continuity of Convex functions

Lemma If P is a bounded nonempty polyhedron in a finite-dimensional normed vector space E and f is a convex function defined on E, then f has an upper bound on P .

Theorem Let X be a finite dimensional normed vector space E and $f : X \to \mathbb{R}$ is convex. If $x \in intX$, then f is continuous at x.
Corollary If a convex function is defined on an open subset of a finite-dimensional normed vector space, then it is continuous.

Differentiable Convex Functions

Let O be an open subset of a normed vector space E and f a real-valued differentiable functions defined on O. If $X \subset O$ is convex and $x, y \in X$, then the following are equivalent:

- f is convex on X.
- $f(y) - f(x) \geq f'(x)(y - x)$
- $(f'(y) - f'(x))(y - x) \geq 0$

Remark. From third condition we deduce that, if f is differentiable on an open interval I of $\mathbb{R}$, then f is convex (resp. strictly convex) if and only if the derivative of f is increasing (resp. strictly increasing) on I.

Differentiable Convex Functions

Example Let $A \in M_n(\mathbb{R})$ be symmetric, $b \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to R$ be defined by

$$f(x) = \frac{1}{2}x^t Ax - b^t x.$$

Then

$$\begin{aligned} f(y)-f(x)-f'(x)(y-x) &= \tfrac{1}{2}y^t Ay - b^t y - \tfrac{1}{2}x^t Ax - b^t x - (Ax-b)^t(y-x) \\ &= \tfrac{1}{2}y^t Ay + \tfrac{1}{2}x^t Ax - x^t Ay \\ &= \tfrac{1}{2}(y-x)^t A(y-x) \end{aligned}$$

It follows that f is convex (resp. strictly convex) if and only if the matrix A is positive (resp. positive definite).

Differentiable Convex Functions

Let O be an open subset of a normed vector space E and f a real-valued 2-differentiable function defined on O. For $x \in O$ and $h \in E$ we set

$$Q_{f(x)}(h) = f^{(2)}(x)(h, h)$$

$Q_{f(x)}$ is quadratic form.

Theorem. Let O be an open subset of a Normed Vector Space E, $X \subset O$ convex and $f : O \to \mathbb{R}$ 2-differentiable. Then

- f is convex on X, if and only if $Q_{f(x)}$ is positive for all $x \in X$.
- f is strictly convex on X, if $Q_{f(x)}$ is positive definite for all $x \in X$.

Differentiable Convex Functions

Example A function f may be strictly convex without the quadratic form Q_f being positive definite at all points.For example, if f is the real-valued function defined on $\mathbb{R}$ by $f(x) = x^4$ then $Q_f(0) = 0$. However,

$$\begin{aligned}(x+h)^4 - x^4 &= x^4 + 4x^3h + 6x^2h^2 + 4xh^3 + h^4 - x^4 \\ &= f'(x)h + h^2(6x^2 + 4xh + h^2) \\ &= f'(x)h + h^2(2x^2 + (2x+h)^2) > f'(x)h\end{aligned}$$

if $h \neq 0$. Therefore f is strictly convex.

The Inverse Mapping Theorem

Suppose that E and F are normed vector spaces and that O and U are open subsets of E and F respectively.

A function $f : O \to U$ is a diffeomorphism if f is bijective and both f and f^{-1} are differentiable. Also, we say that f is a C^k -diffeomorphism if both f and f^{-1} are C^k -mappings.

Results. If f is a diffeomorphism, then at any point x in its domain, $f'(x)$ is invertible.
In addition, for f to be a C^k -diffeomorphism it is sufficient that f be of class C^k.

The Inverse Mapping Theorem

Proposition. Let E and F be Banach spaces, $O \subset E$ and $U \subset F$ open sets and $f : O \to U$ a differentiable homeomorphism. If $a \in O$ and $f'(a)$ is invertible then f^{-1} is differentiable at $b = f(a)$.
In addition, if f is of class C^1 then there is a open neighbourhood O' such that $f_{|O'}$ is a C^1-diffeomorphism onto its image.

Theorem. Let E and F be Banach spaces, $O \subset E$ and $f : O \to F$ of class C^1. If $a \in O$ and $f'(a)$ is invertible, then there is an open neighbourhood O' of such that $f_{|O'}$ is a C^1-diffeomorphism onto its image.

The Inverse Mapping Theorem

Remarks.

- Under the conditions of the theorem, $f_{|O'}$ is a C^1-diffeomorphism onto its image. Infact if f is of class C^k, then $f_{|O'}$ is a C^k-diffeomorphism.
- If a mapping f is such that each point in its domain has an open neighbourhood O such that f restricted to O defines a diffeomorphism onto its image, then we say that f is a **local diffeomorphism**.

The Inverse Mapping Theorem

Example.

Consider the mapping

$$f : \mathbb{R}^2 \ (0,0) \to \mathbb{R}^2, (r,\theta) \mapsto (rcos\theta, rsin\theta)$$

Then

$$J_f(r,\theta) = \begin{bmatrix} cos\theta & -rsin\theta \\ sin\theta & rcos\theta \end{bmatrix}$$

and **det** $J_f(r,\theta) = r \neq 0$. It follows that $f'(r,\theta)$ is invertible for all $(r,\theta) \in \mathbb{R}^2$.

The continuity of the entries in the Jacobian matrix imply that f is a C^1-mapping. Hence f is a *local diffeomorphism*. However, f is not bijective and so not a *diffeomorphism*.

Future Work

- Implicit Mapping and Rank Theorem
- Theory of Vector Fields
- Application of these concepts in various fields (Optimization/Machine Learning etc.)

References

Rodney Colman(2012)
Calculus on Normed Linear Spaces

Avez, A.(1986)
Differential Calculus. J. Wiley and Sons Ltd, New York (1986)

Thank You!

The End