The Fréchet Derivative

Jacob Hauck

Math 6418

Outline and goals

- Introduce Fréchet derivative
  - Basic properties
- Some examples of Fréchet derivatives
  - Relationship with finite-dimensional derivatives
- Important theorems
  - Chain rule
  - Mean value theorem
- Partial Fréchet derivatives

Motivation

Let X, Y be normed vector spaces. We know a lot about bounded, linear operators $A \in B(X, Y)$.

What about nonlinear operators $f : X \to Y$?

Linearize:

$$f(x+h) \approx f(x) + Ah, \qquad A \in B(X,Y), \quad h \text{ “small enough”}$$

From calculus, we know that

$$\frac{f(x+h) - f(x) - f'(x)h}{h} \to 0 \qquad \text{as } h \to 0.$$

Generalize to arbitrary X, Y?

Definition of Fréchet derivative

Definition: Fréchet Derivative

Let $U \subset X$ be open and $f : U \to Y$. Then f is **Fréchet differentiable at** $x \in U$ if there exists $A \in B(X, Y)$ such that

$$\frac{\|f(x+h) - f(x) - Ah\|_Y}{\|h\|_X} \to 0 \quad \text{as} \quad h \to 0 \text{ in } X.$$

In this case, A is called the **Fréchet derivative of f at** x, and is also denoted

$$A = A_x = Df(x) = f'(x).$$

This reduces to the usual derivative if $X = Y = \mathbb{R}$.

Basic Properties

Let f and g be Fréchet differentiable at $x \in U$, and let $\alpha, \beta \in \mathbb{F}$. Then

1. Df is unique,
   - Let A, B both be derivatives. Show $A = B$ via $\|A - B\| = 0$.
2. $D(\alpha f + \beta g)(x) = \alpha Df(x) + \beta Dg(x)$ (linearity),
3. f is continuous at x (with respect to $\|\cdot\|_Y$ and $\|\cdot\|_X$),
   - Add and subtract $Df(x)h$, triangle inequality.
4. f is **locally Lipschitz** at x. That is, there is $\delta > 0$ and $L > 0$ such that

   $$\|h\|_X < \delta \implies \|f(x+h) - f(x)\|_Y \leq L\|h\|_X.$$

   Moreover, given $\varepsilon > 0$, we can take $L = \|Df(x)\|_{B(X,Y)} + \varepsilon$ (maybe need to take δ smaller)

Examples – Linear operators and “quadratic” operators

Example 1. Let $f(x) = Ax$, where $A \in B(X, Y)$ (f is linear). Then $Df(x) = A$ for all $x \in X$.

Example 2. Let $X = H$, a Hilbert space over $\mathbb{R}$. Suppose that $f(x) = (x, Ax)$, where $A \in B(X, X)$. Then $Df(x) = (A^* + A)x$ for all $x \in X$.

- Rearrange inner products
- Cauchy-Schwarz inequality
- Boundedness of A

Examples – C^1, finite-dimensional maps

Example 3. Let $f : \mathbb{R}^n \to \mathbb{R}$, where $f \in C^1$ (so $\partial_i f$ continuous).

- Guess: $Df(x)h = \nabla f(x)^T h$
- $n = 2$ case one coordinate at a time
- Use continuity of $\partial_i f$

Example 4. Let $f : \mathbb{R}^n \to \mathbb{R}^m$, where $f \in C^1$ (so $\partial_i f_j$ is continuous).

- f is a set of m functions from Example 3:

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$$

- Only finitely many components

Examples – Function space

Example 5. Let $p > 1$ be an integer, and let $f : L^p(\mathbb{R}) \to \mathbb{R}$ be defined by

$$f(\varphi) = \int_{\mathbb{R}} \varphi^p, \qquad \varphi \in L^p(\mathbb{R})$$

- Binomial theorem on $(\varphi + h)^p$
- Hölder's inequality on higher order terms (magic happens)
- Check that derivative is bounded and linear

Chain rule

In calculus, the chain rule involves the product of derivatives. Fréchet derivatives are operators – product of operators?

Theorem: Chain Rule for Fréchet Derivatives

Let X, Y, Z be normed vector spaces, $U \subset X$ and $V \subset Y$ open. Suppose that $f : U \to Y$, $g : V \to Z$.

If f is Fréchet differentiable at $x \in U$ and g is Fréchet differentiable at $f(x) \in V$, then $g \circ f$ is Fréchet differentiable at x, with

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$$

- Add/subtract $Dg(f(x))[f(x+h) - f(x)]$ – linear approximation of $g(f(x+h)) - g(f(x))$
- Add/subtract $Df(x)h$ to introduced $f(x+h) - f(x)$ – linear approximation of $f(x+h) - f(x)$
- Differentiability of f and boundedness of $Dg(f(x))$
- Multiply and divide by $\|f(x+h) - f(x)\|_Y$, differentiability of g

Mean value theorem

Recall: if f is differentiable on (a, b) **and continuous on** $[a, b]$, then there exists $c \in (a, b)$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

Again, we can't divide by a vector... MVT often used to bound $|f(b) - f(a)|$ in terms of derivative:

$$|f(b) - f(a)| \leq |f'(c)|(b - a)$$

Generalizing this, we can say

Theorem: Mean Value Theorem for Fréchet Derivatives

Let f be Fréchet differentiable on U.

If $\ell = \{tx_2 + (1 - t)x_1 \mid 0 \leq t \leq 1\} \subset U$ then

$$\|f(x_2) - f(x_1)\|_Y \leq \sup_{x \in \ell} \|Df(x)\| \cdot \|x_2 - x_1\|_X.$$

Proof of mean value theorem

- Focus on case $X = \mathbb{R}$, and $[x_1, x_2] = [0, 1]$
  - Let $\varphi : [0, 1] \to Y$ be Fréchet differentiable on $(0, 1)$ and continuous on $[0, 1]$. Then

$$\|\varphi(1) - \varphi(0)\|_Y \leq \sup_{0<t<1} \|D\varphi(t)\|.$$

  - Local Lipschitz property and compactness of $[0, 1]$ to construct a partition of a subinterval where change in function is almost controlled by derivative between partition points
  - Expand difference between endpoints in telescoping sum
  - Use continuity to take limit as subinterval endpoints approach full interval
- Chain rule to extend to the general case

Partial Fréchet derivatives

The partial derivative we know involves restricting a function to one coordinate direction – what to do about the abstract input space X?

- Use directional derivative (**Gateaux derivative**)
- Partition X into finitely many subspaces

Let $X_1, X_2, \dots, X_n$ be normed vector spaces, and let

$$X = X_1 \oplus X_2 \oplus \cdots \oplus X_n.$$

Note that there are two equivalent ways to think about this direct sum:

$$\begin{array}{l} X_k \subseteq X, \quad X_j \cap X_k = \{0\} \text{ if } j \neq k, \\ \operatorname{span}\{X_1, X_2, \dots, X_n\} = X, \\ \|\cdot\|_{X_k} = \|\cdot\|_X\big|_{X_k} \end{array} \iff \begin{array}{l} X = X_1 \times X_2 \cdots \times X_n \\ \|x\|_X = \|(\|x_1\|_{X_1}, \|x_2\|_{X_2}, \dots, \|x_n\|_{X_n})\| \\ \text{where } \|\cdot\| \text{ is any norm on } \mathbb{R}^n. \end{array}$$

The latter will be useful for developing the partial Fréchet derivative.

Definition of partial Fréchet derivatives

Definition: Partial Fréchet Derivatives

For $x = (x_1, x_2, \ldots, x_n) \in U$, define

$$f_{k,x}(z) = f(x_1, \ldots, x_{k-1}, z, x_{k+1}, \ldots, x_n)$$

on $U_k = \{z \in X_k \mid (x_1, \ldots, z, \ldots, x_n) \in U\}$, which is open in X_k.

If $f_{k,x}$ is Fréchet differentiable at x_k, then f is **partially Fréchet differentiable along X_k at x** with **partial Fréchet derivative** $D_k f : X \to B(X_k, Y)$ given by

$$D_k f(x) = D f_{k,x}(x_k).$$

Differentiable at x implies partially differentiable at x

- Lifting from X_k to X is differentiable. Apply chain rule.

Having all partial derivatives $\implies$ differentiable

Theorem: Fréchet "Gradient"

Suppose that $D_k f(x)$ exists for all $x \in U$, and $D_k f$ is continuous at $x_0 \in U$. Then f is Fréchet differentiable at x_0, and

$$Df(x_0)h = \sum_{k=1}^{n} D_k f(x_0) h_k, \qquad h = (h_1, \dots, h_n) \in X.$$

- Show that proposed derivative is bounded and linear
- One "coordinate axis" at a time
- Mean value theorem on each coordinate + continuity of partial derivatives
- Equivalence of $\|\cdot\|_X$ and ℓ^1 norm of X_k norms