

2.6. Differentiability in Banach spaces

To be able to characterise optimal controls and optimal states, similar as in the finite-dimensional case (Chapter 1.3), we require a generalisation of the concept of differentiability to Banach spaces.

In this section, let U and V denote real Banach spaces and $\mathcal{U} \subset U$ a non-empty, open subset of U . Further, let $f: \mathcal{U} \rightarrow V$.

Definition: (i) Let $u \in \mathcal{U}$ and $h \in U$ be given.

If the limit

$$\delta f(u, h) := \lim_{t \rightarrow 0} \frac{f(u+th) - f(u)}{t}$$

exists in V , then it is called first variation of f at u in the direction h .

(ii) If there exists a continuous linear operator (in h) $A: U \rightarrow V$ for fixed $u \in \mathcal{U}$, such that

$$\delta f(u, h) = Ah \quad \forall h \in U,$$

then f is called Gâteaux-differentiable at u with Gâteaux-derivative $f'_G(u) = A$.

Examples: (i) Evaluation of a function at a point

Let $U = \mathcal{U} = C([0, 1])$ and $f: U \rightarrow \mathbb{R}$ with $f(u(\cdot)) = \sin(u(1))$. Let $h \in C([0, 1])$, then

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} (f(u+th) - f(u)) &= \lim_{t \rightarrow 0} \frac{1}{t} (\sin((u+th)(1)) - \sin(u(1))) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} (\sin(u(1) + th(1)) - \sin(u(1))) = \frac{d}{dt} \sin(u(1) + th(1)) \Big|_{t=0} \\ &= \cos(u(1) + th(1)) h(1) \Big|_{t=0} = \cos(u(1)) h(1) = \delta f(u, h). \end{aligned}$$

The mapping $h(\cdot) \mapsto \cos(u(1))h(1)$ is linear and continuous with respect to $h \in C([0, 1])$, hence the Gateaux-derivative exists at any point $u \in U$ and it holds: $f'_G(u)h = \cos(u(1))h(1)$.

Note that here it is not possible to express f'_G without reference to the direction h .

(ii) Square of the norm in Hilbert spaces

Let $\{H, (\cdot, \cdot)_H\}$ be a real Hilbert space, equipped with natural norm $\|\cdot\|_H = \sqrt{(\cdot, \cdot)_H}$.

Consider $f: H \rightarrow \mathbb{R}$, $f(u) = \|u\|_H^2$.

It holds:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} (f(u+th) - f(u)) &= \lim_{t \rightarrow 0} \frac{1}{t} (\|u+th\|_H^2 - \|u\|_H^2) \\ &= \lim_{t \rightarrow 0} \frac{(u, u)_H + 2t(u, h)_H + t^2(h, h)_H - \|u\|_H^2}{t} \\ &= \lim_{t \rightarrow 0} \frac{2t(u, h)_H + t^2\|h\|_H^2}{t} = 2(u, h)_H, \end{aligned}$$

hence $f'_G(u)h = (2u, h)_H$.

If we use the Riesz representation theorem to identify $(2u) \in H$ with $(h \mapsto (2u, h)_H) \in H^*$, we can write $f'_G(u) = 2u$.

(iii) Application to the norm in $L^2(\Omega)$

By (ii), the Gateaux derivative of the functional

$$f(u) := \|u(\cdot)\|_{L^2(\Omega)}^2 = \int_{\Omega} |u(x)|^2 dx$$

is given by

$$f'_G(u)h = \int_{\Omega} 2u(x)h(x) dx \quad \forall h \in L^2(\Omega).$$

Identifying $(L^2(\Omega))^*$ with $L^2(\Omega)$, we have

$$f'_G(u)(x) = 2u(x).$$

Definition: $f: U \rightarrow V$ is Fréchet-differentiable at $u \in U$ if there exists a continuous linear operator $A: U \rightarrow V$ and a mapping $r(u, \cdot): U \rightarrow V$ such that for all $h \in U$ such that $u+h \in U$ we have

$$f(u+h) = f(u) + Ah + r(u, h)$$

where the remainder satisfies

$$\frac{\|r(u, h)\|_V}{\|h\|_U} \rightarrow 0 \text{ as } \|h\|_U \rightarrow 0.$$

$f'(u) = A$ is called Fréchet-derivative of f at u . If f is Fréchet-differentiable at every $u \in U$, then f is called Fréchet-differentiable in U .

Remarks:

(i) It is often convenient to prove

$$\frac{\|f(u+h) - f(u) - Ah\|_V}{\|h\|_U} \rightarrow 0 \text{ as } \|h\|_U \rightarrow 0.$$

(ii) Fréchet-differentiability \Rightarrow Gâteaux-differentiability (converse is false, see Tröltzsch book for an example)

(iii) If f is Fréchet-differentiable, $f'(u) = f'_G(u)$, hence the explicit form of a Fréchet-derivative can be determined by the Gâteaux-derivative, i.e. using the directional derivative.

(iv) Chain rule: Let $f: U \rightarrow V \subset V$ and $g: V \rightarrow Z$ be Fréchet-differentiable at $u \in U$ and $f(u) \in V$, respectively. Then $e = f \circ g: U \rightarrow Z$, $e(u) = g(f(u))$ is Fréchet-diff. at $u \in U$ and $e'(u) = g'(f(u))f'(u)$.

Examples: (i) $f(u) = \sin(u(1))$ is Fréchet-differentiable at every $u \in C([0,1])$ (Problem Set 3).

(ii) $f(u) = \|u\|_H^2$ is Fréchet-differentiable on every Hilbert space H (exercise)

(iii) Every continuous linear operator A is Fréchet-differentiable, since $A(u+h) = Au + Ah + r(u, h)$ holds with $r(u, h) = 0$. Therefore, the derivative of a continuous linear operator is given by the operator itself.

(iv) Chain rule: consider for some continuous linear operator $S: U \rightarrow H$ the functional $e: U \rightarrow \mathbb{R}$, $e(u) = \|Su - z\|_H^2$, where $z \in H$ is given, and U, H are real Hilbert spaces.

Let $g(v) = \|v\|_H^2$ and $f(u) = Su - z$.

Then by previous examples we have

$$g'(v)h = g'_G(v)h = (2v, h)_H \text{ and } f'(u)h = Sh.$$

$$\begin{aligned} \Rightarrow e'(u)h &= g'(f(u)) \underbrace{f'(u)h}_{\in H} = (2f(u), f'(u)h)_H \\ &= (2(Su - z), Sh)_H \\ &= 2(Su - z, Sh)_H \\ &= 2(S^*(Su - z), h)_U, \end{aligned}$$

where S^* is the so-called adjoint operator of S .

2.7. Adjoint operator.

The adjoint operator can be considered as the operator-equivalent to the transpose of a matrix in finite dimensions. In fact, if A is an $m \times n$ -matrix, then $(A^T u, v)_{\mathbb{R}^m} = (u, Av)_{\mathbb{R}^n}$ for all $u \in \mathbb{R}^m, v \in \mathbb{R}^n$.

Definition: Let $\{U, (\cdot, \cdot)_U\}, \{V, (\cdot, \cdot)_V\}$ be real Hilbert spaces, and $A: U \rightarrow V$ a continuous linear operator. An operator A^* is called the adjoint of A if $(v, Au)_V = (A^* v, u)_U$ for all $u \in U, v \in V$.

Note that $A^*: U \rightarrow V$ is also a continuous linear operator (cf. Tröltzsch book.).

Examples: (i) Let $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ denote a linear operator, represented by an $m \times n$ matrix A . Since $(v, Au)_{\mathbb{R}^m} = (A^T v, u)_{\mathbb{R}^n}$ for all $u \in \mathbb{R}^n, v \in \mathbb{R}^m$, the adjoint operator is $A^* = A^T$.

(ii) Consider in $L^2(0,1)$ the integral operator $(Au)(t) = \int_0^t e^{(t-s)} u(s) ds, t \in (0,1)$.

A is a continuous linear operator on $L^2(0,1)$ (cf. Problem Set 2). We have for all $u, v \in L^2(0,1)$:

$$\begin{aligned}
(v, Au)_{L^2(0,1)} &= \int_0^1 v(t) \left(\int_0^t e^{(t-s)} u(s) ds \right) dt \\
&= \int_0^1 \int_0^t v(t) e^{(t-s)} u(s) ds dt = \int_0^1 \int_s^1 v(t) e^{(t-s)} u(s) dt ds \\
&= \int_0^1 u(s) \left(\int_s^1 v(t) e^{(t-s)} dt \right) ds = \int_0^1 u(t) \left(\int_t^1 e^{(s-t)} v(s) ds \right) dt \\
&= (A^* v, u)_{L^2(0,1)} \text{ where } (A^* v)(t) = \int_t^1 v(s) e^{(s-t)} ds.
\end{aligned}$$

