Differential Calculus in Banach Spaces

Fréchet and Gâteaux Derivatives

We begin nonlinear analysis of operators with definitions of differentiation. Let $F(x)$ be a nonlinear operator acting from $D(F) \subset X$ to $R(F) \subset Y$, where X and Y are real Banach spaces. Assume $D(F)$ is open.

Definition 3.1.1. $F(x)$ is *differentiable in the Fréchet sense* at $x_0 \in D(F)$ if there is a bounded linear operator, denoted by $F'(x_0)$, such that

$$F(x_0 + h) - F(x_0) = F'(x_0)h + \omega(x_0, h) \text{ for all } \|h\| < \varepsilon$$

with some $\varepsilon > 0$, where $\|\omega(x_0, h)\|/\|h\| \to 0$ as $\|h\| \to 0$. Then $F'(x_0)$ is called the *Fréchet derivative* of $F(x)$ at x_0, and $dF(x_0, h) = F'(x_0)h$ is

its *Fréchet differential*. $F(x)$ is Fréchet differentiable in an open domain $S \subset D(F)$ if it is Fréchet differentiable at every point of S.

Example Let $\Omega \subset \mathbb{R}^n$ be an open bounded domain. Denote by $C(\overline{\Omega})$ the continuous function space on $\overline{\Omega}$. Let

$$\varphi : \overline{\Omega} \times \mathbb{R}^1 \longrightarrow \mathbb{R}^1$$

be a C^1 function. Define a mapping $f : C(\overline{\Omega}) \to C(\overline{\Omega})$ by

$$u(x) \mapsto \varphi(x, u(x)) \ .$$

Then f is F-differentiable, and $\forall u_0 \in C(\overline{\Omega})$,

$$(f'(u_0) \cdot v)(x) = \varphi_u(x, u_0(x)) \cdot v(x) \quad \forall v \in C(\overline{\Omega}) \ .$$

Proof. $\forall h \in C(\overline{\Omega})$

$$t^{-1}[f(u_0 + th) - f(u_0)](x) = \varphi_u(x, u_0(x) + t\theta(x)h(x))h(x) \ ,$$

where $\theta(x) \in (0, 1)$. $\forall \varepsilon > 0$, $\forall M > 0$, $\exists \delta = \delta(M, \varepsilon) > 0$ such that

$$| \varphi_u(x, \xi) - \varphi_u(x, \xi') | < \varepsilon, \ \forall x \in \overline{\Omega} \ ,$$

as $|\xi|$, $|\xi'| \leqslant M$ and $|\xi - \xi'| \leqslant \delta$. We choose $M = \| u_0 \| + \| h \|$, then for $|t| < \delta < 1$,

$$|\varphi_u(x, u_0(x) + t\theta(x)h(x)) - \varphi_u(x, u_0(x))| < \varepsilon \ .$$

It follows that $df(u_0, h)(x) = \varphi_u(x, u_0(x))h(x)$.

Problem Assume $\mathbf{y} = \mathbf{f}(\mathbf{x})$ is a vector function from $\mathbb{R}^m$ to $\mathbb{R}^n$ and $\mathbf{f}(\mathbf{x}) \in (C^{(1)}(\Omega))^n$. Show that its Fréchet derivative at $\mathbf{x}_0 \in \Omega$ is the Jacobi matrix $\left(\frac{\partial f_i(\mathbf{x}_0)}{\partial x_j}\right)_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$.

Definition Assume that for all $h \in D(F)$ we have

$$\lim_{t \to 0} \frac{F(x_0 + th) - F(x_0)}{t} = DF(x_0, h), \qquad x_0 \in D(F),$$

where $DF(x_0, h)$ is a linear operator with respect to h. Then $DF(x_0, h)$ is called the *Gâteaux differential* of $F(x)$ at x_0, and the operator is called *Gâteaux differentiable*. Denoting $DF(x_0, h) = F'(x_0)h$, we get the *Gâteaux derivative* $F'(x_0)$. An operator is differentiable in the Gâteaux sense in an open domain $S \subset X$ if it has a Gâteaux derivative at every point of S.

The definitions of derivatives are clearly valid for functionals. Suppose $\Phi(x)$ is a functional which is Gâteaux differentiable in a Hilbert space and that $D\Phi(x, h)$ is bounded at $x = x_0$ as a linear functional in h. Then, by the Riesz representation theorem, it can be represented in the form of an inner product; denoting the representing element by $\operatorname{grad} \Phi(x_0)$, we get

$$D\Phi(x_0, h) = (\operatorname{grad} \Phi(x_0), h).$$

By this, we have an operator grad $\Phi(x_0)$ called the *gradient* of $\Phi(x)$ at x_0.

Theorem If an operator $F(x)$ from X to Y is Fréchet differentiable at $x_0 \in D(F)$, then $F(x)$ is Gâteaux differentiable at x_0 and the Gâteaux derivative coincides with the Fréchet derivative.

Theorem *Suppose that* $f : U \to Y$ *is G-differentiable, and that* $\forall x \in U$, $\exists A(x) \in L(X, Y)$ *satisfying*

$$df(x, h) = A(x)h \quad \forall h \in X \ .$$

If the mapping $x \mapsto A(x)$ *is continuous at* x_0*, then* f *is F-differentiable at* x_0*, with* $f'(x_0) = A(x_0)$*.*

Proof. With no loss of generality, we assume that the segment $\{x_0 + th \mid t \in [0, 1]\}$ is in U. According to the Hahn–Banach theorem, $\exists y^* \in Y^*$, with $\| y^* \| = 1$, such that

$$\| f(x_0 + h) - f(x_0) - A(x_0)h \|_Y = \langle y^*, f(x_0 + h) - f(x_0) - A(x_0)h \rangle \ .$$

Let

$$\varphi(t) = \langle y^*, f(x_0 + th) \rangle \ .$$

From the mean value theorem, $\exists \xi \in (0,1)$ such that

$$\begin{aligned}
|\ \varphi(1) - \varphi(0) - \langle y^*, A(x_0)h \rangle \ | &= |\ \varphi'(\xi) - \langle y^*, A(x_0)h \rangle \ | \\
&= |\ \langle y^*, df(x_0 + \xi h, h) - A(x_0)h) \rangle \ | \\
&= |\ \langle y^*, [A(x_0 + \xi h) - A(x_0)]h \rangle \ | \\
&= \circ(\| \ h \ \|) \ ,
\end{aligned}$$

i.e., $f'(x_0) = A(x_0)$.

Example Let X be a Hilbert space, with inner product $(,)$. Find the F- derivative of the norm $f(x) = \| \ x \ \|$, as $x \neq \theta$.

High-Order Derivatives

The second-order derivative of f at x_0 is defined to be the derivative of $f'(x)$ at x_0. Since $f' : U \to L(X, Y)$, $f''(x_0)$ should be in $L(X, L(X, Y))$. However, if we identify the space of bounded bilinear mappings with $L(X, L(X, Y))$, and verify that $f''(x_0)$ as a bilinear mapping is symmetric, see Theorem 1.1.9 below, then we can define equivalently the second derivative $f''(x_0)$ as follows: For $f : U \to Y$, $x_0 \in U \subset X$, if there exists a bilinear mapping $f''(x_0)(\cdot, \cdot)$ of $X \times X \to Y$ satisfying

$$\| f(x_0+h)-f(x_0)-f'(x_0)h-\frac{1}{2}f''(x_0)(h,h) \|= \circ(\| h \|^2) \quad \forall h \in X, \text{as} \| h \|\to 0 ,$$

then $f''(x_0)$ is called the second-order derivative of f at x_0.

By the same manner, one defines the mth-order derivatives at x_0 successively: $f^{(m)}(x_0) : X \times \cdots \times X \to Y$ is an m-linear mapping satisfying

$$\left\| f(x_0 + h) - \sum_{j=0}^{m} \frac{f^{(j)}(x_0)(h, \ldots, h)}{j!} \right\| = \circ(\| h \|^m) ,$$

as $\| h \|\to 0$. Then f is called m differentiable at x_0.

Theorem 1.1.9 *Assume that $f : U \to Y$ is m differentiable at $x_0 \in U$. Then for any permutation π of $(1, \ldots, m)$, we have*

$$f^{(m)}(x_0)(h_1, \ldots, h_m) = f^{(m)}(x_0)(h_{\pi(1)}, \ldots, h_{\pi(m)}) .$$

Theorem *(Taylor formula) Suppose that $f : U \to Y$ is continuously m-differentiable. Assume the segment $\{x_0 + th \,|\, t \in [0,1]\} \subset U$. Then*

$$f(x_0 + h) = \sum_{j=0}^{m} \frac{1}{j!} f^{(j)}(x_0)(h, \ldots, h) + \frac{1}{m!} \int_0^1 (1-t)^m f^{(m+1)}(x_0 + th)(h, \ldots, h) dt .$$

Example $X = \mathbb{R}^n$, $Y = \mathbb{R}^1$. If $f : X \to Y$ is twice continuously differentiable, then

$$f''(x) = H_f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n} .$$

Theorem (Mean Value Theorem)

If $f\colon X \longrightarrow Y$ *is a Gâteaux differentiable function and* $x, h \in X$, $y^* \in Y^*$, *then we can find* $\lambda_0 \in (0,1)$, *such that*

$$\langle y^*, f(x+h) - f(x) \rangle_Y \;=\; \langle y^*, f'_G(x+\lambda_0 h)h \rangle_Y$$

and

$$\left\| f(x+h) - f(x) \right\|_Y \;\leqslant\; \left\| f'_G(x+\lambda_0 h) \right\|_{\mathcal{L}} \|h\|_X \,.$$

THEOREM (Implicit Function Theorem)

If X, Y, Z *are three Banach spaces,* $U \subseteq X \times Y$ *is an open set,* $(x_0, y_0) \in U$, $f\colon U \longrightarrow Z$ *is a continuous differentiable function,* $f(x_0, y_0) = 0$ *and*

$D_2 f(x_0, y_0) \in \mathcal{L}(X;Y)$ *is invertible with a continuous inverse,*

i.e., $D_2 f(x_0, y_0)$ *is an isomorphism,*
then there exist neighbourhoods U_1 *of* x_0 *and* U_2 *of* y_0, *such that* $U_1 \times U_2 \subseteq U$ *and a unique continuously differentiable function* $g\colon U_1 \longrightarrow U_2$, *such that*

$$f\big(x, g(x)\big) \;=\; 0 \qquad \forall\, x \in U_1$$

and

$$Dg(x) \;=\; -\Big(D_2 f\big(x, g(x)\big)\Big)^{-1} D_1 f\big(x, g(x)\big) \qquad \forall\, x \in U_1.$$

consider an operator equation with a parameter μ being an element of a real Banach space M:

$$F(x, \mu) = 0$$

where $D(F(x, \mu)) \subseteq X$, $R(F(x, \mu)) \subseteq Y$.

Theorem Assume an operator $F(x)$ from X to Y has a Fréchet derivative at $x = x_0$, and an operator $x = S(z)$ from a real Banach space Z to X also has a Fréchet derivative $S'(z_0)$ and $x_0 = S(z_0)$. Then their composition $F(S(z))$ has a Fréchet derivative at $z = z_0$ and

$$(F(S(z_0)))' = F'(x_0)S'(z_0).$$

Proof. Substituting

$$x - x_0 = S(z) - S(z_0) = S'(z_0)(z - z_0) + \omega_1(z_0, z - z_0)$$

into

$$F(x) - F(x_0) = F'(x_0)(x - x_0) + \omega(x_0, x - x_0),$$

we get

$$F(x) - F(x_0) = F'(x_0)S'(z_0)(z - z_0) + F'(x_0)\omega_1(z_0, z - z_0) + \\ + \omega(x_0, S(z) - S(z_0)).$$

Lyapunov–Schmidt Reduction

Let X, Y be Banach spaces, and let $\wedge$ be a topological space. Assume that $F : U \times \wedge \to Y$ is continuous, where $U \subset X$ is a neighborhood of θ. We assume that $F_x(\theta, \lambda_0)$ is a Fredholm operator, i.e.,

(1) $\text{Im } F_x(\theta, \lambda_0)$ is closed in Y,
(2) $d = \dim \ker F_x(\theta, \lambda_0) < \infty$,
(3) $d^* = \text{codim Im } F_x(\theta, \lambda_0) < \infty$.

Set

$$X_1 = \ker F_x(\theta, \lambda_0),\ Y_1 = \text{Im } F_x(\theta, \lambda_0)\ .$$

Since both dim X_1, and codim Y_1 are finite, we have the direct sum decompositions:

$$X = X_1 \oplus X_2,\ Y = Y_1 \oplus Y_2\ ,$$

and the projection operator $P : Y \to Y_1$. $\forall x \in X$, there exists a unique decomposition:

$$x = x_1 + x_2,\ x_i \in X_i,\ i = 1, 2\ .$$

Thus

$$F(x, \lambda) = \theta \Leftrightarrow \begin{cases} PF(x_1 + x_2, \lambda) = \theta \ , \\ (I - P)F(x_1 + x_2, \lambda) = \theta \ . \end{cases}$$

Now, $PF_x(\theta, \lambda_0) : X_2 \to Y_1$ is a surjection as well as an injection. According to the Banach theorem, it has a bounded inverse. If we already have $F(\theta, \lambda_0) = \theta$, then from the IFT, we have a unique solution

$$u : V_1 \times V \to V_2$$

satisfying

$$PF(x_1 + u(x_1, \lambda), \lambda) = \theta \ ,$$

where V_i is a neighborhood of θ in $U \cap X_i, i = 1, 2$, and V is a neighborhood of λ_0.

It remains to solve the equation:

$$(I - P)F(x_1 + u(x_1, \lambda), \lambda) = \theta$$

on $V_1 \times V$. This is a nonlinear system of d variables and d^* equations.

The above procedure is called the Lyapunov–Schmidt reduction.

Definition (x_0, μ_0) is a *bifurcation point* of the equation $F(x, \mu) = 0$ if for any $r > 0$, $\rho > 0$, in the ball $\|\mu - \mu_0\| \leq \rho$ there exists μ such that in the ball $\|x - x_0\| \leq r$ there are at least two solutions of the equation corresponding to μ.

Example *The point $(0, 0)$ is a bifurcation point for the ordinary differential equation*

$$u'' + \lambda(u + u^3) = 0$$

subject to the periodic boundary conditions

$$u(0) = u(2\pi), \quad u'(0) = u'(2\pi).$$

Theorem *If the point $(0, \lambda_0)$ is a bifurcation point for the equation*

$$F(u, \lambda) = 0,$$

then the Fréchet derivative $F_u(0, \lambda_0)$ cannot be a linear homeomorphism of X to Y.