Optimization II

Lecture 8, Feb. 1

4. Fréchet and Gateaux derivatives

Notation for $r : U \subset X \to Y$

$$r(x) = o(\|x\|) \iff \frac{r(x)}{\|x\|} \xrightarrow{\|x\| \to 0} 0$$

$$r(x) = o(1) \iff r(x) \xrightarrow{\|x\| \to 0} 0$$

For Banach spaces X, Y, we define

$$L(X, Y) = \{ f : X \to Y \mid f \text{ linear and continuous} \}$$

Def. 12: let $f : U(x_0) \subset X \to Y$ and X and Y are Banach spaces, then there holds:

i) f is Fréchet differentiable (F-diffb) in x_0

$\iff$ there exists $T \in L(X, Y)$ s.t.

$$f(x_0+h) - f(x_0) = Th + o(\|h\|)$$

T is called F-derivative of f at x_0,

$f'(x_0) := T$

(ii) f is Gateaux diffb. (G-diffb.)

$\Longleftrightarrow$ $\exists\; T \in L(\mathcal{X}, \mathcal{Y})$, s.th.

$$f(x_0 + tk) - f(x_0) = tTk + o(t)$$

for all $k \in \mathcal{X}$ with $\|k\| = 1$.

T is called the Gateaux derivative.

(iii) If the F (- or G) derivative exists for all $x_0 \in \mathcal{U}$, then

$$f' : \mathcal{U} \subset \mathcal{X} \to L(\mathcal{X}, \mathcal{Y}),$$

$$x \longmapsto f'(x)$$ is called

F (or G) derivative of f in $\mathcal{U}$

Examples

1. $F : C[0,1] \to \mathbb{R}, \quad y \mapsto \int_0^1 (\sin^3 x + y^2(x))\, dx$

Claim: F is F-diff. with

$$F'(y)h = 2\int_0^1 y(x)h(x)\, dx$$

Proof:

$$F(y+h) - F(y) - F'(y)h =$$

$$= \int_0^1 [(y+h)^2 - \cancel{y^2} - 2\cancel{y}h]\, dx$$

$$= \int_0^1 h^2\, dx \leq \|h\|_\infty^2 = o(\|h\|_\infty)$$

2. Let H be a Hilbert space and

$$F : H \to \mathbb{R}, \quad u \mapsto \|u\|_H^2$$

$$= (u,u)_H$$

$$F(u+h) - F(u) = (u+h, u+h)_H - (\cancel{u,}\, u)$$

$$= 2(u,h)_H + (h,h)$$

$$= 2(u,h)_H + \|h\|^2$$

$$= (2u,h)_H + o(\|h\|)$$

$\Rightarrow$ F is F-diff. with

$$F'(u)h = (2u,h)_H$$

$F'(u) : H \to \mathbb{R}$ i.e. $F' \in H^*$

by identification using Riesz

$$\Rightarrow \quad F'(u) = 2u$$

3. let $F : X \to Y$ linear and continuous, then F is F-diffb. with F-derivative $F' = F$ because

$$F(x+h) - F(x) = F(h) + 0$$

Remark: It is easy to check that

$$F\text{-diffb.} \Longrightarrow G \text{ diffb.}$$

the converse in general is not true.

Theorem 13 (Chain rule)

Let X, Y, Z be Banach spaces,

$F : U(x_0) \subset X \to U(f(x_0)) \subset Y$

$G : U(f(x_0)) \longrightarrow Z$ and let

$F'(x_0)$ and $G'(F(x_0))$ exist as

F-derivatives.

Then $H(x) = G(F(x))$ is F-diffb.

with $H'(x_0) = G'(F(x_0))\, F'(x_0)$

$$= G'(F(x_0)) \circ F'(x_0)$$

Example:

$$f(u) = J(y(u), u) = \frac{1}{2} \int_\Omega (y - y_d)^2 \, dx$$

$$= \frac{1}{2} (y - y_d, y - y_d)_{L^2(\Omega)}$$

where $S : L^2(\Omega) \to L^2(\Omega)$, $u \mapsto Su = y$

linear and continuous, hence $S' = S$

and $F(u) = Su - y_d$

$$G(y) = \frac{1}{2} \|y\|^2_{L^2(\Omega)}$$

$$F'(u) h = S h$$

$$G'(y) h = (y, h)_{L^2(\Omega)}$$

then $f'(u)h = G'(F(u)) F'h$

$$= (F(u), F'h)_{L^2} = (F(u), Sh)_{L^2}$$

$$= (Su - y_d, Sh)_{L^2}$$

$$= (S^*(Su - y_d), h)_{L^2}$$

applying Riesz theorem, we get

$$\boxed{f'(u) = S^*(Su - y_d)}$$

Remark: For $T \in L(X, Y)$ with Hilbert spaces X and Y, we define the adjoint operator $T^* \in L(Y, X)$ by $(y, Tx)_Y = (T^* y, x)_X$

5. Necessary optimality conditions

We begin with considering

$$(P) \quad \min_{u \in \mathcal{U}_{ad}} f(u) \quad , \mathcal{U}_{ad} \subset \mathcal{U}$$

Lemma 14: Let $\mathcal{U}_{ad} \subset \mathcal{U}$ be closed and convex, $f : \mathcal{U} \to \mathbb{R}$ G-diffb. and $\bar{u}$ be a sol to (P). Then, there holds

$$f'(\bar{u})(u - \bar{u}) \geq 0 \quad \forall u \in \mathcal{U}_{ad}$$

Proof: Let $u \in \mathcal{U}_{ad}$ then

$$\bar{u} + t(u - \bar{u}) \in \mathcal{U}_{ad} \text{ for all } t \in [0,1]$$

since f is G-diff., we obtain for $t > 0$, small enough

$$0 \leq \frac{f(\bar{u} + t(u - \bar{u})) - f(\bar{u})}{t}$$

$$\Big\downarrow \; t \searrow 0$$

$$f'(\bar{u})(u-\bar{u}) \geq 0 .$$

Now, we consider (CP1):

$$\min J(y,u)$$

$$\text{subj. to} \quad a_0(y,\varphi) = \int_\Omega u \varphi \, dx \quad \forall \varphi \in H_0^1(\Omega)$$

$$\text{and} \quad u \in \mathcal{U}_{ad} \subset L^2(\Omega)$$

We define reduced cost functional

$$f(u) = J(y(u),u) = \frac{1}{2} \int_\Omega (Su - y_d)^2 dx + \frac{\gamma}{2} \int_\Omega u^2 dx$$

Using previous examples, we get

$$f'(u)h = (Su - y_d, Sh)_{L^2(\Omega)} + \gamma (u,h)_{L^2}$$

$$= (S^*(Su - y_d), h)_{L^2} + \gamma (u,h)_{L^2}$$

Now, we want to characterize S^*, the adjoint solution operator.

Lemma 16 For the state equation

$$\begin{cases} -\Delta y = u & \text{in } \Omega \\ \quad y = 0 & \text{on } \partial\Omega \end{cases}$$

the Hilbert space adjoint $S^*: L^2(\Omega) \to L^2(\Omega)$ is defined as $S^* z = p$, where p is weak sol to

$$(*) \begin{cases} -\Delta p = z, & \text{in } \Omega \\ \quad p = 0, & \text{on } \partial\Omega \end{cases}$$

i.e., S^* is the sol. operator $z \mapsto S^* z = p$ of equ. $(*)$

Proof: S^* is defined as

$$(z, Su) = (S^* z, u) \quad \forall z, u \in L^2(\Omega)$$

Then there holds

$$(1) \quad \int_\Omega \nabla p \cdot \nabla v \, dx = \int_\Omega z v \, dx \quad \forall v \in H_0^1(\Omega)$$

$$\qquad\qquad\qquad\qquad\qquad\qquad \|$$

$$(2) \quad \int_\Omega \nabla y \cdot \nabla \varphi \, dx = \int_\Omega u \varphi \, dx \quad \forall \varphi \in H_0^1(\Omega)$$

take $v = y$ in (1) and $\varphi = p$ in (2)

$$\int_\Omega z y \, dx = (z, Su) = \int_\Omega p u \, dx = (S^* z, u)$$

$\square$

Now consider $S^*(Su - y_d) = S^* z$,

i.e. we choose $z = Su - y_d = y - y_d$,

then $S^*(Su - y_d) = p$, where p solves the adjoint eqn.

$$(AE) \begin{cases} -\Delta p = y - y_d & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega \end{cases}$$

Corollary 17: For $F(u) = J(y(u), u)$ we have $f'(u) = p + \gamma u$

Proof:

$$f'(u)h = (\underbrace{S^*(Su - y_d)}_{= p} + \gamma u, h)_{L^2}$$

$$= (p + \gamma u, h)_{L^2}$$

Thus $\Longrightarrow$ $f'(u) = p + \gamma u$

Now let $\bar{u}$ be anol to (CP1) and $h \in L^2(\Omega)$ s.th. $\bar{u} + h \in \mathcal{U}_{ad}$, then we have

$$0 \leq f'(\bar{u})h \quad , \quad \text{for } h = u - \bar{u}$$

we get

$$(p + \gamma \bar{u}, u - \bar{u})_{L^2} \geq 0$$

$$\forall u \in \mathcal{U}_{ad}$$