Some Comments on Derivatives

→ $\theta(\upsilon): \mathcal{U} \to \mathbb{R}$

- $\mathcal{U}$ can be something as simple as a mean, something as complex as a density

→ $D_\upsilon \theta(\upsilon_0)$ [often interested in derivative]

- Delta method (asym. CIs)
- Lower bounds on efficiency (LAM)

→ Types of derivatives

- Gateaux:

$$\lim_{t \downarrow 0} \frac{f(x_0 + td) - f(x_0)}{t} = df(x_0)(d)$$

- Hadamard:

$$\lim_{\substack{t \downarrow 0 \\ d' \to d}} \frac{f(x_0 + td') - f(x_0)}{t} = \nabla f(x_0)(d)$$

- Fréchet

$$\lim_{\|d\| \to 0} \left\| \frac{f(x_0 + d) - f(x_0) - Df(x_0)(d)}{\|d\|} \right\| = 0$$

→ Linear vs. Directional

- Linear $\nabla f(x_0)(d)$ (linear operator of d) (df, Df)
- Directional $\nabla f(x_0)(d)$ (1-homogeneous, cont. if Hadamard/Fréchet) (df, Df)

→ Fréchet + Hadamard coincide in finite-dim space $\mathcal{U}$, Gateaux + Hadamard "closer" in infinite dims (Zajíček 2015)

~> Focus: Hadamard directional deriv.

- Close conn. to pathwise deriv. of semiparam. (Bickel et al. 1993)
- "Right" deriv for delta-method (Van der Vaart + Wellner 2023)

→ Full linear deriv.

- Directional delta-method (Romisch 2004)
- Efficiency lower bounds for dir. diff. fnctls (Ponomarev 2022)
- Will focus on $\bar{x}(\upsilon_0)$ but happy to discuss $v(\upsilon_0)$ [much easier to analyze]

Consider optimal solution to the problem

(P_υ) $\min_{G(x) \in K} f(x, \upsilon)$

$f: \mathcal{X} \times \mathcal{U} \to \mathbb{R}$

$G: \mathcal{X} \to \mathcal{Y}$

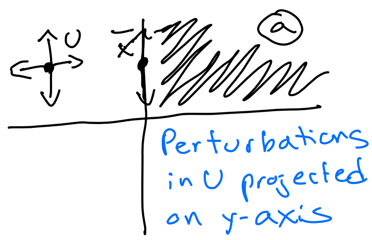
$\mathcal{X}, \mathcal{Y}, \mathcal{U}$ Banach Spaces

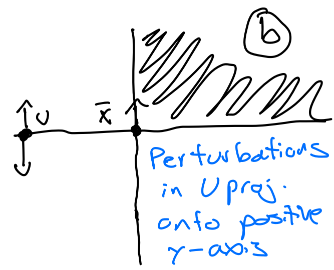
$K \subseteq \mathcal{Y}$ Closed, Convex

- Extremely general... nonlinear programming, semi-def. prog., conic prog.
- Can consider the optimal value $v(\upsilon_0)$ as well as optimal solution $\bar{x}(\upsilon_0)$
- "Since the optimal value function... As well as the optimal solutions... typically are non differentiable... we concentrate on directional analysis" [pg. 4 PAOP]
- $G(x)$ can be augmented to $G(x, \upsilon)$

- Will introduce some examples of (P_υ)
- (Positive Projection)

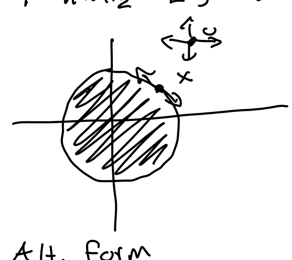
$$\min_{x \in [0, \infty)^2} \|x - \upsilon\|_2^2 \quad (1)$$





- (Ball projection)

$$\min_{1 - \|x\|_2^2 \in [0, \infty)} \|x - \upsilon\|_2^2 \quad (2)$$



Alt. form

$$\min_{x \in B(0,1)} \|x - \upsilon\|_2^2$$

Key Q: Can we express all these derivatives in a unified way?

- Essential Ingredients:

$\to$ Tangent Cone

$$T_K(x) = \lim_{t \downarrow 0} \frac{K-x}{t}$$

$\to$ Normal Cone

$$N_K(x) = T_K(x)^{-}$$
$$= \{ y^* \in \mathcal{X}^* : \langle y^*, h \rangle \le 0 \ \ \forall h \in T_K(x) \}$$

$\to$ Lagrange Multipliers

$$\{ \lambda \in N_K(x) : -Df(x,u) = \langle \lambda, DG(x) \rangle \}$$
$$= \Lambda(x,u)$$

$\to$ Critical Cone

$$C(x_0) = \{ g : DG(x_0)\, g \in T_K(x_0) \\ \langle \lambda, DG(x_0)(g) \rangle = 0 \}$$

$$\lambda \in \Lambda(x_0, u_0)$$

Can think of critical cone as "potential first-order" perturbation of x_0... perturbation by small g to x_0 leaves x_0 in K, in direction f is first-order unchanging

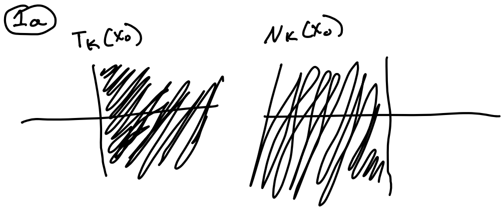
In the examples we've considered so far...

$D\bar{x}(u_0)(d)$ is the soln to

$$\min_{g \in C(x_0)} g^T g - 2 g^T d + g^T \langle \lambda, D^2 G(x_0) \rangle g$$

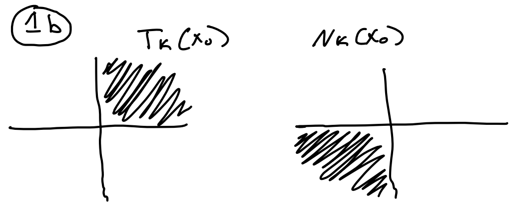
Illustrations of central role of critical cone

(1a)



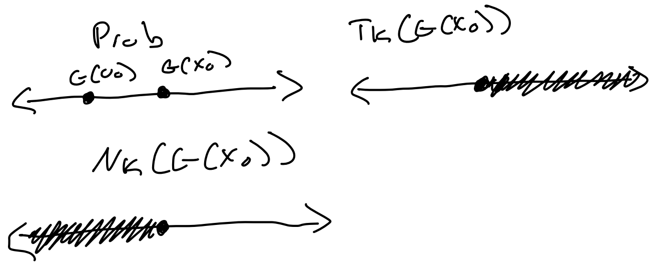
$$-2(x_0 - u_0) = \lambda$$
$$C(x_0) = \{ g : g \in T_K(x_0),\ \langle \lambda, g \rangle = 0 \}$$
$$= y\text{-axis}$$
$$D\bar{x}(u_0)(d) : \min_{g \in \{0\} \times \mathbb{R}} g^T g - 2 g^T d = (0, d_{(2)})$$

(1b)



$$-2(x_0 - u_0) = \lambda$$
$$C(x_0) = \{ g : g \in T_K(x_0),\ \langle \lambda, g \rangle = 0 \}$$
$$= \{0\} \times [0, \infty)$$
$$D\bar{x}(u_0)(d) : \min_{g \in \{0\} \times [0,\infty)} g^T g - 2 g^T d = (0, [d_{(2)}]^+)$$

(2) Assume $\|u_0\| > 1$



$$\lambda = 1 - \|u_0\|$$

$$C(x_0) = \{ g : x_0^T g = 0 \}$$

$$\min_{g : x_0^T g = 0} g^T g - 2 g^T d - \lambda g^T g$$

Thm: Assume $\mathcal{X}$ is a reflexive Banach space, there exists a unique optimal soln x_0 to (Not directly in PAOP, but similar to Thm. 4.133, Thms 4.94, 4.95)

$$\min_{G(x) \in K} f(x_0, u_0)$$

f, G twice cont. diff. Assume:

i) extended polyhedricity condition holds at x_0

ii) Robinson's constraint qualification holds at x_0

iii) For every $\lambda \in \Lambda(x_0, u_0)$
$h \mapsto D_{xx}^2 L(x_0, \lambda, u_0)(h, h)$ is an extended Legendre form

iv) $D_{xx}^2 L(x_0, \lambda, u_0)(g, g) > 0$
$\forall g \in Sp(C(x_0))$

Then for any $o(\|u - u_0\|^2)$-optimal soln $\bar{x}(u_t)$ + path $u_t = u_0 + td + o(t)$

$$\frac{\bar{x}(u_t) - \bar{x}(u_0)}{t} \to g(d)$$

where $g(d)$ solves

$$\min_{g \in C(x_0)} \sup_{\lambda \in \Lambda(x_0, u_0)} D^2 L(x_0, \lambda, u_0)((g, d), (g, d))$$

Second-order tangent sets

- Inner:
$$T_K^{i,2}(x, h) = \liminf_{t \downarrow 0} \frac{K - x - th}{\frac{1}{2} t^2}$$
- Outer,
$$T_K^2(x, h) = \limsup_{t \downarrow 0} \frac{K - x - th}{\frac{1}{2} t^2}$$

- $T_K^{i,2}(x, h), T_K^2(x, h)$ empty if
$$K = \{(x_1, x_2) : x_2 \geq |x_1|^{3/2}\}$$
$$h = (1, 0)$$

(See example 3.29)

- $\exists$ function $\eta(x)$ st, for every $k \in \mathbb{N}$, on (x_k, x_{k+1}), $\eta(x)$ is linear, $\eta(x_k) = x_k^2$, line thru $\eta(x_{k+1})$, $\eta(x_k)$ is tangent to $2x^2$, and $\eta(x) = \eta(-x)$
$\eta(0) = 0$

$T_K^{i,2}(0, h) = \{(x_1, x_2) : x_2 \geq 4\}$

$h = (1, 0)$ $T_K^2(0, h) = \{(x_1, x_2) : x_2 \geq 2\}$

(see example 3.31)

(Note although K closed convex, $T_K^{i,2}(0, h) \neq T_K^2(0, h)$ + not a cone)

Robinson's Constraint Qualification

- RCQ holds for K at x_0 if
$$0 \in \operatorname{int}\{G(x_0) + D_x G(x_0)\mathcal{X} - K\}$$

- Consequences:
  - $\to$ If $x_0 \in G^{-1}(K)$, then for all x in a neighborhood of x_0,
  $$\operatorname{dist}(x_0, G^{-1}(K)) = O(\operatorname{dist}(G(x), K))$$
  (Thm 2.87)
  - $\to$ If $G(x)$ is cont. diff at $x_0 \in G^{-1}(K)$ + RCQ holds, then
  $$T_{G^{-1}(K)} = \{h \in \mathcal{X} : DG(x_0)(h) \in T_K(G(x_0))\}$$
  (Corollary 2.91)
  - $\to$ If $K = \{0\}^q \times (-\infty, 0]^{p-q}$
  $Dg_i(x_0)$, $i = 1 \ldots q$ linearly indep.
  $\exists h \in \mathcal{X}$ st
  $Dg_i(x_0)h = 0$, $i = 1 \ldots q$ linearly indep.
  $Dg_i(x_0)h < 0$ $\forall i \in q+1 \ldots p$ st $g_i(x_0) = 0$
  Known as Mangasarian-Fromowitz Constraint Qualification
  (eq. 2.191)
  - $\to$ Let Ω be a nonempty compact Hausdorff topological space + let $Y = C(\Omega)$ and
  $K = \{y \in C(\Omega) : y(\omega) \leq 0 \ \forall \omega \in \Omega\}$
  Let $G : X \to Y$ be continuously diff + $G(x_0) \in K$. Let
  $\Delta(x_0) = \{\omega \in \Omega : G(x_0)(\omega) = 0\}$

Since K has nonempty interior, RCQ is equiv to

$$G(x_0) + DG(x_0)h \in \operatorname{int} K.$$

(Lemma 2.99)

Follows from compactness of Ω that if $[D_x G(x_0)h](\omega) < 0 \ \forall \omega \in \Delta(x_0)$, then $\sup_{\omega \in \Delta(x_0)} [D_x G(x_0)h](\omega) < 0$ and so

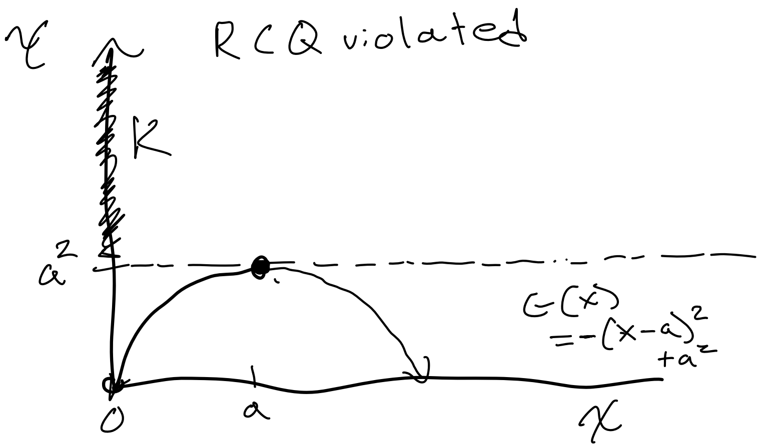
RCQ holds iff

$\exists h \in X$ st $D_x G(x_0)h(\omega) < 0$ $\forall \omega \in \Delta(x_0)$

Extended MFCQ

(Example 2.102)

RCQ violated



$G^{-1}([a^2, \infty)) = \{a\}$ $\quad K = [a^2, \infty)$

As $x \to a$, $\operatorname{dist}(x, G^{-1}(K)) = |x-a|$

$\operatorname{dist}(G(x), K) = (x-a)^2$

So $\nexists\, c > 0$ st $\operatorname{dist}(x, G^{-1}(K)) \le c \operatorname{dist}(G(x), K)$

RCQ holds, but $DG(x_0)$ not full rank

$G: \mathbb{R}^2 \to \mathbb{R}^2$, $G(x,y) = (0,y)$ $\quad K = B(0,1)$

$DG(0,1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, but at $(0,1)$,

$h = (0,-1) \Rightarrow G(0,1) + DG(0,1)h = (0,0) \in \operatorname{int} K$

$\to$ If RCQ holds, x_0 is locally optimal)

$\Lambda(x_0, u_0)$ is a nonempty, convex, bounded, weak*-compact subset of $\mathcal{Y}^*$

Proof Sketch

Use 2nd-order expansion of opt. value function

$$\limsup_{t \downarrow 0} \frac{v(u_t) - v(u_0) - tDf(x_0,u_0)(u_t - u_0)}{\frac{1}{2}t^2} \quad (\Delta)$$

to derive path

$$\bar{x}(u_t) = \bar{x}(u_0) + tD\bar{x}(u_0)(d) + o(t)$$

Want to upper bound (Δ) & show every path $\bar{x}(u_t)$ matches upper bound

(A) Upper Bound (Δ).

- Consider $g \in C(x_0)$, $w \in T^2_{G^{-1}(K)}(x_0, g)$

$\Rightarrow \exists\ x_0 + tg + \frac{1}{2}t^2 w^2 + o(t^2) \in G^{-1}(K)$

$$(\Delta) \le \inf_{g \in C(x_0)} \inf_{w \in T^2_{G^{-1}(K)}(x_0,g)} D_x f(x_0,u_0)w + D^2 f(x_0,u_0)((g,d),(g,d))$$

Can show (Assuming RCQ) that

$$T^2_{G^{-1}(K)}(x_0, g) = DG(x_0)^{-1}\left[T^2_K(G(x_0), DG(x_0)(g)) - D^2G(x_0)(g,g)\right]$$

$$\Rightarrow \inf_{w \in T^2_{G^{-1}(K)}(x_0,g)} \left[D_x f(x_0,u_0)(w) + D^2 f(x_0,u_0)((g,d),(g,d)) \right]$$

$$(D) \quad = \inf_{w : DG(x_0)w + D^2G(x_0)(g,g) \in T^2_K(G(x_0), DG(x_0)g)} \left[D_x f(x_0,u_0)(w) + D^2 f(x_0,u_0)((g,d)(g,d)) \right]$$

It can be shown that

$$\left[T^2_K(G(x_0), DG(x_0)(g))\right]^\infty = \operatorname{cl}\{T_K(G(x_0)) + \operatorname{sp} DG(x_0)(g)\}$$

$\Rightarrow \sigma(\lambda, T^2_K(G(x_0), DG(x_0)(g))) = \infty$ iff $\lambda \notin N_K(G(x_0))$, $\langle \lambda, DG(x_0)(g) \rangle \neq 0$

$$\Lambda(x_0) = \{\lambda \in \mathcal{Y}^* : \sigma(\lambda, T^2_K(G(x_0), DG(x_0)(g))) < \infty\}$$

From Lagrangian duality, (Eq. 2.305)

(D) dual:
$$\sup_{\lambda^* \in \mathcal{Y}^*} \inf_{w} \left[D_x f(x_0,u_0)(w) + D^2 f(x_0,u_0)((g,d)(g,d)) + \langle \lambda^*, DG(x_0)w + D^2G(x_0)(g,g) \rangle - \sigma(\lambda^*, T^2_K(G(x_0), DG(x_0)(g))) \right]$$

- Note if $\langle \lambda^*, DG(x_0)\rangle \neq -D_x f(x_0,u_0)(w)$

  then expression $\to -\infty$

- (P) dual

$$= \sup_{\lambda \in \mathcal{Y}^*} \left[D^2 L(x_0,\lambda,u_0)((g,d),(g,d)) - \sigma\left(\lambda, T_K^2(G(x_0), DG(x_0)g)\right)\right]$$

$$= \sup_{\lambda\in\Lambda(x_0,u_0)} \left[D^2 L(x_0,\lambda,u_0)((g,d),(g,d)) - \sigma\left(\lambda, T_K^2(G(x_0), DG(x_0)g)\right)\right]$$

Can be shown no gap b/w primal dual under RCQ. Also can be shown

$$\sigma\left(\lambda, T_K^2(G(x_0), DG(x_0)(g))\right) \leq 0 \quad \text{[See Thm 4.84]}$$

so if $0 \in T_K^2(G(x_0), DG(x_0)g)$,

$$\sigma\left(\lambda, T_K^2(G(x_0), DG(x_0)(g))\right) = 0$$

$$\Rightarrow (\Delta) \leq \inf_{g\in C(x_0)} \sup_{\lambda\in\Lambda(x_0,u_0)} D^2 L(x_0,\lambda,u_0)((g,d),(g,d))$$

Lower Bound

$$f(x,u) - f(x_0,u_0) \geq L(x,\lambda,u) - L(x_0,\lambda,u_0)$$

Since $G(x) - G(x_0) \in T_K(G(x_0))$ and $\lambda \in N_K(G(x_0))$,

$$\langle \lambda, G(x) - G(x_0)\rangle \leq 0$$

$$\begin{aligned} v(u_t) &= f(\bar{x}(u_t), u_t) + o(t^2) \\ &\geq f(x_0,u_0) + L(\bar{x}(u_t),\lambda,u_t) - L(x_0,\lambda,u_0) + o(t^2) \\ &= v(u_0) + D_u f(x_0,u_0)\left(\frac{u_t - u_0}{t}\right) \\ &\quad + \frac{1}{2}t^2 D^2 L(x_0,\lambda,u_0)\left[\left(\frac{\bar{x}(u_t)-\bar{x}(u_0)}{t}, d\right), \left(\frac{\bar{x}(u_t)-\bar{x}(u_0)}{t}, d\right)\right] \\ &\quad + o(t^2)\end{aligned}$$

Assuming $\mathcal{X}$ reflexive, $g_t = \frac{\bar{x}(u_t) - \bar{x}(u_0)}{t}$ has a weak accumulation pt $\tilde{g}$. Then

$$\liminf_{t\downarrow 0} \frac{v(u_t) - v(u_0) - t D_u f(x_0,u_0)(g_t)}{\frac{1}{2}t^2}$$

$$\geq \liminf_{t\downarrow 0} D^2 L(x_0,\lambda_0,u_0)((g_t,d),(g_t,d))$$

$$= D^2 L(x_0,\lambda_0,u_0)(\tilde{g},d)(\tilde{g},d)$$

- Since $\bar{x}(u_t), x_0 \in G^{-1}(K)$ + $T_K(G(x_0))$ weakly closed,

$$\frac{G(\bar{x}(u_t)) - G(x_0)}{t} \in T_K(G(x_0))$$

$$\Rightarrow DG(x_0)(\tilde{g}) \in T_K(G(x_0))$$

- Since $\bar{x}(u_t)$ - near optimal, x_0 optimal,

  $D_x f(x_0,u_0)(\tilde{g}) = 0$

$\forall$ paths u_t,

$$\liminf_{t\downarrow 0} \frac{v(u_t) - v(u_0) - t D_u f(x_0,u_0)(g_t)}{\frac{1}{2}t^2}$$

$$\geq \inf_{g\in C(x_0)} \sup_{\lambda\in\Lambda(x_0,u_0)} D^2 L(x_0,\lambda,u_0)((g,d),(g,d))$$

(iii) + (iv) imply uniqueness, norm-convergence of limit of $(\bar{x}(u_t) - x_0)/t$

PAOP Thm 4.133

Assume $\mathcal{X}$ finite-dim. Assume $\forall g\in\mathcal{X}$ $G^{-1}(K)$ 2nd-order regular at x_0 in direction g. Assume Lipschitz stable $o(\|u-u_0\|^2)$-optimal soln $\bar{x}(u)$. Then if

$$\min_{g\in C(x_0)} \max_{\lambda\in\Lambda(x_0,u_0)} D^2 L(x_0,\lambda,u_0) - \sigma\left(\lambda, T_K^2(G(x_0), DG(x_0)d)\right)$$

Has a unique opt. soln,

$$\nabla \bar{x}(u_0, d)$$

exists + equals that soln.