Lecture 10

General existence results and introduction to nonsmooth maps

October 8, 2007

Outline

- Motivating Example:
  - Traffic Equilibrium Problem
- Copositivity
- Implications for Affine VI's and CP's
- Nonsmooth Maps
- Generalized Gradients and Jacobians
- Generalized Directional Differentiabilty

Recap

In the past few lectures, we have

- Considered $CP(K, F)$: where
  K is a **closed convex cone** and F is a **monotone map**

- Addressed question: Solvability of $CP(K, F)$ through feasibility

- Learned (among other)
  - F is continuous and pseudo monotone, CP(K, F) strictly feasible implies nonemptiness and compactness of the solution set
  - F affine and K polyhedral (solvability $\equiv$ feasibility)
  $$F(x) = q + Mx$$
    - ⋆ F monotone is equivalent to $JF = M$ being positive semidefinite

- What can we say when M is not positive semidefinite (F not monotone)?

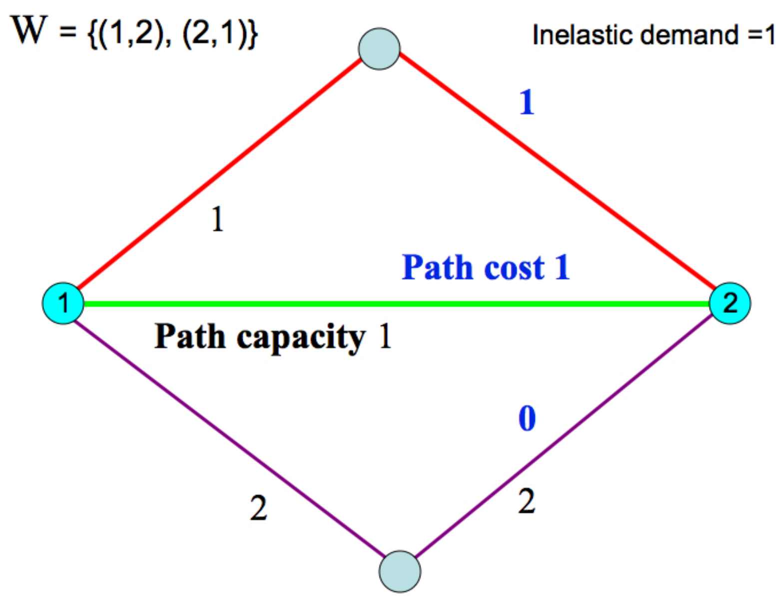
Traffic Equilibrium Problem

- Static traffic equilibrium model for congestion control in a network
- The set of origin-destination (OD) pairs of users is $\mathcal{W}$
- The users compete for the network resource (flow on links/paths)
- Each OD pair $w \in \mathcal{W}$ sends a flow x_w (a decision variable)
  - A flow x_w is a continuous variable, so multiple-paths can be used by the same OD pair w
- In the presence of congestion, network operator imposes costs on (congested) links
  - Each link $a \in \mathcal{A}$ has a flow-cost $c_a(x)$, a function of the network flow x
- The goal of the operator is to impose costs that will force users to redirect flow and avoid congestion

Congestion Control: Path Formulation

Network Model

- For each $w \in \mathcal{W}$, let $\mathcal{P}_w$ be the set of paths p connecting the OD pair w
- Let $\mathcal{P}$ be the set of all paths: $\mathcal{P} = \cup_{w \in \mathcal{W}} \mathcal{P}_w$
- Let h_p denote the flow on path p resulting from all the traffic that uses this path
- The congestion control problem can be viewed as a two-player game:
  - Network receives flow requests $h_{wp}, p \in \mathcal{P}_w$ from OD pairs $w \in \mathcal{W}$
  - Based on "potential" congestion on the paths carrying the flow, network "operator" assigns path prices $u = \{u_p \mid p \in \mathcal{P}\}$
  - Based on the path prices u, each OD pair w
    - ⋆ Determines the paths $\mathcal{P}_w$ and decides on flows h_{wp} on these paths
    - ⋆ Subject to meeting its demand $d_w(u)$



Model of users with *elastic demand*

- Each $w \in \mathcal{W}$ has a flow demand $d_w(u)$ as a function of the minimum travel cost u for all users
  - *Inelastic demand* model assumes $d_w(u)$ is constant for all w
- Each OD pair w decides on its path-flows $h_{wp}, p \in \mathcal{P}_w$ that satisfy the demand $\sum_{p\in\mathcal{P}_w} h_{wp} = d_w(u)$

Stipulated rule for user's behavior: **users are selfish (noncooperative)**

- This behavioral model is the Wardrop user equilibrium principle
- From this principle, it follows that each OD pair w sends flows only along the paths h_p with minimum cost
- Paths with cost that is not minimal will not be used i.e., no user will send flow along such paths

- Mathematically the problem can be casted as a CP:

$$0 \leq C_p(h) - u_w \perp h_p \geq 0 \qquad \text{for all } w \in \mathcal{W}, p \in \mathcal{P}$$

$$u_w = \min_{p \in \mathcal{P}_w} C_p(h)$$

subject to $u_w \geq 0$ for all w, and meeting the demand

$$\sum_{p \in P_w} h_p = d_w(u) \quad \text{for all } w$$

Formulation as NCP

- Under Wardrop principle, when the demand $d_w(u)$ and cost functions $C_p(h)$ are nonnegative, and assuming that for each w we have

$$\sum_{p\in\mathcal{P}_w} h_p C_p(h) = 0 \quad \text{and} \quad h \geq 0 \qquad \text{imply } h_p = 0 \text{ for all } p \in \mathcal{P}_w.$$

  Then the static user equilibrium problem is equivalent to the following nonlinear complementarity problem

$$0 \leq \begin{bmatrix} h \\ u \end{bmatrix} \perp \begin{bmatrix} C(h) - B^T u \\ Bh - d(u) \end{bmatrix} \geq 0$$

  where B is OD pair-path incidence matrix

$$B_{wp} = 1 \quad \text{if } p \in \mathcal{P}_w \qquad \text{and} \qquad B_{wp} = 0 \quad \text{otherwise}$$

- Existence of solutions? See Prop. 2.2.14 for a result

Traffic Equilibrium Problem as Nonlinear CP

- Suppose cost and demand are affine

$$C(h) = Ch + q_1, \qquad -d(u) = Du + q_2$$

- Then, we have following complementarity problem

$$0 \leq \begin{bmatrix} h \\ u \end{bmatrix} \perp \begin{bmatrix} C & -B^T \\ B & D \end{bmatrix} \begin{bmatrix} h \\ u \end{bmatrix} + \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \geq 0$$

- Thus, it is of the form

$$0 \leq x \perp Mx + q \geq 0$$

- What about the size of the problem?
- M need not positive semidefinite, hence $F(x) = Mx + q$ not monotone
- However, the constraint set $K = \{x \mid x \geq 0\}$ is a nice cone (closed and convex; in fact, polyhedral)
- Copositivity of M on K may help

What next

- Affine variational inequality problem: finding an $x \in K$ such that

$$(y - x)^T (Mx + q) \geq 0 \qquad \text{for all } y \in K$$

- The importance of the affine VI's stems from "linear approximation" of nonlinear F by a first-order Taylor expansion near a point of interest:

$$F(x) \approx F(x_0) + JF(x_0)(x - x_0) \qquad \text{for all } x \text{ near } x_0$$

- Thus, a given $VI(K, F)$ can be approximated locally near x_0 by

$$VI(K, q, M) \qquad \text{with } q = F(x_0) - JF(x_0)x_0, \; M = JF(x_0)$$

- This approximation idea is at the heart of algorithms for solving $VI's$
- We next study the $VI(K, q, M)$ where
  - K is a **cone** or a **polyhedral set**
  - M is **copositive**

Copositivity

Def. Let C be a cone in $\mathbb{R}^n$. A matrix $A \in \mathbb{R}^{n\times n}$ is said to be **copositive** on C when

$$x^T Ax \geq 0 \qquad \text{for all } x \in C$$

Examples:

- Let C be the nonnegative orthant in $\mathbb{R}^n$, $C = \mathbb{R}^n_+$.
  Any matrix A with nonnegative entries $a_{ij} \geq 0$ for all i, j is copositive on $\mathbb{R}^n_+$
- Let A be a symmetric matrix. Then, the matrix can be represented as

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^T$$

  where for each i, we have:
  - $\lambda_i \in \mathbb{R}$ is an eigenvalue of A and $v_i \in \mathbb{R}^n$ is a normalized eigenvector corresponding to λ_i

- Vectors v_i are mutually orthogonal and form a basis in $\mathbb{R}^n$
- Is there a cone C such that A is copositive on C?

Range, Domain, and Kernel of $VI(K, q, M)$

Let $K \subseteq \mathbb{R}^n$ and $M \in \mathbb{R}^{n\times n}$ be fixed. Consider $VI(K, q, M)$ when $q \in \mathbb{R}^n$ varies.

- The VI range of the pair (K, M) is the set given by

$$\mathcal{R}(K, M) = \{q \in \mathbb{R}^n \mid SOL(K, q, M) \neq \emptyset\}$$

  This is the set of all q for which VI(K, q, M) has a solution.

- The VI domain of the pair (K, M) is the set given by

$$\mathcal{D}(K, M) = (K_\infty)^* - MK$$

- The VI kernel of the pair (K, M) is the set given by

$$\mathcal{K}(K, M) = SOL(K_\infty, 0, M)$$

  This is the set of solutions to the homogeneous CP on the recession cone K_∞ or $K_\infty \ni v \perp Mv \in (K_\infty)^*$

In general, are these sets closed or convex or cones? What can we say about these sets when K is a cone? The domain set has no “special meaning” in general case.

When K is a cone, the domain $\mathcal{D}(K, M)$ is related to the set of q's for which the resulting $CP(K, M)$ is feasible.

Kernel Description

Proposition 1 (*Kernel for closed convex K*)

Let $K \subseteq \mathbb{R}^n$ be a closed convex set and $M \in \mathbb{R}^{n \times n}$. Then, we have

$$\cup_{q \in \mathbb{R}^n} (SOL(K, q, M))_\infty \subseteq \mathcal{K}(K, M)$$

Consequence 2 When K is closed convex and $\mathcal{K}(K, M) = \{0\}$, the solution set $SOL(K, q, M)$ is **bounded** for any $q \in \mathbb{R}^n$ (may be empty)

Proof of the Proposition Let $q \in \mathbb{R}^n$ and d be a recession direction of $SOL(K, q, M)$. Then, there exists and $x \in SOL(K, q, M)$ such that

$$x + \tau d \in SOL(K, q, M) \quad \text{for all } \tau > 0 \tag{1}$$

Since $x \in SOL(K, q, M)$ we have $x \in K$. Also, relation (1) implies

$$(y - x - \tau d)^T (q + Mx + \tau Md) \geq 0 \qquad \text{for all } y \in K, \tau > 0$$

By letting $y = x$, we have $-\tau d^T q - \tau^2 d^T Md \geq 0$ for all $\tau > 0$. Hence $d^T Md \leq 0$. By letting $y = x + 2\tau d$, we similarly obtain $d^T Md \geq 0$.

Hence

$$d^T M d = 0 \qquad \text{with } d \in K_\infty$$

To show $d \in \mathcal{K}(K, M)$, it remains to prove that $Md \in (K_\infty)^*$. Let $\tilde{d} \in K_\infty$ be arbitrary. Then $x + \tau d + \tilde{d} \in K$, and letting $y = x + \tau d + \tilde{d}$ in Eq. (1), we see that

$$\tilde{d}^T (q + \tau M d + M\tilde{d}) \geq 0 \qquad \text{for all } \tau > 0$$

Since this relation holds for any τ, it follows that

$$\tilde{d}^T M d \geq 0,$$

thus showing $Md \in (K_\infty)^*$. We have therefore shown that

$$K_\infty \ni d \perp Md \in (K_\infty)^*,$$

completing the proof.

More on Kernel Description

Proposition 3 (*Kernel for closed convex cone K*)

Let $K \subseteq \mathbb{R}^n$ be a closed convex cone and $M \in \mathbb{R}^{n\times n}$. Then, we have

$$\cup_{q\in\mathbb{R}^n} (SOL(K,q,M))_\infty = \mathcal{K}(K,M) \tag{2}$$

Furthermore, $\mathcal{K}(K,M) = \{0\}$ if and only if there exists a constant $c > 0$ such that

$$\|x\| \le c\|q\| \qquad \text{for all } x \in SOL(K,q,M) \text{ and all } q \in \mathbb{R}^n$$

Proof By definition of the kernel, we have $\mathcal{K}(K,M) = SOL(K_\infty, 0, M)$. Since K is a closed convex cone, we have $K = K_\infty$. Hence,

$$\mathcal{K}(K,M) = SOL(K_\infty, 0, M) = SOL(K, 0, M).$$

Because K is a closed cone, the solution set $SOL(K,0,M)$ is also a closed cone. Therefore $SOL(K,0,M) \subseteq SOL(K,0,M)_\infty$.

Hence, $\mathcal{K}(K,M) \subseteq SOL(K,0,M)_\infty$.

This and Proposition 1 yield relation (3).

For the "uniform boundedness" statement, assume that such a constant c does not exist. Then, for any k, we can find $q_k \neq 0$ and $x_k \neq 0$ such that

$$x_k \in SOL(K, q_k, M) \quad \text{and} \quad \|x_k\| \geq k\|q_k\|.$$

Therefore

$$\lim_{k\to\infty} \frac{\|x_k\|}{\|q_k\|} = \infty.$$

Consider vectors $z_k = \frac{x_k}{\|x_k\|}$. The sequence $\{z_k\}$ is bounded and contained in K (since K is a cone). Because K is a cone, and x_k is a solution to $CP(K, q_k, M)$, by scaling with $\|x_k\|$ we see that z_k is a solution to $CP(K, q_k/\|x_k\|, M)$. The bounded vectors z_k have a limit point $\tilde{z}$, which belongs to the cone K by closedness of K. Also, since $z_k \in SOL(K, q_k/\|x_k\|, M)$ and $\frac{\|q_k\|}{\|x_k\|} \to 0$, it follows that $\tilde{z} \perp M\tilde{z}$ with $M\tilde{z} \in K^*$. Hence,

$$\tilde{z} \in SOL(K, 0, M) = SOL(K_\infty, 0, M)$$

implying that $\tilde{z} \in \mathcal{K}(K, M)$ with $\tilde{z} \neq 0$ - a contradiction.

More on Kernel Description

Proposition 1 (*Kernel for closed convex cone K*)

Let $K \subseteq \mathbb{R}^n$ be a closed convex cone and $M \in \mathbb{R}^{n\times n}$. Then, we have

$$\cup_{q\in\mathbb{R}^n} (SOL(K,q,M))_\infty = \mathcal{K}(K,M) \qquad (3)$$

Furthermore, $\mathcal{K}(K,M) = \{0\}$ if and only if there exists a constant $c > 0$ such that

$$\|x\| \le c\|q\| \qquad \text{for all } x \in SOL(K,q,M) \text{ and all } q \in \mathbb{R}^n$$

Note: By Prop. 1, $\cup_{q\in\mathbb{R}^n} SOL(K,q,M)$ is uniformly bounded when q's are bounded

This property has a special name.

Def. We say that (K,M) is an *R_0-pair* if and only if the union $\cup_q SOL(K,q,M)$ **is uniformly bounded for all q belonging to a bounded set**.

Proposition 3 provides a sufficient condition for (K, M) to be an R_0-pair. Namely, for a closed convex cone K, if $\mathcal{K}(K, M) = \{0\}$ then (K, M) is an R_0-pair. Furthermore, if M is a symmetric copositive matrix, then (K, M) has the $\mathbf{R}_0$ property if and only if M is strictly copositive on K

Solutions: Existence and Properties

Theorem 2 Let $K \subseteq$ be a closed convex cone and $M \in \mathbb{R}^{n \times n}$. Suppose that (K, M) is an R_0-pair. Then, if M **is copositive on** K**, the solution set** $SOL(K, q, M)$ **is nonempty and bounded**.

Theorem 2 is a trim-down version of Theorem 2.5.10 in Facchinei and Pang's book, volume 1.

Theorem 3 Let $K \subseteq \mathbb{R}^n$ be a closed convex cone and $F : K \to \mathbb{R}^n$ be a continuous map. Suppose that is a copositive matrix $E \in \mathbb{R}^{n \times n}$ on K such that (K, E) is an R_0-pair and the union

$$\cup_{\tau > 0} SOL(K, F + \tau E)$$

is bounded, then $CP(K, F)$ has a solution.

- Theorem 3 provides sufficient conditions for existence of solutions without any monotonicity requirement on F

Existence Certificate

- As a direct consequence of Theorem 3, we have the following result

  Corollary 4 Let $K \subseteq \mathbb{R}^n$ be a closed convex cone and $F : K \to \mathbb{R}^n$ be a continuous map. Then, either $CP(K, F)$ has a solution or there are unbounded sequences $\{x_k\}$ and $\{\tau_k\}$ such that

$$K \ni \ x_k \perp F(x_k) + \tau_k x_k \ \in K^*, \quad \tau_k > 0 \quad \text{for all } k \geq 0$$

- The preceding material is in Chapters 2.5 and 2.6 of Facchinei and Pang's book v1.

System of Nonsmooth Equations

- We are interested in numerical methods for solving a system of nonsmooth equations

$$G(x) = 0, \qquad x \in X$$

where G is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$, and $X \subset \mathbb{R}^n$ is closed.

- A $CP(K, F)$ in $\mathbb{R}^n$ is equivalent to

$$G(x,y) = \Big(G_1(x,y),\ \ldots\ G_n(x,y),\ G_{n+1}(x,y),\ \ldots,\ G_{2n}(x,y)\Big)^T$$
$$G_i(x,y) = x_i y_i \qquad \text{for } i = 1, \ldots, n$$
$$G_{n+i}(x,y) = y_i - F_i(x) \qquad \text{for } i = 1, \ldots, n; X = K \times K^*.$$

- Why nonsmooth? Appears naturally in applications, example for predicting traffic equilibrium

$$C_p(h) = \max_{a \in p} c_a(A^T h), \text{ where } C_p(h) \text{ not differentiable in h}$$

Locally Lipschitz Functions

- Clearly, a continuous function need not be differentiable
- However, a continuous function with "local Lipschitz" properties is differentiable almost everywhere

Theorem: *(Rademacher Theorem)*
Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Let $f : \Omega \to \mathbb{R}$ be a continuous function on Ω and locally Lipschitz on Ω, i.e., for every $x \in \Omega$, there exists an open ball $B(x, r_x)$ such that

$$|f(y) - f(x)| \le L_x \|x - y\| \qquad \text{for all } \; y \in B(x, r_x)$$

Then, the function f is differentiable almost everywhere on the set Ω

Example: $f(x) = |x|$, $\Omega = \mathbb{R}$

- This result applies to maps $F = [F_1, \ldots, F_m]^T$ component-wise

Locally Lipschitz Maps

- Similarly, a continuous map with "local Lipschitz" properties is differentiable almost everywhere
- In other words, it has Jacobian well defined almost everywhere

Theorem:
Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Let $F : \Omega \to \mathbb{R}^m$ be a continuous map on Ω and locally Lipschitz on Ω, i.e., for every $x \in \Omega$, there exists an open ball $B(x, r_x)$ such that

$$\|F(y) - F(x)\| \le L_x \|x - y\| \qquad \text{for all} \quad y \in B(x, r_x)$$

Then, the map F is differentiable almost everywhere on the set Ω, i.e., **the Jacobian JF(x) exists almost everywhere**

- Continuous locally Lipschitz functions (maps) are "almost differentiable"
- The main fact: a point $\bar{x}$ where such a function is not differentiable can be approached with a sequence $\{x_k\}$ of points where f is differentiable
- This is used to define a notion of generalized gradient/Jacobian

Limiting Gradient and Jacobian

Let $\Omega \subseteq \mathbb{R}^n$ be an open set and let $\bar{x} \in \Omega$ be arbitrary. Let $f : \Omega \to \mathbb{R}$ be a continuous function locally Lipschitz on Ω.

Def. The set of **limiting gradients** of f at $\bar{x}$ is

$$Jac\, g(\bar{x}) = \left\{ g \mid g = \lim_{k\to\infty} \nabla f(x_k) \quad \text{for some } \{x_k\} \subseteq \Omega \text{ with } x_k \to \bar{x} \right\}$$

- This is a set of vectors in $\mathbb{R}^n$

Let $F : \Omega \to \mathbb{R}^m$ be a continuous map locally Lipschitz on Ω.

Def. The set of **limiting Jacobians** of map F at $\bar{x}$ is

$$Jac\, F(\bar{x}) = \left\{ g \mid g = \lim_{k\to\infty} JF(x_k) \quad \text{for some } \{x_k\} \subseteq \Omega \text{ with } x_k \to \bar{x} \right\}$$

- This is a set of matrices in $\mathbb{R}^{m\times n}$

Drawback of Limiting Gradients and Jacobians

- Difficult to compute and manipulate (chain rules, operations)
- Do not admit mean-value theorems among others
- Cannot capture optimality conditions

Example: $f(x) = |x|, \quad x \in \mathbb{R}$

Generalized Gradient and Jacobian

Let $\Omega \subseteq \mathbb{R}^n$ be an open set and let $\bar{x} \in \Omega$ be arbitrary. Let $f : \Omega \to \mathbb{R}$ be a continuous function locally Lipschitz on Ω.

Def. The **generalized differential** of f at $\bar{x}$ is

$$\partial f(\bar{x}) = \mathrm{conv}\, Jac\, g(\bar{x})$$

- This is a convex set of vectors in $\mathbb{R}^n$. A vector $g \in \partial f(\bar{x})$ is a **generalized gradient**

Let $F : \Omega \to \mathbb{R}^m$ be a continuous map locally Lipschitz on Ω.

Def. The set of **generalized Jacobians** of map F at $\bar{x}$ is

$$\partial F(\bar{x}) = \mathrm{conv}\, Jac\, F(\bar{x})$$

- This is a convex set of matrices in $\mathbb{R}^{m \times n}$.

Example: $f(x) = |x|$ for $x \in \mathbb{R}$, $\qquad \partial f(0) = [-1, 1]$

Properties of Generalized Jacobians

Proposition: Let Ω be open and $G : \Omega \to \mathbb{R}^m$ be continuous and locally Lipschitz on Ω. The following is true for any $x \in \Omega$:

(a) The generalized Jacobian $\partial G(x)$ is nonempty, convex, and compact set

(b) The multivalued map $\partial G : \mathbb{R}^n \to \mathbb{R}^{m \times n}$ is closed at x, i.e., when $x_k \to x$, $A_k \in \partial G(x_k)$, and $A_k \to A$, then $A \in \partial G(x)$.

Def. Let $f : \Omega \to \mathbb{R}$ be continuous and locally Lipschitz on open set $\Omega \subseteq \mathbb{R}^n$. **Generalized directional derivative** of f at $\bar{x}$ in direction d is denoted by $f^\circ(\bar{x}; d)$, and it is defined as

$$f^\circ(\bar{x}; d) = \limsup_{\substack{y \to \bar{x} \\ t \downarrow 0}} \frac{f(y + td) - f(y)}{t}$$

- When f is locally Lipschitz at $\bar{x}$, the generalized directional derivative is finite for all directions $d \in \mathbb{R}^n$

Class of C-regular Functions

We have the following **important property** of generalized derivative: for all $(x, d) \in \Omega \times \mathbb{R}^n$,

$$f^\circ(x; d) = \max_{s \in \partial f(x)} s^T d$$

Recall, the classic definition of directional derivative $f'(\bar{x} : d)$ of f at $\bar{x}$ in direction d:

$$f'(\bar{x}; d) = \lim_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

When $f'(\bar{x}; d)$ exists, we have

$$f'(\bar{x}; d) \le f^\circ(\bar{x}; d)$$

Recall, f is said to be directionally differentiable at $\bar{x}$ when $f'(\bar{x}; d)$ exists for every direction $d \in \mathbb{R}^n$

Def. (*C-regularity*) Let $f : \Omega \to \mathbb{R}$ be continuous and locally Lipschitz on Ω. We say that **f is C-regular** at $\bar{x} \in \Omega$ when f is directionally differentiable at $\bar{x}$ and

$$f^\circ(\bar{x}; d) = f'(\bar{x}; d) \qquad \text{for all } d \in \mathbb{R}^n$$

The class of C-regular functions includes

- Convex functions
- Continuously differentiable functions

Example: $f(x) = -|x|$ and $f(x_1, x_2) = \min\{x_1, x_2\}$ are not C-regular functions

- A C-regular function at a point x admits accurate first order approximation in any given direction, i.e.,

$$f(\bar{x} + d) = f(\bar{x}) + f'(\bar{x}; d) + o(\|d\|) \qquad \text{with} \quad \lim_{\|d\| \to 0} \frac{o(\|d\|}{\|d\|} = 0$$

- This property is crucial for the family of Newton methods for solving equations $G(x) = 0$ with $x \in X$
- The preceding material is from Chapter 7.1 of Facchinei and Pang's book v2.