MATH 450/558
FALL 2015

LECTURE 15

FUNCTIONS DIFFERENTIABLE ALMOST EVERYWHERE

Let $f: \mathbb{R}^n \to \mathbb{R}$ be differentiable almost everywhere (in the sense of Lebesgue measure).

Ω - the set of measure zero where f is not differentiable.

<u>THM</u>

$$\partial f(x) = \underbrace{\boxed{Co\left\{ \lim_{j\to\infty} \nabla f(x^{(j)}) \mid \{x^{(j)}\} \text{ is such that } \lim_{j\to\infty} x^{(j)} = x,\ x^{(j)} \notin \Omega\ \forall j \text{ and } \{\nabla f(x^{(j)})\} \text{ is convergent} \right\}}}_{C}$$

<u>Proof of $C \subseteq \partial f(x)$</u>

Let $\{x^{(j)}\}$ be any sequence such that $\lim_{j\to\infty} x^{(j)} = x$, $x^{(j)} \notin \Omega$ $\forall j$ and $\lim_{j\to\infty} \nabla f(x^{(j)}) = \nabla f_*$ exists.

Since f is differentiable at $x^{(j)}$, we have $\nabla f(x^{(j)}) \in \partial f(x^{(j)})$ (see Thm on page 19, Lecture 11 notes). Now by thm on page 18 Lecture 11 notes

$$\lim_{j\to\infty} \nabla f(x^{(j)}) \in \partial f\left(\lim_{j\to\infty} x^{(j)}\right)$$

$$\Longrightarrow \nabla f_* \in \partial f(x)$$

Since $\partial f(x)$ is convex, convex hull of such ∇f_* is contained in $\partial f(x)$. $\square$

<u>Remark</u>

Since the generalized gradient $\partial f(x^{(j)})$ is bounded, even if the sequence $\{\nabla f(x^{(j)})\}$ itself is not convergent in the proof of the thm above, it must have convergent subsequences with limit in $\partial f(x)$ by the Bolzano-Weierstrass thm.

Thus C is not empty.

<u>Ex</u>

① $f(x) = \max\{f_1(x), \ldots, f_n(x)\}$

where $f_j : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable everywhere $j = 1, \ldots, n$.

$A(x) := \{ j \in \{1, \ldots, n\} \mid f_j(x) = f(x) \}$

Assume $|A(\bar{x})| = 1$ for all $\bar{x} \in B(x, \delta) \setminus \{x\}$ for some $\delta > 0$.

By thm above

$$\partial f(x) = Co \{ \nabla f_j(x) \mid j \in A(x) \}$$

e.g.

$$f(x) = \max \{ e^x, 1/(1+x^2) \}$$

$$\partial f(0) = Co \left\{ \left.\frac{de^x}{dx}\right|_{x=0}, \left.\frac{d\, 1/(1+x^2)}{dx}\right|_{x=0} \right\}$$

$$= Co \{ 1, 0 \}$$

$$= [0, 1].$$

(2) $f(x) = \| \underset{\text{in } \mathbb{R}^m}{b} - \underset{m \times n}{A} x \|_1$ $\quad A = \begin{bmatrix} \bar{a}_1 \\ \vdots \\ \bar{a}_m \end{bmatrix}$ rows of A

$$f(x) = |b_1 - \bar{a}_1 x| + |b_2 - \bar{a}_2 x| + \cdots + |b_m - \bar{a}_m x|$$

$$A(x) = \{ j \mid b_j - \bar{a}_j x = 0 \}$$

$$s_j(x) = \begin{cases} 1 & -b_j + \bar{a}_j x > 0 \\ -1 & -b_j + \bar{a}_j x < 0 \end{cases} \quad \text{for each } j \in \{1, \ldots, m\} \setminus A(x)$$

By thm above

$$\partial f(x) = Co\left\{ \sum_{j \notin A(x)} s_j(x)\, \bar{a}_j^T + \sum_{j \in A(x)} b_j\, \bar{a}_j^T \;\middle|\; b_j \in \{-1, 1\} \text{ for each } j \in A(x) \right\}$$

Largest eigenvalue function

$f_t : \mathbb{R}^n \to \mathbb{R}$ a continuously differentiable function depending on a parameter $t \in T$ given set

$$f(x) = \sup\{ f_t(x) \mid t \in T \}$$

THM

$$\partial f(x) = Co\left\{ \lim_{k\to\infty} \nabla f_{t^{(k)}}(x^{(k)}) \;\middle|\; \{t^{(k)}\}, \{x^{(k)}\} \text{ are convergent sequences in } T, \mathbb{R}^n \text{ s.t. } \lim_{k\to\infty} x^{(k)} = x \text{ and } \lim_{k\to\infty} f_{t^{(k)}}(x) = f(x) \right\}$$

$$\lambda_{max}(x) := \lambda_{max}(A(x))$$

$A(x) : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ symmetric PD continuously differentiable

Characterization

$$\lambda_{max}(x) = \max\{ \underbrace{w^T A(x) v}_{f_{w,v}(x)} \mid \underbrace{\|w\|_2, \|v\|_2 \leq 1}_{T := \{(w,v) \in \mathbb{R}^n \times \mathbb{R}^n \mid \|w\|_2, \|v\|_2 \leq 1\}} \}$$

Let

$$M(x) := \{ (w,v) \in T \mid w^T A(x) v = \lambda_{max}(x) \}$$

also r denote the algebraic multiplicity of $\lambda_{max}(x)$.

By thm above

$$\partial \lambda_{max}(x) = Co \left\{ \left[\begin{matrix} \frac{\partial (w^T A(\bar{x}) v)}{\partial \bar{x}_1} \\ \vdots \\ \frac{\partial (w^T A(\bar{x}) v)}{\partial \bar{x}_n} \end{matrix} \right] \Bigg|_{\bar{x}=x} \ \Bigg| \ (w,v) \in M(x) \right\}$$

$$= Co \left\{ \left[\begin{matrix} w^T \frac{\partial A(x)}{\partial \bar{x}_1} v \\ \vdots \\ w^T \frac{\partial A(x)}{\partial \bar{x}_n} v \end{matrix} \right] \ \Bigg| \ (w,v) \in M(x) \right\}$$

Furthermore, letting $A(x) = Q \cdot \Lambda \cdot Q^T$ be the orthogonal eigenvalue decomposition of $A(x)$ (where eigenvalues of $A(x)$ appear on the diagonal of Λ in non-increasing order),

$$M(x) = \{ (w,v) \in T \mid w^T Q \Lambda Q^T v = \lambda_{max}(x) \}$$

$$= \{ (\overset{Q}{\uparrow}v, \overset{Q}{\uparrow}v) \mid v \in S_r \}$$

where $S_r = \{ \alpha_1 e_1 + \dots + \alpha_r e_r \mid \alpha_1, \dots, \alpha_r \in \mathbb{R} \text{ s.t. } \alpha_1^2 + \dots + \alpha_r^2 = 1 \}$

Thus

$$\partial \lambda_{max}(x) = Co \left\{ \left[\begin{matrix} \overset{Q^T}{\uparrow} v^T \frac{\partial A(x)}{\partial \bar{x}_1} \overset{Q}{\uparrow} v \\ \vdots \\ \underset{Q^T}{\downarrow} v^T \frac{\partial A(x)}{\partial \bar{x}_n} \underset{Q}{\downarrow} v \end{matrix} \right] \ \Bigg| \ v \in S_r \right\}.$$

When $\lambda_{max}(x)$ has algebraic multiplicity one,

$$S_r = \{e_1, -e_1\}$$

thus

$$\partial \lambda_{max}(x) = \left\{ \begin{bmatrix} e_1^T Q^T \frac{\partial A(x)}{\partial \bar{x}_1} Q e_1 \\ \vdots \\ e_1^T Q^T \frac{\partial A(x)}{\partial \bar{x}_n} Q e_1 \end{bmatrix} \right\}$$

$$= \left\{ \begin{bmatrix} q_1^T \frac{\partial A(x)}{\partial \bar{x}_1} q_1 \\ \vdots \\ q_1^T \frac{\partial A(x)}{\partial \bar{x}_n} q_1 \end{bmatrix} \right\}$$

where q_1 is a unit eigenvector corresponding to $\lambda_{max}(x)$. This implies $\lambda_{max}(\bar{x})$ is continuously differentiable at $\bar{x} = x$.