Nonlinear Optimization Lecture 7

Garrick Aden-Buie

Tuesday, February 2, 2016

Theorem. For the following optimization problem, where $f \in C^1$, f convex and X convex

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t} \quad & x \in X \end{aligned}$$

$\bar{x} \in S$ is a global minimum $\Leftrightarrow [\nabla f(\bar{x})]^T(x - \bar{x}) \geq 0 \; \forall x \in S$.

The necessary and sufficient condition is called the *variational inequality problem.*

More generally, if f is not differentiable:

$\bar{x} \in S$ is a global min $\Leftrightarrow \exists$ a subgradient ξ at $\bar{x}$ such that $\xi^T(x - \bar{x}) \geq 0 \; \forall x \in S$.

Corollary (1).

(1) $\nabla f(x) = 0 \Rightarrow \bar{x}$ is a global min.

Proof: $\nabla f(\bar{x}) = 0 \Rightarrow \nabla f(x)^T(x - \bar{x}) \geq 0 \; \forall x \in S$.

(2) $\exists$ subgradient $\xi = 0$ at $\bar{x} \in S \Rightarrow \bar{x} \in S$ is a global min.

Corollary (2).

Let $\bar{x} \in \text{Int}S$.

(1) $\bar{x}$ global min $\Rightarrow \nabla f(\bar{x}) = 0$

(2) $\bar{x}$ global min $\Rightarrow \exists$ subgradient $\xi = 0$ at $\bar{x}$.

Example: consider an LP:

$$\begin{aligned} \min \quad & c^T x = f(x) \\ \text{s.t} \quad & Ax = 6 \\ & x \geq 0 \end{aligned}$$

Definition. Let S be a convex set. An extreme point of S is a point that cannot be represented as a strict convex combination of two distinct points in S.

Representation Theorem

Suppose S is closed, convex and bounded. Then any point in S can be represented as a convex combination of extreme points of S.

Theorem. Let $S \subset \mathbb{R}^n$ be compact (closed and bounded) and convex, and f be convex.

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t} \quad & x \in X \end{aligned}$$

Then there exists a global optimal solution at an extreme point of S.

Proof. $\bar{x}$ global maximum $\Leftrightarrow f(\bar{x}) \geq f(x)\ \forall x \in S$.

Let the extreme point of S be $x^1, x^2, \ldots$.

$$\bar{x} \sum_j \lambda_j x^j,\ \sum_j \lambda_j = 1,\ \lambda_j \geq 0$$

Let $x^k \colon f(x^k) = \max_j f(x^j)$.

$$\begin{aligned} f(\bar{x}) = f(\textstyle\sum \lambda_j x_j) &\leq \lambda_j f(x_j) && \text{(by convexity of } f) \\ &\leq f(x^k) \textstyle\sum \lambda_j = f(x^k) \\ \Rightarrow f(\bar{x}) &\leq f(x^k) \end{aligned}$$

if $\bar{x}$ is a global maximum, then x^k is a global maximum.

If f is *convex*:

$$\begin{gathered} \bar{x} \text{ local min} \Rightarrow \bar{x} \text{ global min} \\ \nabla f(\bar{x}) = 0 \Rightarrow \bar{x} \text{ global min} \end{gathered}$$

If f is *strictly convex*:

$$\begin{gathered} \bar{x} \text{ local min} \Rightarrow \bar{x} \text{ unique global min} \\ \nabla f(\bar{x}) - 0 \Rightarrow \bar{x} \text{ unique global min} \end{gathered}$$

Definition. *Quasiconvex.* Let $f \colon S \to \mathbb{R}, S$ convex, $S \subset \mathbb{R}^n, S \neq \emptyset$.

f is *quasiconvex* on S if

$$f(\lambda \hat{x} + (1-\lambda)\bar{x}) \leq \max\{f(\hat{x}), f(\bar{x})\},\ \forall \hat{x}, \bar{x} \in S,\ \lambda \in [0,1]$$

Theorem. f is quasiconvex on $S \Leftrightarrow S_\alpha = \{x \in S \colon f(x) \leq \alpha\}$ (level set).

Proof. $(\Rightarrow)$ $\bar{x}, \hat{x} \in S_\alpha,\ \lambda \in [0,1] \Rightarrow f(\hat{x}) \leq \alpha$ and $f(\bar{x}) \leq \alpha$.

We know that $f(\lambda\hat{x} + (1-\lambda)\bar{x}) \leq \max\{f(\hat{x}), f(\bar{x})\} \Rightarrow \lambda\hat{x} + (1-\lambda)\bar{x} \in S_\alpha \Rightarrow S_\alpha$ is convex.

Proof. $(\Leftarrow)$ $\bar{x}, \hat{x} \in S_\alpha,\ \lambda \in [0,1]$. Let $\alpha = \max\{f(\hat{x}), f(\bar{x})\}$, then $\hat{x}, \bar{x} \in S_\alpha$. Since S_α is convex, $f(\lambda\hat{x} + (1-\lambda)\bar{x}) \leq \alpha \Rightarrow f$ is quasiconvex.

Definition. Given $f \colon S \to \mathbb{R}, S$ convex, $S \subset \mathbb{R}^n, S \neq \emptyset$. f is *strictly quasiconvex* on S if

$$f(\lambda\hat{x} + (1-\lambda)\bar{x}) < \max\{f(\hat{x}), f(\bar{x})\}$$

$\forall \hat{x}, \bar{x} \in S$, $\lambda \in (0,1)$ and $f(\hat{x}) \neq f(\bar{x})$. (*Eliminates all flat spots except at the bottom*).

Note.

- A strictly convex function is a convex function.
- A strictly quasiconvex function may not be a quasiconvex function.

Example: $f \colon \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases}$$

This is a strictly quasiconvex function, but not a convex function.

Definition. f *strongly quasiconvex* on S if

$$f(\lambda\hat{x} + (1-\lambda)\bar{x}) < \max\{f(\hat{x}), f(\bar{x})\},\ \forall \hat{x}, \bar{x} \in S,\ \lambda \in (0,1), \hat{x} \neq \bar{x}$$

- Strictly convex $\Rightarrow$ strongly quasiconvex.
- Strongly quasiconvex $\Rightarrow$ strictly quasiconvex.
- Strongly quasiconvex $\Rightarrow$ quasiconvex.

Theorem. $S \subset \mathbb{R}^n$ non-empty, open convex. $f \colon S \to \mathbb{R},\ f \in C^1(S)$. f is *quasiconvex* if and only if

$$f(\hat{x}) \leq f(\bar{x}) \Rightarrow [\nabla f(\bar{x})]^T (\hat{x} - \bar{x}) \leq 0,\ \forall \bar{x}, \hat{x} \in S$$

Theorem. $f \colon \mathbb{R}^n \to \mathbb{R}$ strongly quasiconvex. $S \subset \mathbb{R}^n$ non-empty convex.

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t} \quad & x \in S \end{aligned}$$

If $\bar{x}$ is a local min, then $\bar{x}$ is the unique global min.