Optimization Theory and Applications

Kun Zhu (zhukun@nuaa.edu.cn)

January 7, 2019

Introduction

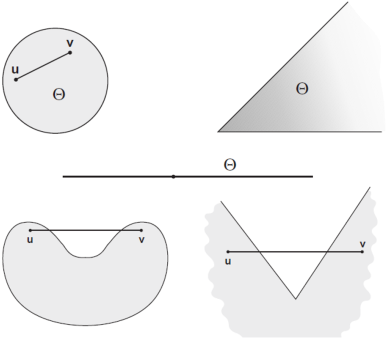
- Consider the general problem

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in \Omega. \end{aligned}$$

- We have seen several types of FONC
- When is a FONC sufficient for global optimality?
- Answer: In a convex optimization problem

Set Convexity

- Review of convex set: Ω is a convex set if, for any distinct $\mathbf{y}, \mathbf{z} \in \Omega$ and $\alpha \in (0, 1)$, we have $\alpha\mathbf{y} + (1-\alpha)\mathbf{z} \in \Omega$
- For a convex set: the line segment joining any two points in the set lies completely inside the set



Set Convexity

- Example:
  - The empty set
  - A set consisting of a single point
  - A line or a line segment
  - A subspace
  - A hyperplane
  - A half-space
  - $\mathbb{R}^n$

Set Convexity

- Example: Prove that $\Omega = \{\mathbf{x} : \mathbf{x} \geq 0\}$ is convex
  - Let $\mathbf{y}, \mathbf{z} \in \Omega$, and $\alpha \in (0, 1)$
  - Want to show that $\mathbf{x} = \alpha\mathbf{y} + (1 - \alpha)\mathbf{z} \in \Omega$
  - What does $\mathbf{x} \in \Omega$ mean?
  - To qualify as a member of Ω, each of its component must be ≥ 0
  - Hence, we must show that each component of $\mathbf{x}$ is ≥ 0
  - Each component of $\mathbf{x} = [x_1, \ldots, x_n]^T$ satisfies $x_i = \alpha y_i + (1 - \alpha) z_i$
  - Note that we have $y_i, z_i, \alpha, 1 - \alpha \geq 0$
  - Hence, $x_i \geq 0$; i.e., $\mathbf{x} \geq 0$, which means $\mathbf{x} \in \Omega$, and therefore, Ω is convex

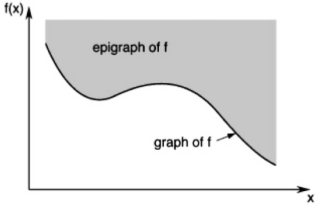
Convex Functions

- ***Definition***: The graph of $f : \Omega \to \mathbb{R}, \Omega \subset \mathbb{R}^n$, is given by

$$\left\{ \begin{bmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{bmatrix} : \boldsymbol{x} \in \Omega \right\}$$

- ***Definition***: The epigraph of a function $f : \Omega \to \mathbb{R}, \Omega \subset \mathbb{R}^n$, denoted epi$(f)$, is the set of points in $\Omega \times \mathbb{R}$ given by

$$\text{epi}(f) = \left\{ \begin{bmatrix} \boldsymbol{x} \\ \beta \end{bmatrix} : \ \boldsymbol{x} \in \Omega, \ \beta \in \mathbb{R}, \ \beta \geq f(\boldsymbol{x}) \right\}$$



Convex Functions

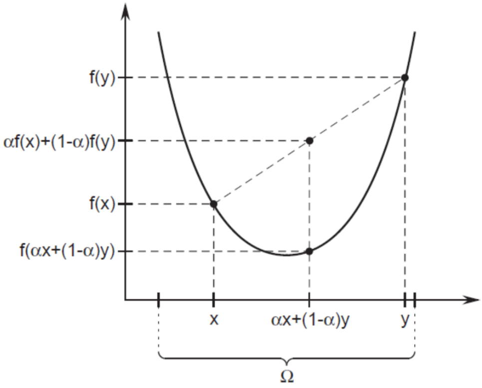
- ***Definition***: A function $f : \Omega \to \mathbb{R}, \Omega \subset \mathbb{R}^n$ is convex on Ω if its epigraph is a convex set
- ***Theorem***: A function f is a convex function on Ω if, for any distinct $\mathbf{x}, \mathbf{y} \in \Omega$ and $\alpha \in (0, 1)$

$$f(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y})$$

- f is strictly convex if $\leq$ is replaced by $<$
- f is said to be (strictly) concave if $-f$ is (strictly) convex

Convex Functions

- A geometric interpretation of convex function: line segment joining two points on the graph lies above the graph



Convex Functions

- Consider the function $f(\mathbf{x}) = x_1 x_2$. Is f convex over $\Omega = \{\mathbf{x} : x_1 \geq 0, x_2 \geq 0\}$?
  - Answer: No
  - Consider $\mathbf{x} = [1, 2]^T, y = [2, 1]^T$, then
  $$\alpha\mathbf{x} + (1-\alpha)\mathbf{y} = \begin{bmatrix} 2-\alpha \\ 1+\alpha \end{bmatrix}$$
  - Hence
  $$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) = (2-\alpha)(1+\alpha) = 2 + \alpha - \alpha^2$$
  and
  $$\alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}) = 2$$
  - If $\alpha = 1/2 \in (0, 1)$, then
  $$f(\tfrac{1}{2}\mathbf{x} + \tfrac{1}{2}\mathbf{y}) = \tfrac{9}{4} > \tfrac{1}{2}f(\mathbf{x}) + \tfrac{1}{2}f(\mathbf{y})$$
  which shows that f is not convex over Ω

Checking Convexity for Quadratics

- ***Proposition***: consider the quadratic function $f(\mathbf{x}) = \mathbf{x}^T\mathbf{Q}\mathbf{x}$, where $\mathbf{Q} = \mathbf{Q}^T$. Suppose Ω is a convex set. Then, f is convex on Ω iff

$$(\mathbf{x} - \mathbf{y})^T\mathbf{Q}(\mathbf{x} - \mathbf{y}) \geq 0$$

for all $\mathbf{x}, \mathbf{y} \in \Omega$

Checking Convexity for Quadratics

- Example: $f(\mathbf{x}) = x_1 x_2$, and can be written as $f(\mathbf{x}) = \mathbf{x}^T \mathbf{Q} \mathbf{x}$, where

  $\mathbf{Q} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

- Let $\Omega = \{\mathbf{x} : \mathbf{x} \geq 0\}$, and $\mathbf{x} = [2, 2]^T \in \Omega, \mathbf{y} = [1, 3]^T \in \Omega$, we have

  $$\mathbf{y} - \mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

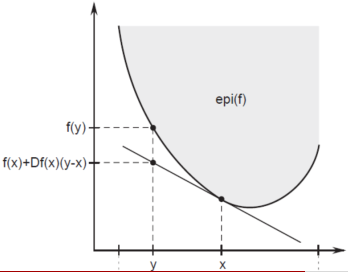
  and

  $$(\boldsymbol{y} - \boldsymbol{x})^\top \boldsymbol{Q} (\boldsymbol{y} - \boldsymbol{x}) = \frac{1}{2}[-1, 1] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -1 < 0.$$

- Hence, f is not convex on Ω

Alternative Way of Interpreting Function Convexity

- Suppose $f : \Omega \to \mathbb{R}$, Ω is convex and open, and $f \in \mathcal{C}^1$
- ***Theorem***: f is convex iff for all distinct $\mathbf{x}, \mathbf{y} \in \Omega$
$$f(\mathbf{y}) \geq f(\mathbf{x}) + Df(\mathbf{x})(\mathbf{y} - \mathbf{x})$$
- Interpretation: f convex means that it lies above any linear approximation of it
- For strict convexity, replace $\geq$ by $>$



Alternative Way of Interpreting Function Convexity

- Suppose $f : \Omega \to \mathbb{R}$, Ω convex and open, and $f \in \mathcal{C}^2$. Let $\mathbf{F}(\mathbf{x})$ be the Hessian of f at $\mathbf{x}$
- The following theorem gives another characterization of convexity
- ***Theorem***: f is convex if and only if $\mathbf{F}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \Omega$
- For strict convexity, $\mathbf{F}(\mathbf{x}) > 0$ is sufficient, but not necessary (e.g., $f(\mathbf{x}) = x^4$ is strictly convex but $f''(0) = 0$)

Alternative Way of Interpreting Function Convexity

- Examples:
  - $f(x) = x^3, \Omega = (0, 1)$. We have $f''(x) = 6x > 0$ on Ω. Hence, f is convex on Ω
  - $f(x) = -x^2, \Omega = \mathbb{R}$. We have $f''(x) = -2 < 0$. Hence, f is concave on Ω
  - $f(\mathbf{x}) = 2x_1x_2 - x_1^2 - x_2^2$. The Hessian of f is
  $$F(\mathbf{x}) = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$$
  which is negative semidefinite for all $\mathbf{x} \in \mathbb{R}^2$. Hence, f is concave on $\mathbb{R}^2$

Operations that Preserves Convexity

- Some operations preserve convexity or allow us to construct new convex functions
- ***Scale a convex function***: f is convex, then for any $\alpha \geq 0$,

  $$\bar{f} = \alpha f$$

  is convex
- ***Nonnegative weighted sums***: $f_1, \ldots, f_m$ are convex, then for nonnegative numbers $c_1, \ldots, c_m$, the function

  $$f = c_1 f_1 + \cdots + c_m f_m$$

  is convex
- ***Pointwise maximum***: f_1 and f_2 are convex, then their pointwise maximum f, defined by

  $$f(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}$$

  is convex

Convex Optimization Problems

- Consider

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega, \end{array}$$

  where Ω is a convex set, and f is a convex function on Ω

- Name: Convex optimization problem, or convex programming problem

- Examples: LP, QP, SDP

Convex Optimization Problems

- ***Theorem***: In a convex optimization problem, a point is a global minimizer if and only if it is a local minimizer
- ***Lemma***: Let $g : \Omega \to \mathbb{R}$ be a convex function defined on a convex set $\Omega \subset \mathbb{R}^n$. Then, for each $c \in \mathbb{R}$, the set

  $$\Gamma_c = \{\mathbf{x} \in \Omega : g(\mathbf{x}) \leq c\}$$

  is a convex set
- ***Corollary***: In a convex optimization problem, the set of all global minimizers is convex (simply by setting $c = \min_{\mathbf{x}\in\Omega} f(\mathbf{x})$)

Convex Optimization Problems

- Summary of FONCs
  - Set constraint: $\mathbf{d}^T \nabla f(\mathbf{x}^*) \geq 0$ for all feasible directions $\mathbf{d}$
  - Interior: $\nabla f(\mathbf{x}^*) = 0$
  - $\Omega = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = 0\}$: Lagrange conditions
  - $\Omega = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = 0, \mathbf{g}(\mathbf{x}) \leq 0\}$: KKT conditions

Convex Optimization Problems with Set Constraints

- ***Theorem***: consider the convex optimization problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega, \end{array}$$

  where $f \in \mathcal{C}^1$ on a convex set that contains Ω. Suppose the point $\mathbf{x}^* \in \Omega$ satisfies

$$\mathbf{d}^T \nabla f(\mathbf{x}^*) \geq 0$$

  for any feasible direction $\mathbf{d}$ at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a global minimizer

- ***Corollary***: If the point $\mathbf{x}^*$ above satisfies $\nabla f(\mathbf{x}^*) = 0$, then $\mathbf{x}^*$ is a global minimizer

Convex Optimization Problems with Equality Constraints

- Let us now consider problems with equality constraints $\mathbf{h}(\mathbf{x}) = 0$
- Assume that the constraint set $\Omega = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = 0\}$ is convex, and f is convex
- ***Theorem***: consider the convex optimization problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}. \end{array}$$

- Suppose there exists a feasible point $\mathbf{x}^*$ and a vector λ^* such that

$$Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}^T.$$

Then, $\mathbf{x}^*$ is a global minimizer

Convex Optimization Problems with Equality and Inequality Constraints

- Now consider problems with both equality and inequality constraints:

$$\mathbf{h}(\mathbf{x}) = 0, \quad \mathbf{g}(\mathbf{x}) \leq 0$$

- The constraint set is

$$\begin{aligned} \Omega &= \{x : \boldsymbol{h}(x) = 0, \boldsymbol{g}(x) \leq 0\} \\ &= \{x : \boldsymbol{h}(x) = 0\} \bigcap \{x : \boldsymbol{g}(x) \leq 0\}. \end{aligned}$$

- Note that the intersection of convex sets is convex
- Hence Ω is convex if both the above sets are convex

Convex Optimization Problems with Equality and Inequality Constraints

- We have already seen an example where the set $\{\mathbf{x} : \mathbf{h}(\mathbf{x})\} = 0$ is convex
- When is $\{\mathbf{x} : \mathbf{g}(\mathbf{x})\} \leq 0$ convex?
- Note that

$$\{\boldsymbol{x} : \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\} = \bigcap_{i=1}^{p} \{\boldsymbol{x} : g_i(\boldsymbol{x}) \leq 0\}.$$

- Therefore, if each g_i is convex, then we conclude that each $\mathbf{x} : g_i(\mathbf{x}) \leq 0$ is convex, and hence $\mathbf{x} : \mathbf{g}(\mathbf{x} \leq 0)$ is convex

Convex Optimization Problems with Equality and Inequality Constraints

- ***Theorem***: consider the convex optimization problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}. \end{array}$$

- Suppose there exists a feasible point $\mathbf{x}^*$ and vectors λ^* and μ^* such that

1. $\boldsymbol{\mu}^* \geq 0$;
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = 0^T$; and
3. $\boldsymbol{\mu}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) = 0$.

Then, $\mathbf{x}^*$ is a global minimizer

Further Reading

- “Convex optimization” by Stephen Byod.