香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

Introduction to Data Science

Lecture 18 Optimization (Cont.)

Zicheng Wang

Review: Optimality Condition

Theorem 1. *Consider an optimization problem*

$$\begin{aligned} &min.\ f(x) \\ &s.t.\ \ x \in \Omega, \end{aligned}$$

where f is a convex function and Ω is a convex set. Then, any local minimum is also a global minimum.

Optimality Condition

Theorem: If f is twice differentiable and convex, given $\boldsymbol{x}^*$ as an interior global minimizer, then

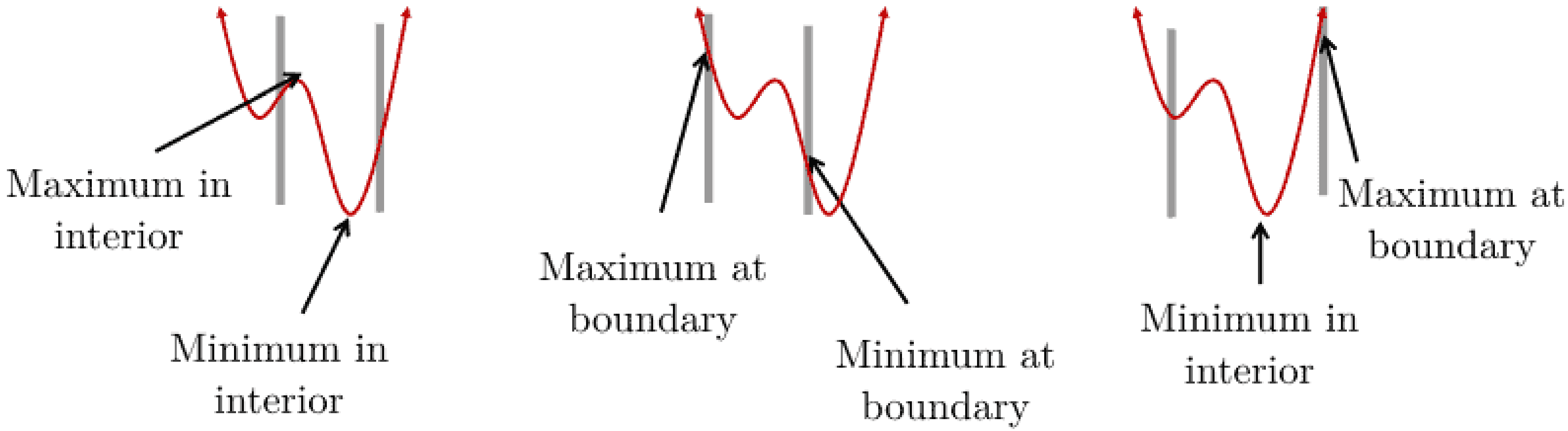
$$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$$

Theorem: If f is twice differentiable and concave, given $\boldsymbol{x}^*$ as an interior global maximizer, then

$$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$$

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{bmatrix}$$

Proof not required.

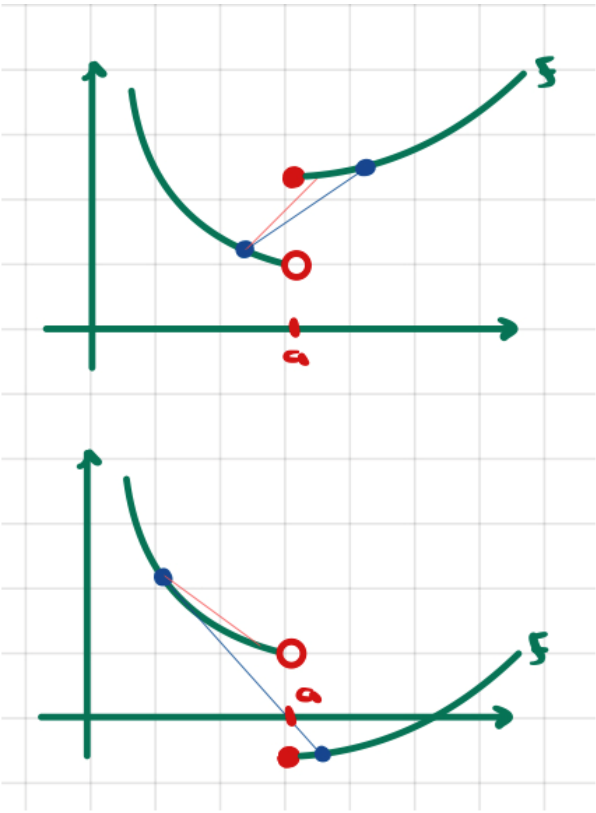

Maximum in
interior
Minimum in
interior
Maximum at
boundary
Minimum at
boundary
Minimum in
interior
Maximum at
boundary

Review: Non-Differentiable Convex Functions

1. Discontinuity

- Consider a one-dimension convex function f defined on [a,b].

- Can f be discontinuous?

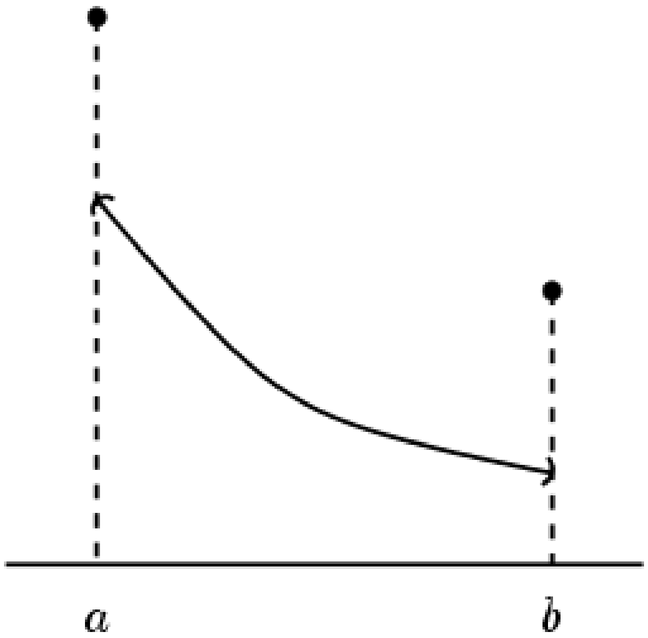


1. Discontinuity

- Consider a one-dimension convex function f defined on [a,b].

- It can be discontinuous at the boundary with

$f(a) > f_{x\downarrow a}(x)$ and $f(b) > f_{x\uparrow b}(x)$



2. Maximum of a set of Convex Functions

If $f_1, \ldots, f_m$ are convex functions, then $f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})\}$ is also convex.

<u>Proof:</u> Pick any $x, y \in dom(f), \lambda \in [0, 1]$. Then,

$$\begin{aligned}
f(\lambda x + (1-\lambda)y) &= f_j(\lambda x + (1-\lambda)y) \text{ (for some } j \in \{1, \ldots, m\}) \\
&\leq \lambda f_j(x) + (1-\lambda) f_j(y) \\
&\leq \lambda \max\{f_1(x), \ldots, f_m(x)\} + (1-\lambda) \max\{f_1(y), \ldots, f_m(y)\} \\
&= \lambda f(x) + (1-\lambda) f(y). \ \square
\end{aligned}$$

- Recall that $C = \{(\boldsymbol{x}, y) : y \geq f(\boldsymbol{x})\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function.

- If $C = \{(\boldsymbol{x}, y) : y \geq f(\boldsymbol{x})\}$ is a convex set, is $f(\boldsymbol{x})$ necessarily a convex function?

- Recall that $C = \{(\boldsymbol{x}, y): y \geq f(\boldsymbol{x})\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function.

- If $C = \{(\boldsymbol{x}, y): y \geq f(\boldsymbol{x})\}$ is a convex set, is $f(\boldsymbol{x})$ necessarily a convex function? This statement is TRUE

1. Choose any two points $(\boldsymbol{x_1}, f(\boldsymbol{x_1}))$ and $(\boldsymbol{x_2}, f(\boldsymbol{x_2}))$
2. These two points are in C
3. As C is a convex set, then $\lambda(\boldsymbol{x_1}, f(\boldsymbol{x_1}))$+ $(1-\lambda)(\boldsymbol{x_2}, f(\boldsymbol{x_2}))$ is in C. That is,

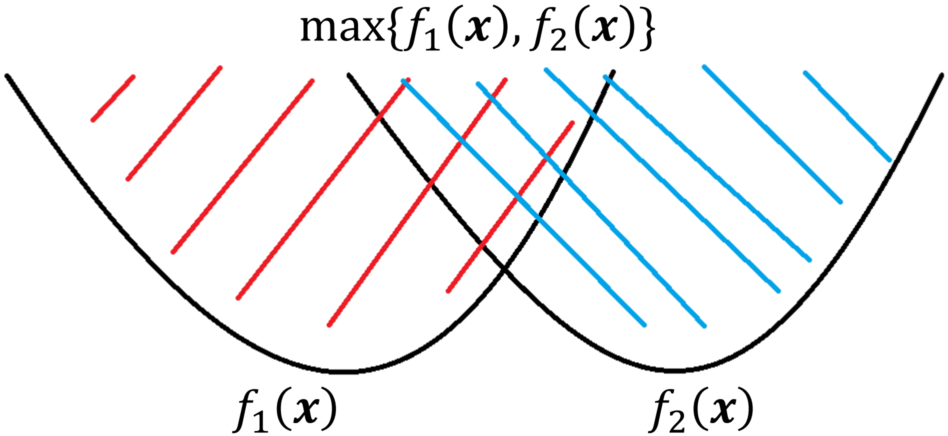
$$\lambda f(\boldsymbol{x_1})+(1-\lambda)f(\boldsymbol{x_2}) \geq \lambda \boldsymbol{x_1}+ (1-\lambda)\ \boldsymbol{x_2}$$

2. Maximum of a set of Convex Functions

If $f_1, \ldots, f_m$ are convex functions, then $f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})\}$ is also convex.

$C = \{(\boldsymbol{x}, y): y \geq f(\boldsymbol{x})\}$ is called the epigraph of function f, often denoted as **epi** f.

An alternative proof : **epi** f = $\cap_{i=1}^{m}$ $(\mathbf{epi}\ f_i)$. As $\mathbf{epi}\ f_i$ is a convex set, so is $\mathbf{epi}\ f$. Therefore, f is a convex function.



3. Nonnegative Weighted Sums

If $f_1, f_2, \ldots, f_n$ are convex functions, and $w_1, w_2, \ldots, w_n \geq 0$, then $f = w_1 f_1 + w_2 f_2 + \cdots + w_n f_n$ is also a convex function.

Proof: (also for high dimension)

- For any x and y, $0 \leq \lambda \leq 1$ and m

$$f_m(\lambda x + (1-\lambda)y) \leq \lambda fm(x) + (1-\lambda) f_m(y)$$

- Multiplying each side of the above inequality by w_m and summing over m, we have

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

Examples: Regression Analysis

- Objective: we need to minimize

$$\Sigma_i (Y_i - \beta_1 X_i - \beta_0)^2$$

- It is easy to verify that each component in the summation is a convex function. Therefore, the sum of them is still a convex function.

3. Nonnegative Weighted Sums

If $f_1, f_2, \ldots, f_n$ are convex functions, and $w_1, w_2, \ldots, w_n \geq 0$, then $f = w_1 f_1 + w_2 f_2 + \cdots + w_n f_n$ is also a convex function.

These properties extend to infinite sums and integrals. For example if $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, and $w(y) \geq 0$ for each $y \in \mathcal{A}$, then the function g defined as

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y)\, dy$$

is convex in x (provided the integral exists).

Proof not required.

Example 1: blind box problem (newsvendor)

minimize $\boxed{-(pE[\min(Q,D)] + sE[(Q-D)^+] - cQ)}$

subject to $-Q \le 0$

$\min(Q,D) = \mathrm{Q} - (\mathrm{Q} - \mathrm{D})^+$

Do not require differentiability of the cdf of D

$$\boxed{\overbrace{(p-c)(-Q) + (p-s)E\underbrace{[\max(Q-D,0)]}_{\text{maximum of convex functions (in Q)}}}^{\text{nonnegative weighted sums of convex functions (in Q)}}}$$

Convex!

Example 1: blind box problem (newsvendor)

Goal: show the convexity of $f(Q) = E[\max(Q - D, 0)]$

Step 1: Given D, $Q - D$ and 0 are both convex functions in Q

Step 2: Because the maximum of a set of convex functions is convex, $\max(Q - D, 0)$ is a convex function in Q

Step 3: Because the weighted integration of convex functions is convex,

$$f(Q) = \int_0^\infty h(x) \max(Q - x, 0)\, dx$$

is a convex function, where $h(x)$ is the pdf of the random variable D

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y)\, dy$$

More Examples

Example 2

- If f is convex (**may not be differentiable**), then $F(x) = \frac{1}{x}\int_0^x f(t)dt$ is also convex on $(0, \infty)$.

If we set $t = xz$ and let z be the integration variable, then the integration domain is changed to $[0, 1]$. As a result,

$$\begin{aligned} F(x) &= \frac{1}{x}\int_0^x f(t)dt \\ &= \frac{1}{x}\int_0^1 f(xz)d(xz) \\ &= \int_0^1 f(xz)dz \end{aligned}$$

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y)\, dy$$

Example 3

For $\boldsymbol{x} \in R^n$, we denote by $x_{[i]}$ the ith largest component of $\boldsymbol{x}$, i.e.,

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}.$$

Let $f(\boldsymbol{x})$ be a function defined as follows

$$f(\boldsymbol{x}) = \sum_{i=1}^{r} x_{[i]},$$

i.e., the sum of the r largest elements of $\boldsymbol{x}$.

Show that $f(\boldsymbol{x})$ is a convex function.

Example 3

Let's consider a concrete case where $\boldsymbol{x} \in R^3$ and $f(\boldsymbol{x}) = \sum_{i=1}^{2} x_{[i]}$, i.e., the sum of the 2 largest elements of $\boldsymbol{x}$. Observe that

$$f(\boldsymbol{x}) = \max(x_1 + x_2, x_1 + x_3, x_2 + x_3),$$

where x_i is the ith component of $\boldsymbol{x}$.

Are you able to show that $f(\boldsymbol{x})$ is a convex function? How?

Example 3

First we show that $\mathrm{g}(\boldsymbol{x}) = x_1 + x_2$ is a convex function:

$$\begin{aligned} \mathrm{g}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) &= \lambda(x_1 + x_2) + (1-\lambda)(y_1 + y_2) \\ &= \lambda\mathrm{g}(\mathbf{x}) + (1-\lambda)\mathrm{g}(\mathbf{y}) \end{aligned}$$

We then apply the result that the maximum of a set of convex functions is convex to obtain that

$$f(\boldsymbol{x}) = \max(x_1 + x_2, x_1 + x_3, x_2 + x_3)$$

is a convex function.

Example 4

For $\boldsymbol{x} \in R^n$, we denote by $|\boldsymbol{x}|$ the vector whose ith element is the absolute value of the ith element of $\boldsymbol{x}$. We then denote by $|x|_{[i]}$ the ith largest component of $|\boldsymbol{x}|$. Let $f(\boldsymbol{x})$ be a function defined as follows

$$f(\boldsymbol{x}) = \sum_{i=1}^{r} |x|_{[i]},$$

i.e., the sum of the r largest elements of $|\boldsymbol{x}|$.

Show that $f(\boldsymbol{x})$ is a convex function.

Example 4

Let's consider a concrete case where $\boldsymbol{x} \in R^3$ and $f(\boldsymbol{x}) = \sum_{i=1}^{2} |x|_{[i]}$, i.e., the sum of the 2 largest elements of $|\boldsymbol{x}|$. Observe that

$$f(\boldsymbol{x}) = \max(|x_1| + |x_2|, |x_1| + |x_3|, |x_2| + |x_3|)\,,$$

where x_i is the ith component of $\boldsymbol{x}$.

Are you able to show that $f(\boldsymbol{x})$ is a convex function? How?

Example 4

First we show that $\mathrm{g}(\boldsymbol{x}) = |x_1|$ is a convex function:

$$\begin{aligned}\mathrm{g}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) &= |\lambda x_1 + (1-\lambda)\mathrm{y}_1| \\ &\leq |\lambda x_1| + |(1-\lambda)\mathrm{y}_1| = \lambda\mathrm{g}(\mathbf{x}) + (1-\lambda)\mathrm{g}(\mathbf{y})\end{aligned}$$

We then apply the result that weighted sum of convex functions is convex to obtain that $h(x) = |x_1| + |x_2|$ is a convex function.

Lastly, we apply the result that the maximum of a set of convex functions is convex to obtain that

$$f(\boldsymbol{x}) = \max(|x_1| + |x_2|, |x_1| + |x_3|, |x_2| + |x_3|)$$

is a convex function.

More on Operations Preserving Convexity

Composition

If f is convex, then $g(\boldsymbol{x}) = f(a\boldsymbol{x} + \boldsymbol{b})$ – c is also convex

Proof:

$$
\begin{aligned}
& g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\
&= f[a(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + \boldsymbol{b}] - \mathrm{c} \\
&= f[\lambda\,(a\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda)(a\boldsymbol{y} + \boldsymbol{b})] - \mathrm{c} \\
&\leq \lambda\, f(a\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda) f(a\boldsymbol{y} + \boldsymbol{b}) - \mathrm{c} \\
&= \lambda\,(f(a\boldsymbol{x} + \boldsymbol{b}) - \mathrm{c}) + (1-\lambda)\,(f(a\boldsymbol{y} + \boldsymbol{b}) - \mathrm{c}) \\
&= \lambda\, g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y})
\end{aligned}
$$

Example 5

For $\boldsymbol{x}, \boldsymbol{y} \in R^n$, we denote by $||\boldsymbol{x} - \boldsymbol{y}||$ the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, i.e.,

$$||\boldsymbol{x} - \boldsymbol{y}|| = \sqrt{\sum_{\mathrm{i}=1}^{\mathrm{n}} (x_i - y_i)^2}.$$

Fix $\boldsymbol{y}$, show that $f(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{y}||$ is a convex function.

Example 5

Let's consider a concrete case where $\boldsymbol{x}, \boldsymbol{y} \in R^2$. It is easy to show that $g(\boldsymbol{x}) = \sqrt{x_1^2 + x_2^2}$ is a convex function (what is the epigraph of $g(\boldsymbol{x})$?)

Then by the Composition result, we know that $f(\boldsymbol{x}) = g(\boldsymbol{x} - \boldsymbol{y})$ is a convex function.

Example 6

For $\boldsymbol{x}, \boldsymbol{y} \in R^n$, we denote by $||\boldsymbol{x} - \boldsymbol{y}||$ the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, i.e.,

$$||\boldsymbol{x} - \boldsymbol{y}|| = \sqrt{\sum_{\mathrm{i}=1}^{\mathrm{n}} (x_i - y_i)^2}.$$

$f(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||$ is a also a convex function.

Proof not required.

Minimization

- We already know that the maximum or supremum of a set of convex functions is convex

- Some <u>special</u> forms (NOT ALL) of minimization also preserve convexity

- If f is convex in (x, y), and C is **a convex nonempty set**, then the function

$$g(x) = \inf_{y \in C} f(x, y)$$

is convex.

Proof not required.

Example 7

Give a point $\boldsymbol{x} \in R^n$, and a convex set $\mathrm{C} \in R^n$. Let

$$f(\boldsymbol{x}) = inf_{\boldsymbol{y} \in C} ||\boldsymbol{x} - \boldsymbol{y}||,$$

i.e., the distance between $\boldsymbol{x}$ and the nearest point of C.

Show that $f(\boldsymbol{x})$ is a convex function.

Example 7

We know that $h(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||$ is a convex function

By the Minimization result, we can obtain that $f(\boldsymbol{x}) = inf_{\boldsymbol{y} \in C} ||\boldsymbol{x} - \boldsymbol{y}||$ is a convex function

Example 8

What about the distance between $\boldsymbol{x}$ and the farthest point of C?

$$f(\boldsymbol{x}) = sup_{\boldsymbol{y}\in C}||\boldsymbol{x} - \boldsymbol{y}||,$$

Do we need the condition that C is a convex set?

The maximum of a set of convex functions is convex.

We impose no condition on the set. We only require $f(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{y}||$ is convex for fixed $\boldsymbol{y}$.