Lecture 19: Convex Problems

Zizhuo Wang

Institute of Data and Decision Analytics (iDDA)
Chinese University of Hong Kong, Shenzhen

Nov 14, 2018

Announcements

- Homework 6 due today
- Homework 7 due next Wednesday (11/21) at noon

Recap: Optimality Conditions

For unconstrained problems

- First-order necessary condition
- Second-order necessary condition
- Second-order sufficient condition

For constrained problems

- We mainly consider the first-order necessary conditions

These conditions are for *local* minimizers (maximizers)

Recap

To derive the first-order necessary conditions for a general problem, we used the following approach

1. Find the feasible direction and descent direction sets
2. Find the conditions such that their intersection is an empty set
3. Use duality arguments/alternative systems to obtain neater conditions (exist something rather than not exist something)

We did these for linear equality constrained problems and briefly mentioned the case with linear inequality constraints.

- Then we derived the optimality conditions for more general cases — the KKT conditions.
- We presented the procedures to find the KKT conditions for general optimization problems.

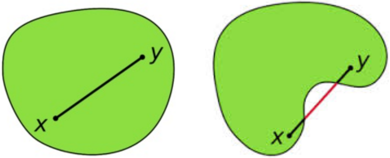
Toward Global Optimality

We have been discussing local minimizers

- When is a local minimizer also a global minimizer?
- We present a class of optimization problems that has this property - convex optimization

Definition (Convex Set)

A set $S \subseteq \mathbb{R}^n$ is *convex* if for any $\mathbf{x}$, $\mathbf{y} \in S$, and any $\lambda \in [0, 1]$, $\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \in S$.



(a) Convex set (b) Non-convex set

- Intersection of convex sets is still a convex set

Definition (Convex Combinations)

For any $\mathbf{x}_1, ..., \mathbf{x}_n$ and $\lambda_1, ..., \lambda_n \geq 0$ satisfying $\lambda_1 + \cdots + \lambda_n = 1$, we call $\sum_{i=1}^{n} \lambda_i \mathbf{x}_i$ a convex combination of $\mathbf{x}_1, ..., \mathbf{x}_n$.

Convex Functions

Definition (Convex function)

A function f on Ω is said to be *convex* if for every $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ and any $0 \leq \alpha \leq 1$,

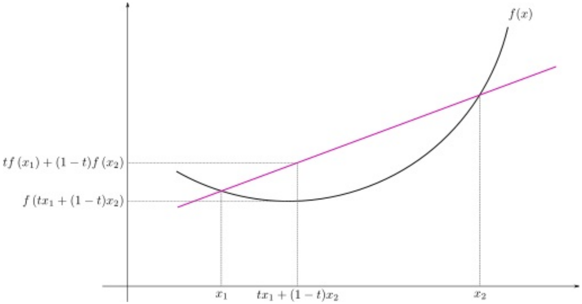
$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$$

Definition (Concave function)

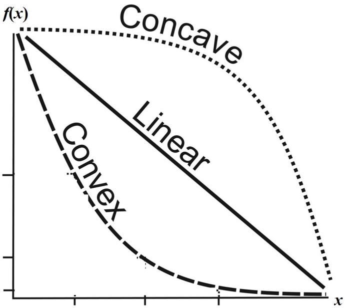
We call f a concave function if and only if $-f$ is a convex function or for any $\mathbf{x}_1, \mathbf{x}_2$ and $0 \leq \alpha \leq 1$,

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$$

Illustration of Convex Function



Convex Functions



Some examples of convex functions:

- $y = x$, $y = x^2$, $y = e^x$, $y = |x|$

Some examples of concave functions:

- $y = x$, $y = \sqrt{x}$, $y = \log x$

Test Convexity

Proposition

If $f(\mathbf{x})$ is second-order differentiable. Then $f(\cdot)$ is convex if and only if its Hessian matrix is positive semi-definite (PSD) throughout the defined region.

- In one-dimension case, this means the second-order derivative is non-negative
- Taking second-order derivative is usually the easiest way to test convexity
- Example: whether $x \log x$, $||\mathbf{x}||_2$ is convex?

Otherwise, one typically tests convexity by definition

For Concavity

Proposition

If $f(\mathbf{x})$ is second-order differentiable. Then $f(\cdot)$ is concave if and only if its Hessian matrix is negative semi-definite throughout the defined region.

- In one-dimension case, this means the second-order derivative is non-positive
- Example: $x^{1/3}$, $\log x$

Properties of Convex Functions

Proposition

If $a_1, ..., a_m \geq 0$, and $f_1, ..., f_m$ are convex (concave) functions, then $a_1 f_1 + \cdots + a_m f_m$ is a convex (concave) function.

- Examples: $x_1^2 + x_2^2$, $e^x + |x|$

Proposition

If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex (concave), then $g(\mathbf{x}) = f(A\mathbf{x} + \mathbf{b})$ is convex (concave).

- Examples: e^{2x+3}, $(x_1 - x_2)^2 + (x_2 + x_3)^2$, $||A\mathbf{x} - \mathbf{b}||_2$, $\log(-2x_1 + 3x_2 + 5)$ (concave)

More Properties

Proposition

If $f_1, ..., f_m$ are convex functions, then $f(x) = \max\{f_1(x), ..., f_m(x)\}$ is a convex function (this can be extended to uncountably many).

- Examples: $|x|$, $\max\{\mathbf{a}_i^T\mathbf{x} + b_i\}$

Proposition

If $f_1, ..., f_m$ are concave function, then $f(x) = \min\{f_1(x), ..., f_m(x)\}$ is a concave function (this can be extended to uncountably many).

- Examples: $-|x|$, $\min\{\mathbf{a}_i^T\mathbf{x} + b_i\}$

Another Example

Consider the linear program

$$\begin{aligned} \text{minimize}_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

Given A and $\mathbf{b}$ fixed, the optimal value can be viewed as a function of $\mathbf{c}$. We denote the function by $V(\mathbf{c})$.

- In sensitivity analysis, we studied how $V(\mathbf{c})$ changes with $\mathbf{c}$.

Theorem
$V(\mathbf{c})$ is a concave function of $\mathbf{c}$.

- Because it is the minimum of a set of linear functions

$$V(\mathbf{c}) = \min_{\{\mathbf{x}:A\mathbf{x}=\mathbf{b},\mathbf{x}\geq 0\}} \{\mathbf{c}^T\mathbf{x}\}$$

- One can also prove by definition

How Does Convexity Help?

Theorem

Let $f(\cdot)$ be a convex function and Ω be a convex set. Then any local minimizer for the optimization problem:

$$\begin{array}{rl} \textit{minimize}_{\mathbf{x}} & f(\mathbf{x}) \\ \textit{s.t.} & \mathbf{x} \in \Omega \end{array}$$

is a global minimizer.

Proof. We prove by contradiction. Assume $\mathbf{x}^*$ is a local minimizer, however, there exists $\mathbf{x}' \in \Omega$ such that $f(\mathbf{x}') < f(\mathbf{x}^*)$. Then by the property of convex function, for any $0 < \alpha < 1$,

$$f(\alpha \mathbf{x}' + (1-\alpha)\mathbf{x}^*) \leq \alpha f(\mathbf{x}') + (1-\alpha) f(\mathbf{x}^*) < f(\mathbf{x}^*)$$

This contradicts with that $\mathbf{x}^*$ is a local minimizer. □

Furthermore..

Theorem

Let $f(\cdot)$ be a convex function and Ω be a convex set. Then the KKT conditions for the optimization problem

$$\begin{array}{rl} \text{minimize}_{\mathbf{x}} & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in \Omega \end{array}$$

are sufficient for global optimality.

Corollary

Let $f(\cdot)$ be a concave function and Ω be a convex set. Then the KKT conditions for the optimization problem

$$\begin{array}{rl} \text{maximize}_{\mathbf{x}} & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in \Omega \end{array}$$

are sufficient for global optimality.

Convex Optimization Problem

Therefore, convexity/concavity plays a very important role in optimization problems

We call the optimization problems that either

- Minimize a convex function over a convex feasible region
- Maximize a concave function over a convex feasible region

convex optimization problems

Otherwise, we call it non-convex optimization problem

In optimization world, convex/non-convex is the main watershed to determine whether a problem is easy or hard

Constraint Types

What constraints would make the feasible region convex?

Proposition

Let f be a convex (concave) function. Then for any number c, the set $\Gamma_c = \{\mathbf{x} : f(\mathbf{x}) \leq (\geq) c\}$ is a convex set.

Therefore,

- If we have constraint $g(\mathbf{x}) \leq 0$, and $g(\mathbf{x})$ is convex, then this is a convex constraint
- If we have constraint $g(\mathbf{x}) \geq 0$, and $g(\mathbf{x})$ is concave, then this is a convex constraint
- Linear constraints are always convex constraints
- Sometimes, even if a constraint doesn't appear to be in the above form, it still could be a convex constraint

Being able to identify convex problem is an important skill.

Example

Is this a convex optimization problem?

$$\begin{array}{ll} \text{minimize} & 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \\ \text{subject to} & x_1^2 + x_2^2 \leq 5 \\ & 3x_1 + x_2 \geq 3 \end{array}$$

- Answer: Yes

What if we change $x_1^2 + x_2^2 \leq 5$ to $x_1^2 + x_2^2 \geq 5$?

- Then it won't be a convex optimization problem.

Example

How about

$$\begin{aligned} \text{minimize}_{\mathbf{x}} \quad & \mathbf{x}^T Q \mathbf{x} - \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & C\mathbf{x} \geq \mathbf{d} \\ & \mathbf{x} \geq 0 \end{aligned}$$

- It is a convex optimization problem if and only if Q is PSD.

Another Example

Consider the optimization problem:

$$\begin{aligned} \text{maximize}_{x,y,z} \quad & xyz \\ \text{s.t.} \quad & x + 2y + 3z \leq 3 \\ & x, y, z \geq 0 \end{aligned}$$

In order for a maximization problem to be a convex optimization problem, we need the objective function to be concave

- However, xyz is not a concave function in x, y, z

But we can transform this into maximizing $\log(xyz)$. The optimization problem will become

$$\begin{aligned} \text{maximize} \quad & \log x + \log y + \log z \\ \text{s.t.} \quad & x + 2y + 3z \leq 3 \\ & x, y, z \geq 0 \end{aligned}$$

which is a convex optimization problem

Constraint/Variable Transformations

In practice, it is often an important task to verify whether a problem is a convex optimization problem.

- Even if it doesn't appear to be a convex problem, it is sometimes possible to transform into one.
- Some monotone mappings or variable substitutions can be helpful.

For the constraints, verifying convexity may go beyond checking the sufficient conditions.

- Remember the ultimate criterion is that the feasible region is a convex set

Verifying Constraints

Sometimes one has to look into the region the constraints define (remember that was the ultimate definition of convex constraint):

- $x^3 - 1 \leq 0$: x^3 is not a convex function, however, this constraint defines a convex feasible region ($x \leq 1$) thus is a convex constraint.
- $z^2 - xy \leq 0$, $x, y, z \geq 0$: $z^2 - xy$ is not a convex function. The Hessian matrix is

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

  The eigenvalues are -1, 1 and 2
- However, this gives a convex region

Proof

We prove by definition. We show for any (x_1, y_1, z_1) and (x_2, y_2, z_2) that satisfy these constraints, and for any $0 \leq \alpha \leq 1$,

$$(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2, \alpha z_1 + (1-\alpha)z_2)$$

also satisfy this constraint (i.e., the feasible region is convex)

We have

$$(\alpha z_1 + (1-\alpha)z_2)^2 = \alpha^2 z_1^2 + (1-\alpha)^2 z_2^2 + 2\alpha(1-\alpha)z_1 z_2$$

By assumption, $z_1^2 \leq x_1 y_1$, $z_2^2 \leq x_2 y_2$. Also when $x, y, z \geq 0$, we have

$$2z_1 z_2 \leq 2\sqrt{x_1 y_1 x_2 y_2} \leq x_1 y_2 + x_2 y_1$$

The last inequality is because of the inequality of arithmetic and geometric means

Proof Continued

Therefore

$$
\begin{aligned}
& (\alpha z_1 + (1-\alpha) z_2))^2 \\
= \ & \alpha^2 z_1^2 + (1-\alpha)^2 z_2^2 + 2\alpha(1-\alpha) z_1 z_2 \\
\leq \ & \alpha^2 x_1 y_1 + (1-\alpha)^2 x_2 y_2 + \alpha(1-\alpha)(x_1 y_2 + x_2 y_1) \\
= \ & (\alpha x_1 + (1-\alpha) x_2)(\alpha y_1 + (1-\alpha) y_2)
\end{aligned}
$$

Thus $z^2 - xy \leq 0$, $x, y, z \geq 0$ is a convex region. $\square$

We don't often use this method, but this can be viewed as a last resort.