Optimization methods
Lecture 3: Optimality conditions

Alexandr Katrutsa

Modern State of Artificial Intelligence Masters Program
Moscow Institute of Physics and Technology

Brief reminder of the previous lecture

- Matrix calculus

Brief reminder of the previous lecture

- Matrix calculus
- Non-differentiable convex functions and subdifferential

Brief reminder of the previous lecture

- Matrix calculus
- Non-differentiable convex functions and subdifferential
- Automatic differentiation technique

Brief reminder of the previous lecture

- Matrix calculus
- Non-differentiable convex functions and subdifferential
- Automatic differentiation technique
- L-smooth convex functions

Brief reminder of the previous lecture

- Matrix calculus
- Non-differentiable convex functions and subdifferential
- Automatic differentiation technique
- L-smooth convex functions
- Strongly convex functions properties

Plan for today

- Optimality conditions: general concept

Plan for today

- Optimality conditions: general concept
- First order optimality condition

Plan for today

- Optimality conditions: general concept
- First order optimality condition
- Equality constrained optimization peoblems

Plan for today

- Optimality conditions: general concept
- First order optimality condition
- Equality constrained optimization peoblems
- KKT optimality conditions

Plan for today

- Optimality conditions: general concept
- First order optimality condition
- Equality constrained optimization peoblems
- KKT optimality conditions
- Slater regularity conditions

Motivation

Question 1

In what cases a solution of an optimization problem exists?

Motivation

Question 1

In what cases a solution of an optimization problem exists?

Question 2

How to verify that the point is not a solution?

Motivation

Question 1

In what cases a solution of an optimization problem exists?

Question 2

How to verify that the point is not a solution?

Question 3

How to find a solution of an optimization problem?

Existence

Let $\mathcal{X} \subset \mathbb{R}^n$ be a compact set and let $f : \mathcal{X} \to \mathbb{R}$ is continuous function in $\mathcal{X}$. Then the point of global minimum does exist in $\mathcal{X}$.

- To be sure that we find solution, we need compact set

Existence

Let $\mathcal{X} \subset \mathbb{R}^n$ be a compact set and let $f : \mathcal{X} \to \mathbb{R}$ is continuous function in $\mathcal{X}$. Then the point of global minimum does exist in $\mathcal{X}$.

- To be sure that we find solution, we need compact set
- Most problems are defined exactly in the compact sets

Existence

Let $\mathcal{X} \subset \mathbb{R}^n$ be a compact set and let $f : \mathcal{X} \to \mathbb{R}$ is continuous function in $\mathcal{X}$. Then the point of global minimum does exist in $\mathcal{X}$.

- To be sure that we find solution, we need compact set
- Most problems are defined exactly in the compact sets
- The source of open feasible sets is domain of some convex functions

Existence

Let $\mathcal{X} \subset \mathbb{R}^n$ be a compact set and let $f : \mathcal{X} \to \mathbb{R}$ is continuous function in $\mathcal{X}$. Then the point of global minimum does exist in $\mathcal{X}$.

- To be sure that we find solution, we need compact set
- Most problems are defined exactly in the compact sets
- The source of open feasible sets is domain of some convex functions
- This issue can be fixed by sequential approximation of the open feasible set by the compact one

General form of optimality conditions

Optimization problem in general form

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

General form of optimality conditions

Optimization problem in general form

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Optimality criterion

Let f be defined in a set $\mathcal{X} \subset \mathbb{R}^n$. Then

General form of optimality conditions

Optimization problem in general form

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Optimality criterion

Let f be defined in a set $\mathcal{X} \subset \mathbb{R}^n$. Then

1. If $\mathbf{x}^*$ is a minimizer of f in $\mathcal{X}$, then $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$

General form of optimality conditions

Optimization problem in general form

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Optimality criterion

Let f be defined in a set $\mathcal{X} \subset \mathbb{R}^n$. Then

1. If $\mathbf{x}^*$ is a minimizer of f in $\mathcal{X}$, then $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$
2. If for some point $\mathbf{x}^* \in \mathcal{X}$ there exists subdifferential $\partial_{\mathcal{X}} f(\mathbf{x}^*)$ and $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, then $\mathbf{x}^*$ is a minimizer of f in $\mathcal{X}$.

General form of optimality conditions

Optimization problem in general form

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Optimality criterion

Let f be defined in a set $\mathcal{X} \subset \mathbb{R}^n$. Then

1. If $\mathbf{x}^*$ is a minimizer of f in $\mathcal{X}$, then $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$
2. If for some point $\mathbf{x}^* \in \mathcal{X}$ there exists subdifferential $\partial_{\mathcal{X}} f(\mathbf{x}^*)$ and $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, then $\mathbf{x}^*$ is a minimizer of f in $\mathcal{X}$.

Q: what drawbacks this criterion has?

Proof

1. If $\mathbf{x}^*$ is a minimizer

Proof

1. If $\mathbf{x}^*$ is a minimizer
   - For any $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$

Proof

1. If $\mathbf{x}^*$ is a minimizer
   - For any $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - This inequality can be re-written in the form: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$

Proof

1. If $\mathbf{x}^*$ is a minimizer
   - For any $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - This inequality can be re-written in the form: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - It means that $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ by definition

Proof

1. If $\mathbf{x}^*$ is a minimizer
   - For any $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - This inequality can be re-written in the form: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - It means that $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ by definition
2. If $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ for some point $\mathbf{x}^*$

Proof

1. If $\mathbf{x}^*$ is a minimizer
   - For any $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - This inequality can be re-written in the form: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - It means that $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ by definition
2. If $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ for some point $\mathbf{x}^*$
   - By definition of subgradient: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle \mathbf{a}, \mathbf{x} - \mathbf{x}^* \rangle$

Proof

1. If $\mathbf{x}^*$ is a minimizer
   - For any $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - This inequality can be re-written in the form: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - It means that $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ by definition
2. If $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ for some point $\mathbf{x}^*$
   - By definition of subgradient: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle \mathbf{a}, \mathbf{x} - \mathbf{x}^* \rangle$
   - If $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, then we have inequality $f(\mathbf{x}) \geq f(\mathbf{x}^*)$

Proof

1. If $\mathbf{x}^*$ is a minimizer
   - For any $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - This inequality can be re-written in the form: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - It means that $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ by definition
2. If $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ for some point $\mathbf{x}^*$
   - By definition of subgradient: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle \mathbf{a}, \mathbf{x} - \mathbf{x}^* \rangle$
   - If $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, then we have inequality $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - It means that $\mathbf{x}^*$ is a minimizer

Geometric interpretation

Necessary condition for unconstrained problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Theorem

If $\mathbf{x}^*$ is a solution of problem (1) and f is differentiable, then $f'(\mathbf{x}^*) = 0$.

Necessary condition for unconstrained problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Theorem

If $\mathbf{x}^*$ is a solution of problem (1) and f is differentiable, then $f'(\mathbf{x}^*) = 0$.

Proof

Necessary condition for unconstrained problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Theorem

If $\mathbf{x}^*$ is a solution of problem (1) and f is differentiable, then $f'(\mathbf{x}^*) = 0$.

Proof

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ and $\lim\limits_{\mathbf{y}\to\mathbf{x}^*} \frac{r(\mathbf{x}^*,\mathbf{y})}{\|\mathbf{x}^*-\mathbf{y}\|_2} = 0 \ (*)$

Necessary condition for unconstrained problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Theorem

If $\mathbf{x}^*$ is a solution of problem (1) and f is differentiable, then $f'(\mathbf{x}^*) = 0$.

Proof

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ and $\lim\limits_{\mathbf{y}\to\mathbf{x}^*} \frac{r(\mathbf{x}^*,\mathbf{y})}{\|\mathbf{x}^*-\mathbf{y}\|_2} = 0 \ (*)$
- If $f'(\mathbf{x}^*) \neq 0$, then consider $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$

Necessary condition for unconstrained problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Theorem

If $\mathbf{x}^*$ is a solution of problem (1) and f is differentiable, then $f'(\mathbf{x}^*) = 0$.

Proof

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ and $\lim\limits_{\mathbf{y} \to \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0 \; (*)$
- If $f'(\mathbf{x}^*) \neq 0$, then consider $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$
- $f(\mathbf{y}(\tau)) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y}(\tau) - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y}(\tau)) = f(\mathbf{x}^*) - \tau \|f'(\mathbf{x}^*)\|_2^2 + r(\mathbf{x}^*, \mathbf{y}(\tau))$

Necessary condition for unconstrained problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Theorem

If $\mathbf{x}^*$ is a solution of problem (1) and f is differentiable, then $f'(\mathbf{x}^*) = 0$.

Proof

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ and $\lim\limits_{\mathbf{y}\to\mathbf{x}^*} \frac{r(\mathbf{x}^*,\mathbf{y})}{\|\mathbf{x}^*-\mathbf{y}\|_2} = 0 \ (*)$
- If $f'(\mathbf{x}^*) \neq 0$, then consider $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$
- $f(\mathbf{y}(\tau)) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y}(\tau) - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y}(\tau)) = f(\mathbf{x}^*) - \tau \|f'(\mathbf{x}^*)\|_2^2 + r(\mathbf{x}^*, \mathbf{y}(\tau))$
- From $(*)$ follows that there exists $\bar{\tau}$ such that for all $\tau \in (0, \bar{\tau})$ holds $f(\mathbf{y}(\tau)) - f(\mathbf{x}^*) \leq -\frac{\tau}{2}\|f'(\mathbf{x}^*)\|_2^2 < 0$

Necessary condition for unconstrained problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Theorem

If $\mathbf{x}^*$ is a solution of problem (1) and f is differentiable, then $f'(\mathbf{x}^*) = 0$.

Proof

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ and $\lim\limits_{\mathbf{y} \to \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0 \; (*)$
- If $f'(\mathbf{x}^*) \neq 0$, then consider $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$
- $f(\mathbf{y}(\tau)) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y}(\tau) - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y}(\tau)) = f(\mathbf{x}^*) - \tau \|f'(\mathbf{x}^*)\|_2^2 + r(\mathbf{x}^*, \mathbf{y}(\tau))$
- From $(*)$ follows that there exists $\bar{\tau}$ such that for all $\tau \in (0, \bar{\tau})$ holds $f(\mathbf{y}(\tau)) - f(\mathbf{x}^*) \leq -\frac{\tau}{2} \|f'(\mathbf{x}^*)\|_2^2 < 0$
- Thus, $\mathbf{x}^*$ is not a minimizer, that is a contradiction.

Remarks

- This theorem is a special case of the general optimality condition

Remarks

- This theorem is a special case of the general optimality condition
- The additional proof is provided to introduce the fact on the descent property of the direction $-f'(\mathbf{x})$

If function is convex...

Theorem

If the objective function in problem (1) is convex, then $\mathbf{x}^*$ is a global minimum iff $f'(\mathbf{x}^*) = 0$

If function is convex...

Theorem

If the objective function in problem (1) is convex, then $\mathbf{x}^*$ is a global minimum iff $f'(\mathbf{x}^*) = 0$

Proof

If function is convex...

Theorem

If the objective function in problem (1) is convex, then $\mathbf{x}^*$ is a global minimum iff $f'(\mathbf{x}^*) = 0$

Proof

- If $\mathbf{x}^*$ is a global minimum, then $\mathbf{x}^*$ is local minimum

If function is convex...

Theorem

If the objective function in problem (1) is convex, then $\mathbf{x}^*$ is a global minimum iff $f'(\mathbf{x}^*) = 0$

Proof

- If $\mathbf{x}^*$ is a global minimum, then $\mathbf{x}^*$ is local minimum
- Thus, $f'(\mathbf{x}^*) = 0$ according to the previous theorem

If function is convex...

Theorem

If the objective function in problem (1) is convex, then $\mathbf{x}^*$ is a global minimum iff $f'(\mathbf{x}^*) = 0$

Proof

- If $\mathbf{x}^*$ is a global minimum, then $\mathbf{x}^*$ is local minimum
- Thus, $f'(\mathbf{x}^*) = 0$ according to the previous theorem
- Let $\mathbf{x}^*$ be a point such that $f'(\mathbf{x}^*) = 0$ and f is convex

If function is convex...

Theorem

If the objective function in problem (1) is convex, then $\mathbf{x}^*$ is a global minimum iff $f'(\mathbf{x}^*) = 0$

Proof

- If $\mathbf{x}^*$ is a global minimum, then $\mathbf{x}^*$ is local minimum
- Thus, $f'(\mathbf{x}^*) = 0$ according to the previous theorem
- Let $\mathbf{x}^*$ be a point such that $f'(\mathbf{x}^*) = 0$ and f is convex
- Then according to the FO criterion

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle = f(\mathbf{x}^*)$$

If function is convex...

Theorem

If the objective function in problem (1) is convex, then $\mathbf{x}^*$ is a global minimum iff $f'(\mathbf{x}^*) = 0$

Proof

- If $\mathbf{x}^*$ is a global minimum, then $\mathbf{x}^*$ is local minimum
- Thus, $f'(\mathbf{x}^*) = 0$ according to the previous theorem
- Let $\mathbf{x}^*$ be a point such that $f'(\mathbf{x}^*) = 0$ and f is convex
- Then according to the FO criterion

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle = f(\mathbf{x}^*)$$

- It means that $\mathbf{x}^*$ is a global minimum

Second-order sufficient condition

Theorem

Let f be twice continuously differentiable function. A point $\mathbf{x}^*$ satisfies equation $f'(\mathbf{x}^*) = 0$. If $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ for all $\mathbf{s} \neq 0$, then $\mathbf{x}^*$ is a point of local minimum.

Proof by contradiction

Second-order sufficient condition

Theorem

Let f be twice continuously differentiable function. A point $\mathbf{x}^*$ satisfies equation $f'(\mathbf{x}^*) = 0$. If $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ for all $\mathbf{s} \neq 0$, then $\mathbf{x}^*$ is a point of local minimum.

Proof by contradiction

- Assume that there exists some point $\mathbf{y}$ close to $\mathbf{x}^*$ such that $f(\mathbf{y}) < f(\mathbf{x}^*)$

Second-order sufficient condition

Theorem

Let f be twice continuously differentiable function. A point $\mathbf{x}^*$ satisfies equation $f'(\mathbf{x}^*) = 0$. If $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ for all $\mathbf{s} \neq 0$, then $\mathbf{x}^*$ is a point of local minimum.

Proof by contradiction

- Assume that there exists some point $\mathbf{y}$ close to $\mathbf{x}^*$ such that $f(\mathbf{y}) < f(\mathbf{x}^*)$
- Then consider Taylor expansion $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$

Second-order sufficient condition

Theorem

Let f be twice continuously differentiable function. A point $\mathbf{x}^*$ satisfies equation $f'(\mathbf{x}^*) = 0$. If $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ for all $\mathbf{s} \neq 0$, then $\mathbf{x}^*$ is a point of local minimum.

Proof by contradiction

- Assume that there exists some point $\mathbf{y}$ close to $\mathbf{x}^*$ such that $f(\mathbf{y}) < f(\mathbf{x}^*)$
- Then consider Taylor expansion $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$
- From stationarity follows that $f(\mathbf{y}) - f(\mathbf{x}^*) = \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$

Second-order sufficient condition

Theorem

Let f be twice continuously differentiable function. A point $\mathbf{x}^*$ satisfies equation $f'(\mathbf{x}^*) = 0$. If $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ for all $\mathbf{s} \neq 0$, then $\mathbf{x}^*$ is a point of local minimum.

Proof by contradiction

- Assume that there exists some point $\mathbf{y}$ close to $\mathbf{x}^*$ such that $f(\mathbf{y}) < f(\mathbf{x}^*)$
- Then consider Taylor expansion $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$
- From stationarity follows that $f(\mathbf{y}) - f(\mathbf{x}^*) = \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$
- If $\mathbf{y} \to \mathbf{x}^*$, then we have a direction $\mathbf{z} \neq 0$ such that $\mathbf{z}^\top f''(\mathbf{x}^*)\mathbf{z} \leq 0$, that is contradiction

Saddle points

Definition

A point $\mathbf{y}$ is called saddle point for a function f if there are directions $\mathbf{z}_1$ and $\mathbf{z}_2$ such that $f(\mathbf{y}+\mathbf{z}_1) > f(\mathbf{y})$, but $f(\mathbf{y}+\mathbf{z}_2) < f(\mathbf{y})$

Summary on unconstrained problems

- Use subdifferential for non-differentiable case

Summary on unconstrained problems

- Use subdifferential for non-differentiable case
- Use FOOC for convex differentiable function

Summary on unconstrained problems

- Use subdifferential for non-differentiable case
- Use FOOC for convex differentiable function
- Use second order sufficient condition for non-convex twice continuously differentiable

Summary on unconstrained problems

- Use subdifferential for non-differentiable case
- Use FOOC for convex differentiable function
- Use second order sufficient condition for non-convex twice continuously differentiable
- Saddle points are possible in the non-convex settings

Differential criterion for constrained problem

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \tag{2}$$

Theorem

A point $\mathbf{x}^*$ is a solution of the problem (2), where f is convex function, iff $\mathbf{x}^* \in \mathcal{X}$ and $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ for all $\mathbf{y} \in \mathcal{X}$.

Differential criterion for constrained problem

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \tag{2}$$

Theorem

A point $\mathbf{x}^*$ is a solution of the problem (2), where f is convex function, iff $\mathbf{x}^* \in \mathcal{X}$ and $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ for all $\mathbf{y} \in \mathcal{X}$.

Proof

Differential criterion for constrained problem

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \tag{2}$$

Theorem

A point $\mathbf{x}^*$ is a solution of the problem (2), where f is convex function, iff $\mathbf{x}^* \in \mathcal{X}$ and $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ for all $\mathbf{y} \in \mathcal{X}$.

Proof

- Let $\mathbf{x}^* \in \mathcal{X}$ and the inequality holds. Then according to the first order criterion of convexity function f: $f(\mathbf{y}) \geq f(\mathbf{x}^*)$

Differential criterion for constrained problem

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \tag{2}$$

Theorem

A point $\mathbf{x}^*$ is a solution of the problem (2), where f is convex function, iff $\mathbf{x}^* \in \mathcal{X}$ and $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ for all $\mathbf{y} \in \mathcal{X}$.

Proof

- Let $\mathbf{x}^* \in \mathcal{X}$ and the inequality holds. Then according to the first order criterion of convexity function f: $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Let $\mathbf{x}^*$ be a solution of the problem (2), but there exists $\tilde{\mathbf{y}}$ such that $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$

Differential criterion for constrained problem

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \tag{2}$$

Theorem

A point $\mathbf{x}^*$ is a solution of the problem (2), where f is convex function, iff $\mathbf{x}^* \in \mathcal{X}$ and $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ for all $\mathbf{y} \in \mathcal{X}$.

Proof

- Let $\mathbf{x}^* \in \mathcal{X}$ and the inequality holds. Then according to the first order criterion of convexity function f: $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Let $\mathbf{x}^*$ be a solution of the problem (2), but there exists $\tilde{\mathbf{y}}$ such that $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- Consider a point $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1-t)\mathbf{x}^*$, $t \in [0, 1]$

Differential criterion for constrained problem

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \tag{2}$$

Theorem

A point $\mathbf{x}^*$ is a solution of the problem (2), where f is convex function, iff $\mathbf{x}^* \in \mathcal{X}$ and $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ for all $\mathbf{y} \in \mathcal{X}$.

Proof

- Let $\mathbf{x}^* \in \mathcal{X}$ and the inequality holds. Then according to the first order criterion of convexity function f: $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Let $\mathbf{x}^*$ be a solution of the problem (2), but there exists $\tilde{\mathbf{y}}$ such that $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- Consider a point $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1-t)\mathbf{x}^*$, $t \in [0, 1]$
- Then $\frac{d}{dt} f(\mathbf{z}(t))\Big|_{t=0} = \langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$

Differential criterion for constrained problem

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \tag{2}$$

Theorem

A point $\mathbf{x}^*$ is a solution of the problem (2), where f is convex function, iff $\mathbf{x}^* \in \mathcal{X}$ and $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ for all $\mathbf{y} \in \mathcal{X}$.

Proof

- Let $\mathbf{x}^* \in \mathcal{X}$ and the inequality holds. Then according to the first order criterion of convexity function f: $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Let $\mathbf{x}^*$ be a solution of the problem (2), but there exists $\tilde{\mathbf{y}}$ such that $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- Consider a point $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1-t)\mathbf{x}^*$, $t \in [0, 1]$
- Then $\frac{d}{dt} f(\mathbf{z}(t))\Big|_{t=0} = \langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- It means that for sufficiently small t the following inequality holds $f(\mathbf{z}(t)) < f(\mathbf{x}^*)$, and we have a contradiction.

Equality constraints

Problem statement

$$
\begin{aligned}
& f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathbb{R}^n} \\
\text{s.t. } & g_i(\mathbf{x}) = 0, \; i = 1, \ldots, m
\end{aligned}
$$

Equality constraints

Problem statement

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathbb{R}^n}$$

$$\text{s.t. } g_i(\mathbf{x}) = 0,\ i = 1, \dots, m$$

Lagrangian

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

Geometric interpretation

From equality constraints to inequalities

Minimization problem

$$\begin{aligned} &\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \\ \text{s.t. } &g_i(\mathbf{x}) = 0,\ i = 1, \ldots, m \\ &h_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, p \end{aligned}$$

Lagrangian

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{p} \mu_j h_j(\mathbf{x})$$

KKT necessary conditions

- $g_i(\mathbf{x}^*) = 0$, $i = 1, \ldots, m$

KKT necessary conditions

- $g_i(\mathbf{x}^*) = 0$, $i = 1, \ldots, m$
- $h_j(\mathbf{x}^*) \leq 0$, $j = 1, \ldots, p$

KKT necessary conditions

- $g_i(\mathbf{x}^*) = 0$, $i = 1, \ldots, m$
- $h_j(\mathbf{x}^*) \leq 0$, $j = 1, \ldots, p$
- $\mu_j^* \geq 0$, $j = 1, \ldots, p$

KKT necessary conditions

- $g_i(\mathbf{x}^*) = 0$, $i = 1, \ldots, m$
- $h_j(\mathbf{x}^*) \leq 0$, $j = 1, \ldots, p$
- $\mu_j^* \geq 0$, $j = 1, \ldots, p$
- $\mu_j^* h_j(\mathbf{x}^*) = 0$, $j = 1, \ldots, p$

KKT necessary conditions

- $g_i(\mathbf{x}^*) = 0$, $i = 1, \ldots, m$
- $h_j(\mathbf{x}^*) \leq 0$, $j = 1, \ldots, p$
- $\mu_j^* \geq 0$, $j = 1, \ldots, p$
- $\mu_j^* h_j(\mathbf{x}^*) = 0$, $j = 1, \ldots, p$
- $L'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = 0$

Slater regularity condition

Slater regularity

There exists a point $\bar{\mathbf{x}}$ inside the interior of convex feasible set such that $f_i(\bar{\mathbf{x}}) < 0$ and $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$

Slater regularity condition

Slater regularity

There exists a point $\bar{\mathbf{x}}$ inside the interior of convex feasible set such that $f_i(\bar{\mathbf{x}}) < 0$ and $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$

Theorem

If a problem is convex and there exists $\mathbf{x}$ inside the interior of the feasible set, i.e. inequality constraints hold with strict inequalities, then the KKT conditions are also sufficient.

Summary on optimality conditions

- Optimality conditiuons: what and why?

Summary on optimality conditions

- Optimality conditiuons: what and why?
- Unconstrained case

Summary on optimality conditions

- Optimality conditiuons: what and why?
- Unconstrained case
- Equality constrained problems

Summary on optimality conditions

- Optimality conditiuons: what and why?
- Unconstrained case
- Equality constrained problems
- Inequality contrained problems

Summary on optimality conditions

- Optimality conditiuons: what and why?
- Unconstrained case
- Equality constrained problems
- Inequality contrained problems
- KKT conditions and Slater condition