Optimization

Lecture 4: The theory of Unconstrained optimization

Parin Chaipunya
KMUTT
└ Mathematics @ Faculty of Science
└ The Joint Graduate School of Energy and Environments

Areas of research:

- Multi-agent optimization: Bilevel programs, Game theory
- Optimization modeling: mainly focused on energy and environmental applications

parin.cha@kmutt.ac.th
parinchaipunya.com
github.com/parinchaipunya

Last update: Januaray 2026

Table of contents

Section 1

Unconstrained optimization

Unconstrained optimization

Unconstrained optimization refers to the case where $\mathcal{X}^{\text{feas}} = \mathcal{X}$, hence all of the decisions are feasible.

The problem is then

$$\begin{cases} \min & J(x) \\ \text{s.t.} & x \in \mathcal{X}\,. \end{cases}$$

Here, we only consider the case $\mathcal{X} \subset \mathbb{R}^n$ is nonempty and open. (Note that $\mathcal{X}$ can be $\mathbb{R}^n$.)

Recall that this problem has a global solution if the following conditions are satisfied:

- J is continuous,
- $\lim_{\|x\|\to\infty} J(x) = \infty$,
- $\lim_{x\to\partial\mathcal{X}} J(x) = \infty$.

Unconstrained optimization

We also have the following.

- Strongly convex functions are continuous and coercive, hence they have unique minimizers.
- The function $J(x) = \|Ax - b\|$ has a minimizer, where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. The result is clear when $\operatorname{im} A = \{0\}$. In case $\operatorname{im} A$ is nontrivial, we have

$$\min_{x\in\mathbb{R}^n} \|Ax - b\| = \min_{y\in\operatorname{im} A} \|y - b\|,$$

  and that $g(y) = \|y - b\|$ is coercive on $\operatorname{im} A$. Then there exists $y^* \in \operatorname{im} A$ that minimizes $y \mapsto \|y - b\|$. Since $y^* \in \operatorname{im} A$, there exists $x^* \in \mathbb{R}^n$ in which $Ax^* = y^*$. Finally, this x^* actually minimizes J.
- The function $J(x) = \|Ax - b\|^2$ has a minimizer since it is the square of a non-negative function from the previous bullet.

Section 2

Local optimality conditions

Subsection 1

One-dimensional case

Optimality conditions for minimization

It is extremely difficult to find optimal points directly from the definition. We therefore need to transform an optimization problem into something we can solve or verify.
These criteria are called **optimality conditions**.

Optimality conditions for minimization

We start with optimality condition in one-dimension that everyone learned in Calculus I.

Theorem

Necessary optimality condition (NOC.) (What a local solution makes?)

Let $J : \mathcal{X} \subset \mathbb{R} \to \mathbb{R}$ be a function whose derivative $J'(x)$ exists at every $x \in \mathcal{X}$. Then

$$\bar{x} \text{ is a local minimizer of } J \implies \underbrace{J'(\bar{x}) = 0}_{\substack{\textit{i.e. } \bar{x} \text{ is a stationary} \\ \text{(or critical) point.}}} .$$

Theorem

Sufficient optimality condition (SOC.) (What makes a local solution?)

Let $J : \mathcal{X} \subset \mathbb{R} \to \mathbb{R}$ be a function whose Hession $J''(x)$ exists at every $x \in \mathcal{X}$ and is continuous. Then

$$J'(\bar{x}) = 0 \text{ and } J''(x) > 0 \implies \bar{x} \text{ is a local minimizer of } J .$$

Optimality conditions for minimization

This has brought us to our famous procedure:

Step 1. **Find candidate points.**
Find all critical points by solving $J'(x) = 0$.

Step 2. **Check whether they are local minimizers.**
For each critical point $\bar{x}$, if $J''(\bar{x}) > 0$ then $\bar{x}$ is a local minimizer of J.

Optimality conditions for minimization and maximization

Since $\arg\max J = \arg\min[-J]$ and $(-J)' = -J'$, therefore we could refine the previous steps to include maximization:

Step 1. **Find candidate points.**
Find all critical points by solving $J'(x) = 0$.

Step 2. **Check whether they are local minimizers or maximizers.**
At each critical point $\bar{x}$,
- if $J''(\bar{x}) > 0$, then $\bar{x}$ is a local minimizer of J ;
- if $J''(\bar{x}) < 0$, then $\bar{x}$ is a local maximizer of J .

Remark

A critical point might neither be a local minimizer or maximizer. Such critical points are called **inflection points**.

Exercises

Exercises

Find all local minimizers and maximizers of the following functions.

- $J(x) = x^3 - 4x$.
- $J(x) = (x^2 - 1)(x^2 - 4)$.
- $J(x) = \sin^2(x)$.

Subsection 2

Multi-dimensional case

Optimality conditions for minimization

The NOC and SOC from the one-dimensional case could be extended to the general setting with $J : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$. Here, we replace J' with ∇J and the sign of J'' is replaced with the positive and negative definiteness of $\nabla^2 J$ at critical points.

Theorem

Necessary optimality condition (NOC.) (What a local solution makes?)

Let $J : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$ be a function whose derivative $\nabla J(x)$ exists at every $x \in \mathcal{X}$. Then

$$\bar{x} \text{ is a local minimizer of } J \implies \underbrace{\nabla J(\bar{x}) = 0}_{\substack{\textit{i.e. } \bar{x} \text{ is a stationary} \\ \text{(or critical) point.}}} .$$

Theorem

Sufficient optimality condition (SOC.) (What makes a local solution?)

Let $J : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$ be a function whose Hession $\nabla^2 J(x)$ exists at every $x \in \mathcal{X}$ and is continuous. Then

$$\nabla J(\bar{x}) = 0 \text{ and } \nabla^2 J(x) \text{ is PD} \implies \bar{x} \text{ is a local minimizer of } J .$$

Optimality conditions for minimization and maximization

Applying a similar technique as in the one-dimensional case, we obtain a procedure that works with both local minimization and maximization.

Step 1. **Find candidate points.**

Find all critical points by solving $\nabla J(x) = 0$.

Step 2. **Check whether they are local minimizers or maximizers.**

At each critical point $\bar{x}$,

- if $\nabla^2 J(\bar{x})$ is PD, then $\bar{x}$ is a local minimizer of J ;
- if $\nabla^2 J(\bar{x})$ is ND, then $\bar{x}$ is a local maximizer of J .

Remark

- Critical points at which Hessians are indefinite are called saddle points.
- Other cases are inconclusive.

Exercises

Exercise

Find all local minimizers and maximizers of the following functions.

- $J(x_1, x_2) = x_1^2 + 2x_2^2 - 4x_1 + 8$.
- $J(x, y) = x^2 - y^2$.
- $J(x_1, x_2) = x_1^2 x_2^2$.

Section 3

Global optimality

Global solutions

Since we largely rely on differential geometry (which is a local tool) to detect and identify optimal points, they are limited to the local behavior of an optimization problem.

The following is perhaps the only case where we could guarantee global optimality.

Theorem

Suppose that $J : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$ is convex. Then

$$\bar{x} \text{ is a } \textbf{global} \text{ minimizer of } J \iff \bar{x} \text{ is a } \textbf{local} \text{ minimizer of } J \overset{\substack{J \text{ is} \\ \text{differentiable}}}{\iff} \nabla J(\bar{x}) = 0\,.$$

Exercise

Exercise

Revisit the optimality of $J(x_1, x_2) = x_1^2 x_2^2$.

Unconstrained quadratic program

Find local minimizers of a convex quadratic function $J(x) = \frac{1}{2}x^t Qx + a^t x + a_0$.

Least-squares problems

Least-squares

Consider $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. A way to resolve $Ax = b$ is to solve

$$\left\{ \min_{x \in \mathbb{R}^n} \quad \|Ax - b\|^2 \, . \right.$$

Recall that this is a convex problem. Moreover, even when $Ax = b$ is infeasible, the above optimization problem still gives a solution, which is called a **least-squares solution**.

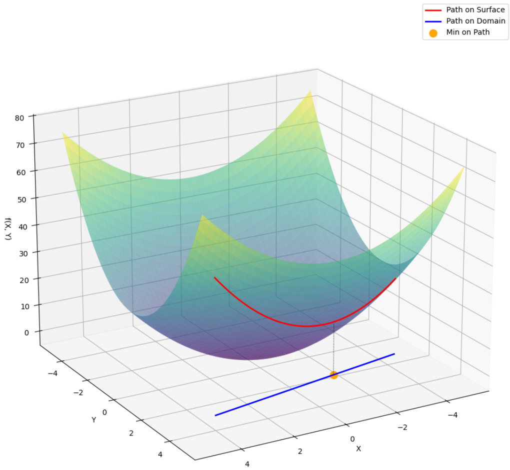
Use the global optimality condition to show that a least-squares solution can be found through a linear system

$$A^t(Ax - b) = 0 \, .$$

Exact line-search

Consider a function $J : \mathbb{R}^n \to \mathbb{R}$. Take any point $x^0 \in \mathbb{R}^n$ and any nonzero direction $d \in \mathbb{R}^n \setminus \{0\}$.

The aim of **exact line-search** is to find a minimizer of J along the line $s \in \mathbb{R} \mapsto c(s) := x^0 + sd$.



We shall show an example of an exact line-search for a quadratic function

$$J(x) = \frac{1}{2} x^t Q x + a^t x + a_0 \,, \quad \text{where } Q \text{ is PD}\,.$$

Exact line-search

To solve the exact line-search problem, we minimize the function $f = J \circ c$, which could be expressed as

$$
\begin{aligned}
f(s) &= \frac{1}{2}(x^0 + sd)^t Q(x^0 + sd) + a^t(x^0 + sd) + a_0 \\
&= \frac{1}{2}(x^0)^t Qx^0 + \frac{1}{2}s^2 d^t Qd + a^t x^0 + sd^t \underbrace{(Qx^0 + a)}_{=\nabla J(x^0)} + a_0 \,.
\end{aligned}
$$

Hence we get

$$
f'(s) = s \underbrace{d^t Qd}_{\substack{\neq 0 \\ \text{since } Q \text{ is PD}}} + d^t \nabla J(x^0)
$$

and thus

$$
f'(s) = 0 \iff s = -\frac{d^t \nabla J(x^0)}{d^t Qd} \,.
$$

Since f is convex, the above critical point is a minimizer of f, which in turns is a solution of the exact line-search problem.

That's it!

Key takeaways.

- Optimality conditions, from one- and multi-dimensional settings.
- Unconstrained quadratic programs and least-squares problems.
- Exact line-search.

Thank you.