Optimization

Lecture 5: The theory of Constrained 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

Constrained optimization

Constrained optimization

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

The problem is then

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

Recall that this problem has a global solution if J is continuous and $\underbrace{\mathcal{X}^{\text{feas}} \text{ is compact}}_{\text{closed and bounded}}$.

Explicit constrained optimization

We mainly focus on the case where the constraint is **explicitly** defined by functions.

$$\mathcal{X}^{\text{feas}} = \left\{ x \in \mathcal{X} \;\middle|\; \begin{array}{ll} g_j(x) \leq 0 & j = 1, \ldots, r \\ h_k(x) = 0 & k = 1, \ldots, \ell \end{array} \right\},$$

where $J, g_j, h_k : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$.

We assume that $\mathcal{X} \subset \mathbb{R}^n$ is a common open domain where all functions are defined.

($\mathcal{X}$ can also be $\mathbb{R}^n$.)

Then we shall study the following form of **explicit constrained optimization problem**

$$\begin{cases} \min & J(x) \\ \text{s.t.} & g_j(x) \leq 0 \quad j = 1, \ldots, r \\ & h_k(x) = 0 \quad k = 1, \ldots, \ell\,. \end{cases}$$

Section 2

Optimality conditions

Subsection 1

Multipliers and the KKT conditions

Multipliers

We assign to each constraint a multiplier. Hence, we have equal numbers of multipliers to the number of constraints.

Equality constraints.
The constraint $h_k(x) = 0$ is assigned with a multiplier $\lambda_k \in \mathbb{R}$.
We write $\lambda = (\lambda_1, \ldots, \lambda_\ell)$.

Inequality constraints.
The constraint $g_j(x) \leq 0$ is assigned with a multiplier $\mu_j \geq 0$.
We write $\mu = (\mu_1, \ldots, \mu_r)$.

We usually indicate after each constraint its associated multiplier, as presented below.

$$\begin{cases} \min \quad J(x) & \text{(1a)} \\ \text{s.t.} \quad g_j(x) \leq 0 \quad j = 1, \ldots, r \qquad (\mu_j) & \text{(1b)} \\ \qquad\; h_k(x) = 0 \quad k = 1, \ldots, \ell \qquad (\lambda_k)\,. & \text{(1c)} \end{cases}$$

Lagrangian function

We then define the **Lagrangian function** of this explicit optimization problem as $L : \mathcal{X} \times \mathbb{R}^r \times \mathbb{R}^\ell \to \overline{\mathbb{R}}$, given by

$$L(x, \mu, \lambda) = J(x) + \sum_{j=1}^{r} \mu_j g_j(x) + \sum_{k=1}^{\ell} \mu_k g_k(x)\,.$$

Note that we have the following gradient in x:

$$\nabla_x L(x, \mu, \lambda) = \nabla J(x) + \sum_{j=1}^{r} \mu_j \nabla g_j(x) + \sum_{k=1}^{\ell} \lambda_k \nabla h_k(x)\,.$$

KKT Conditions

Attached to the optimization problem (1) are its **KKT conditions***(also called its **KKT system**) given as follows:

$$
\begin{aligned}
&\nabla_x L(x, \mu, \lambda) = 0 & & \text{(Stationary)} & (2a)\\
&g_j(x) \leq 0 & \forall j = 1, \ldots, r & \text{(Primal feasibility I)} & (2b)\\
&h_k(x) = 0 & \forall k = 1, \ldots, \ell & \text{(Primal feasibility II)} & (2c)\\
&\mu_j g_j(x) = 0 & \forall j = 1, \ldots, r & \text{(Complementarity)} & (2d)\\
&\mu_j \geq 0 & \forall j = 1, \ldots, r & \text{(Dual feasibility)} & (2e)
\end{aligned}
$$

A triple (x, μ, λ) that validates all (2a)–(2e) is said to **satisfy the KKT conditions**. We also say that (x, μ, λ) is a **KKT stationary point**.

*The KKT conditions are named after the three mathematicians William Karush, Harold W. Kuhn, and Albert W. Tucker.

Example

Example 1

Consider the problem

$$\begin{cases} \min & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 = x_2 \qquad\qquad (\lambda) \\ & x_1 + x_2 \geq 1 \qquad (\mu)\,. \end{cases}$$

The KKT conditions are then derived as

$$\begin{aligned} 2x_1 + \lambda - \mu &= 0 \\ 2x_2 - \lambda - \mu &= 0 \\ x_1 &= x_2 \\ x_1 + x_2 &\geq 1 \\ \mu(1 - x_1 - x_2) &= 0 \\ \mu &\geq 0\,. \end{aligned}$$

Subsection 2

Problems with equality or inequality constriants only

Problems with equality constriants only

When there are only equality constraints, the problem reads

$$\begin{cases} \min & J(x) \\ \text{s.t.} & h_k(x) = 0 \quad k = 1, \ldots, \ell \qquad (\lambda_k)\,. \end{cases}$$

The Lagrangian function and its gradient reduce to

$$L(x, \lambda) = J(x) + \sum_{k=1}^{\ell} \lambda_k h_k(x) \quad \text{and} \quad \nabla_x L(x, \lambda) = \nabla J(x) + \sum_{k=1}^{\ell} \lambda_k \nabla h_k(x)\,.$$

Finally, the KKT conditions are simplified into

$$\begin{aligned} &\nabla_x L(x, \lambda) = 0 && && \text{(Stationary)} \\ &h_k(x) = 0 && \forall k = 1, \ldots, \ell && \text{(Primal feasibility)} \end{aligned}$$

Example

Example 2

Consider an optimization problem

$$\begin{cases} \min & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 = 1 \qquad (\lambda)\,. \end{cases}$$

Then the KKT conditions of this optimization problem are

$$\begin{aligned} 2x_1 + \lambda &= 0 \\ 2x_2 + \lambda &= 0 \\ x_1 + x_2 &= 1\,. \end{aligned}$$

Solving this system of equations, we obtain a unique KKT stationary point

$$(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)\,.$$

Problems with inequality constriants only

When there are only inequality constraints, the problem reads

$$\begin{cases} \min & J(x) \\ \text{s.t.} & g_j(x) \leq 0 \quad j = 1, \ldots, r \qquad (\mu_j)\,. \end{cases}$$

The Lagrangian function and its gradient reduce to

$$L(x, \mu) = J(x) + \sum_{j=1}^{r} \mu_j g_j(x) \quad \text{and} \quad \nabla_x L(x, \mu) = \nabla J(x) + \sum_{j=1}^{r} \mu_j \nabla g_j(x)\,.$$

Finally, the KKT conditions are simplified into

$$\begin{aligned} &\nabla_x L(x, \mu) = 0 && && \text{(Stationary)} \\ &g_j(x) \leq 0 && \forall j = 1, \ldots, r && \text{(Primal feasibility)} \\ &\mu_j g_j(x) = 0 && \forall j = 1, \ldots, r && \text{(Complementarity)} \\ &\mu_j \geq 0 && \forall j = 1, \ldots, r && \text{(Dual feasibility)} \end{aligned}$$

Example

Example 3

Consider an optimization problem

$$\begin{cases} \min & x(1-x) \\ \text{s.t.} & x \geq 0 \qquad (\mu)\,. \end{cases}$$

Then the associated KKT conditions are given by

$$\begin{aligned} 1 - 2x - \mu &= 0 \\ \mu x &= 0 \\ x &\geq 0 \\ \mu &\geq 0\,. \end{aligned}$$

There are two KKT stationary points, which are

$$(\bar{x}, \bar{\mu}) = \left(\frac{1}{2}, 0\right) \quad \text{and} \quad (\hat{x}, \hat{\mu}) = (0, 1)\,.$$

Subsection 3

Optimality conditions

Optimality conditions

We have not linked the KKT stationary points with solutions of an optimization problem.

The goal is to obtain a set of assumptions, usually known as **constraint qualifications** (**CQ**), under which we can write NOC and SOC of the problem (1) in terms of the KKT conditions.

We may have the following *tentative* statement.

NOC: CQ holds and $\bar{x}$ is a local minimizer of (1)

$$\implies \text{there exist } \bar{\mu} \text{ and } \bar{\lambda} \text{ such that } (\bar{x}, \bar{\mu}, \bar{\lambda}) \text{ is KKT stationary}.$$

and also

SOC: (1) is convex and $(\bar{x}, \bar{\mu}, \bar{\lambda})$ is KKT stationary

$$\implies \bar{x} \text{ is a global minimizer of (1)}$$

Section 3

Constraint qualifications (CQ)

Subsection 1

Linear independence CQ

LICQ

We start with LICQ, which is one of the most common CQs.

Definition 4

- An inequality constraint $g_j(x) \leq 0$ is **active** at $\bar{x}$ if $g_j(\bar{x}) = 0$.
- From the system of inequality constraints (1b), the **active index set** at a feasible point $\bar{x}$ is defined by

$$\mathcal{A}(\bar{x}) = \left\{ j \in \{1, \ldots, r\} \mid g_j(\bar{x}) = 0 \right\}.$$

Definition 5

We say that the constraints (1b)–(1c) satisfy the **linear independence constraint qualification** (or **LICQ**) at a feasible point $\bar{x}$ if the set

$$\left\{ \nabla g_j(\bar{x}) \mid j \in \mathcal{A}(\bar{x}) \right\} \cup \left\{ \nabla h_k(\bar{x}) \mid k = 1, \ldots, \ell \right\}$$

is linearly independent.

Example

Example 6

Consider a problem

$$\begin{cases} \min & 1 - x^2 \\ \text{s.t.} & x \geq -1 \\ & x \leq 1 \end{cases}$$

We may write the constraint system as

$$g_1(x) = -1 - x \quad \text{and} \quad g_2(x) = x - 1\,.$$

Then we have the following active index sets

$$\mathcal{A}(-1) = \{1\}, \quad \mathcal{A}(1) = \{2\}, \quad \text{and} \quad \mathcal{A}(x) = \emptyset \ \text{ for } -1 < x < 1.$$

NOC under LICQ

Theorem 7

Suppose that (1b)–(1c) *satisfy the LICQ at a feasible point* $\bar{x}$. *Then*

$\bar{x}$ *is a local minimizer of* (1)

$\implies$ *there exist multipliers* $\bar{\mu} \in \mathbb{R}^r$ *and* $\bar{\lambda} \in \mathbb{R}^\ell$ *such that* $(\bar{x}, \bar{\mu}, \bar{\lambda})$ *is a KKT stationary point.*

Failure of NOC without a CQ

Even before we formally give some CQs, we would like to give an example to convince that the KKT conditions are, in general, not necessary for a local minimizer.

Example 8

Consider the following problem

$$\begin{cases} \min & x_1 + x_2 \\ \text{s.t.} & \left(x_1^2 + x_2^2\right)^2 = 1 \qquad (\lambda)\,. \end{cases}$$

Then the point $\bar{x} = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ is a minimizer, but $(\bar{x}, \lambda)$ is not KKT stationary for any λ.

The reason behind this failure is the lack of a CQ.

Equivalent formulation with a CQ

Next, we would like to emphasize the fact that the satisfaction of a CQ depends on the functional representation of constraints, not the constraint set itself.

Example 9

Consider the problem

$$\begin{cases} \min & x_1 + x_2 \\ \text{s.t.} & x_1^2 + x_2^2 = 1 \qquad (\lambda)\,. \end{cases}$$

One should observe that the constraint set of this problem and the one from Example 8 are identical. However, in this present setting, we have two KKT stationary points

$$(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \quad \text{and} \quad (\hat{x}_1, \hat{x}_2, \hat{\lambda}) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

Note here that LICQ holds at all feasible points for this formulation.

Subsection 2

Slater CQ

Slater CQ

The LICQ could often be difficult to evaluate. Here, we present another CQ that works with convex constraints. Consider a problem (1) with convex inequality constraints and linear equality constraints as follows:

$$
\begin{cases}
\min \quad J(x) & & (4a) \\
\text{s.t.} \quad g_j(x) \le 0 & j = 1, \dots, r & (4b) \\
\qquad c_k^t x = d_k & k = 1, \dots, \ell \,, & (4c)
\end{cases}
$$

where g_j's are convex.

Definition 10

The constraints (4b)–(4c) satisfy the **Slater CQ** if there is a feasible point $\hat{x}$ in which no inequality constraints are active, *i.e.,* $g_j(\hat{x}) < 0$ for all $j = 1, \dots, r$.

NOC under Slater CQ

Theorem 11

If the constraints (4b)–(4c) *satisfy the Slater CQ, then*

$\bar{x}$ *is a local minimizer of* (4)

$\implies$ *there exist multipliers* $\bar{\mu} \in \mathbb{R}^r$ *and* $\bar{\lambda} \in \mathbb{R}^\ell$ *such that* $(\bar{x}, \bar{\mu}, \bar{\lambda})$ *is a KKT stationary point.*

Example

Example 12

Consider again a problem from Example 1:

$$\begin{cases} \min & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 = x_2 & (\lambda) \\ & x_1 + x_2 \geq 1 & (\mu)\,. \end{cases}$$

Both constraints are linear and there is a feasible point $\hat{x} = (1, 1)$ in which the inequality constraint $\hat{x}_1 + \hat{x}_2 = 2 > 1$ is inactive. Therefore, the Slater CQ holds for this system of constraints.

Subsection 3

Linear constraints as a CQ

Linear constraints as a CQ

Next, we consider the problem (1) where all constraints are linear:

$$\begin{cases} \min & J(x) & & (5a) \\ \text{s.t.} & a_j^t x \leq b_j & j = 1, \ldots, r & (5b) \\ & c_k^t x = d_k & j = 1, \ldots, \ell\,. & (5c) \end{cases}$$

The constriants being linear is qualified as a CQ itself.

Theorem 13

$\bar{x}$ is a local minimizer of (5)

$\implies$ there exist multipliers $\bar{\mu} \in \mathbb{R}^r$ and $\bar{\lambda} \in \mathbb{R}^\ell$ such that $(\bar{x}, \bar{\mu}, \bar{\lambda})$ is a KKT stationary point.

Subsection 4

Optimality conditions

Optimality conditions

Let us wrap up the results of this section into two theorems.

Theorem 14 (NOC)

Suppose that $\bar{x}$ is a local minimizer of (1) *and one of the following is true:*

(a) *LICQ holds at $\bar{x}$,*

(b) *all the inequality constraints are convex, all equality constraints are linear, and Slater CQ holds,*

(c) *the constraints are all linear.*

Then there exist $\bar{\mu} \in \mathbb{R}^r$ and $\bar{\lambda} \in \mathbb{R}^\ell$ such that $(\bar{x}, \bar{\mu}, \bar{\lambda})$ is KKT stationary.

Theorem 15 (SOC)

Suppose that (1) *is a convex program and $(\bar{x}, \bar{\mu}, \bar{\lambda})$ is KKT stationary. Then $\bar{x}$ is a global minimizer of* (1).

Example

Example 16

Consider again the problem (3) from Example 1. Recall that it is a convex program whose KKT conditions are given by

$$2x_1 + \lambda - \mu = 0 \tag{6}$$

$$2x_2 - \lambda - \mu = 0 \tag{7}$$

$$x_1 = x_2 \tag{8}$$

$$x_1 + x_2 \geq 1 \tag{9}$$

$$\mu(1 - x_1 - x_2) = 0 \tag{10}$$

$$\mu \geq 0\,. \tag{11}$$

According to Theorems 14 and 15, a point $\bar{x}$ is a global minimizer of (3) if and only if there are multipliers μ and λ such that $(\bar{x}, \bar{\mu}, \bar{\lambda})$ solves the KKT conditions (6)–(11).
Solving this system of equations and inequalities, we obtain that $(\bar{x}, \bar{\mu}, \bar{\lambda}) = \big((\frac{1}{2}, \frac{1}{2}), 1, 0\big)$ is the unique KKT stationary point. Discarding the multipliers, we get that $\bar{x} = \big(\frac{1}{2}, \frac{1}{2}\big)$ is the unique minimizer of the problem (3).

That's it!

Key takeaways.

- The KKT conditions.
- The choice of CQs under different settings and their roles in NOC and SOC.

Thank you.