

香 港 大 學
THE UNIVERSITY OF HONG KONG

Optimality Conditions for Constrained Programs

Prof. Li Jiayang

DASE7113 Optimization Methods for Intelligent Systems

Optimality Condition

Unconstrained Convex Programs

Given a differentiable convex function $f: \mathbb{R}^n \to \mathbb{R}$, x^* is the solution to the following optimization problem if and only if $\nabla f(x^*) = 0$.

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &x \in \mathbb{R}^n \end{aligned}$$

Optimality Condition

Constrained Convex Programs

Now let's proceed to constrained convex programs of the following form, where $f(x)$ and all $g_i(x)$ are differentiable convex functions.

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_1(x) \leq 0, \\ & \vdots \\ & g_m(x) \leq 0. \end{aligned}$$

⬅ **KKT Conditions**

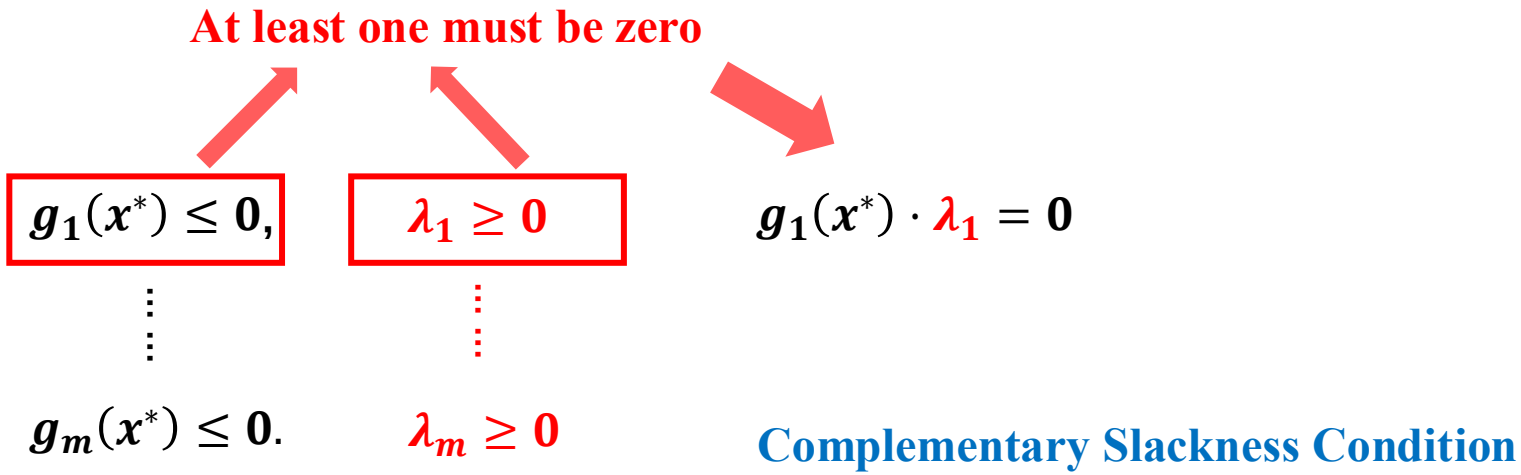
First appeared in a paper by Kuhn and Tucker in 1951

Later people found out Karush had the conditions in his unpublished master's thesis of 1939

Optimality Condition

Constrained Convex Programs

Now let's proceed to constrained convex programs of the following form, where $f(x)$ and all $g_i(x)$ are differentiable convex functions.



Optimality Condition

Constrained Convex Programs

Now let's proceed to constrained convex programs of the following form, where $f(x)$ and all $g_i(x)$ are differentiable convex functions.

KKT conditions

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \cdot \nabla g_i(x^*) = 0 \\ g_1(x^*) \le 0, \quad \lambda_1 \ge 0 \quad g_1(x^*) \cdot \lambda_1 = 0 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \\ g_m(x^*) \le 0. \quad \lambda_m \ge 0 \quad g_m(x^*) \cdot \lambda_m = 0 \end{cases}$$

Optimality Condition

Constrained Convex Programs

Now let's proceed to constrained convex programs of the following form, where $f(x)$ and all $g_i(x)$ are differentiable convex functions.

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \ \ i = 1, \ldots, m \end{aligned}$$

When some (mild) conditions are satisfied,

x^* is the optimal solution if and only if

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \cdot \nabla g_i(x^*) = 0 \\ g_i(x^*) \leq 0, \lambda_i \geq 0, g_i(x^*) \cdot \lambda_i = 0, \ \ i = 1, \ldots, m \end{cases}$$

Optimality Condition

Comparison

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$

$$\nabla f(x^*) = 0$$

What if $g_i(x^*) < 0$ for all i?

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad g_i(x) \le 0, \ \ i = 1, \dots, m$$

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \cdot \nabla g_i(x^*) = 0 \\ g_i(x^*) \le 0, \ \lambda_i \ge 0, \ g_i(x^*) \cdot \lambda_i = 0, \ \ i = 1, \dots, m \end{cases}$$

Optimality Condition

Exercise

Write down the KKT conditions corresponding to the following program

$$\begin{aligned} &\text{minimize} \quad x^2 \\ &\text{subject to} \quad x \geq -1 \end{aligned}$$

$$\begin{aligned} &\text{minimize} \quad x^2 \\ &\text{subject to} \quad -x - 1 \leq 0 \end{aligned}$$

$$\begin{cases} 2x^* + \lambda \cdot (-1) = 0 \\ -x^* - 1 \leq 0, \lambda \geq 0, (-x^* - 1) \cdot \lambda = 0 \end{cases}$$

Optimality Condition

Exercise

Write down the KKT conditions corresponding to the following program

$$\begin{aligned} &\text{minimize} \quad x^2 \\ &\text{subject to} \quad x \geq 1 \end{aligned}$$

$$\begin{aligned} &\text{minimize} \quad x^2 \\ &\text{subject to} \quad -x + 1 \leq 0 \end{aligned}$$

$$\begin{cases} 2x^* + \lambda \cdot (-1) = 0 \\ -x^* + 1 \leq 0, \lambda \geq 0, (-x^* + 1) \cdot \lambda = 0 \end{cases}$$

Optimality Condition

Constrained Convex Programs

Write down the KKT conditions corresponding to the following program

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1 + x_2 \geq 1 \\ & x_1 \geq 0 \end{aligned}$$

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & -x_1 - x_2 + 1 \leq 0 \quad (\lambda_1) \\ & -x_1 \leq 0 \quad (\lambda_2) \end{aligned}$$

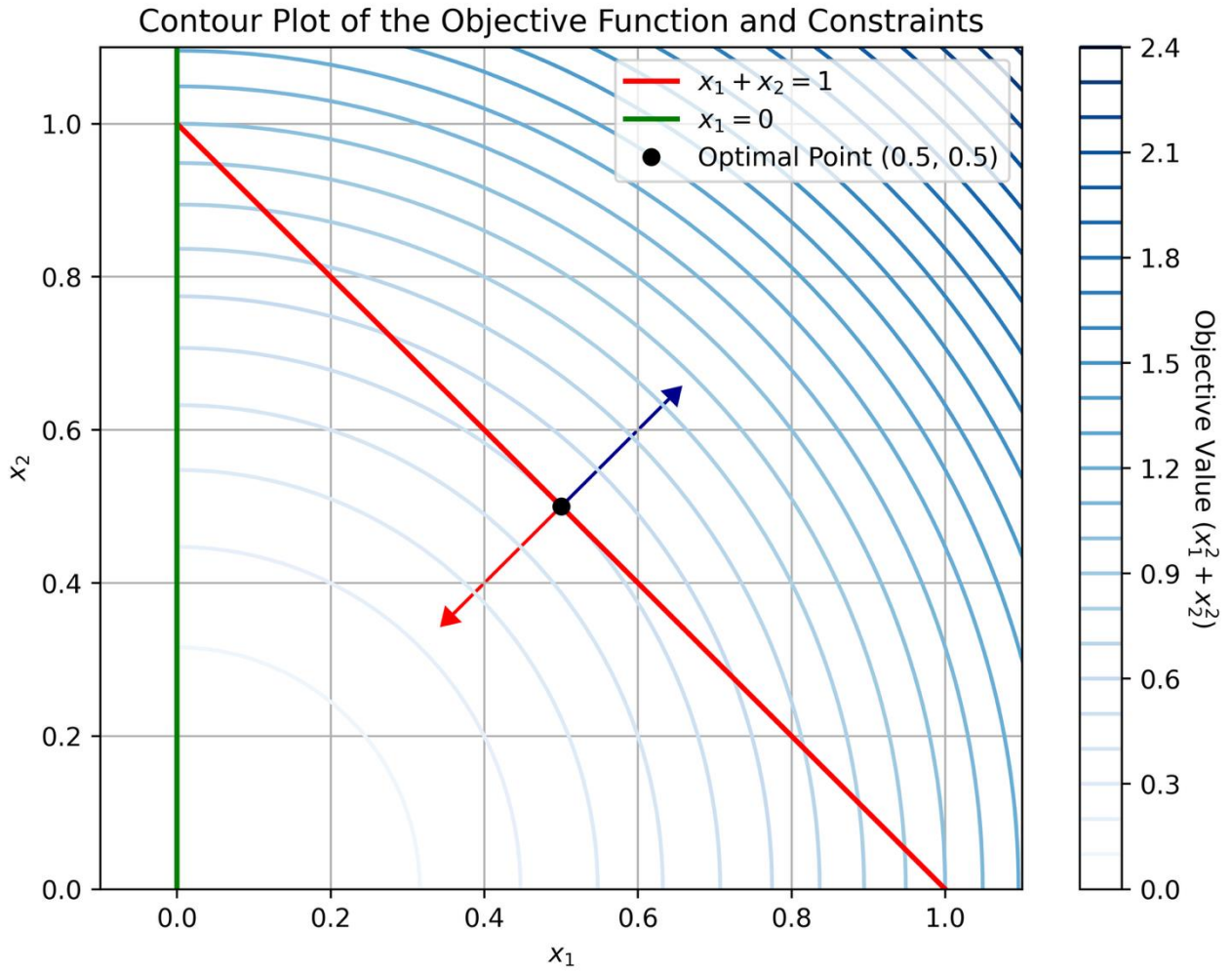
$$\begin{cases} \begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} + \lambda_1 \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} = 0 \\ -x_1^* - x_2^* - 1 \leq 0,\ \lambda_1 \geq 0,\ (-x_1^* - x_2^* + 1) \cdot \lambda_1 = 0 \\ -x_1^* \leq 0,\ \lambda_2 \geq 0,\ (-x_1^*) \cdot \lambda_2 = 0 \end{cases}$$

Optimality Condition

$$\begin{aligned} \textbf{minimize} \quad & x_1^2 + x_2^2 \\ \textbf{subject to} \quad & -x_1 - x_2 + 1 \leq 0 \\ & -x_1 \leq 0 \end{aligned}$$

Optimal solution: $[0.5, 0.5]^T$

$$\begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} + \lambda_1 \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} = 0$$



Optimality Condition

Constrained Convex Programs

Write down the KKT conditions corresponding to the following program

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1 + x_2 \geq 1 \\ & x_1 \geq 0.8 \end{aligned}$$

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & -x_1 - x_2 + 1 \leq 0 \quad \lambda_1 \\ & -x_1 + 0.8 \leq 0 \quad \lambda_2 \end{aligned}$$

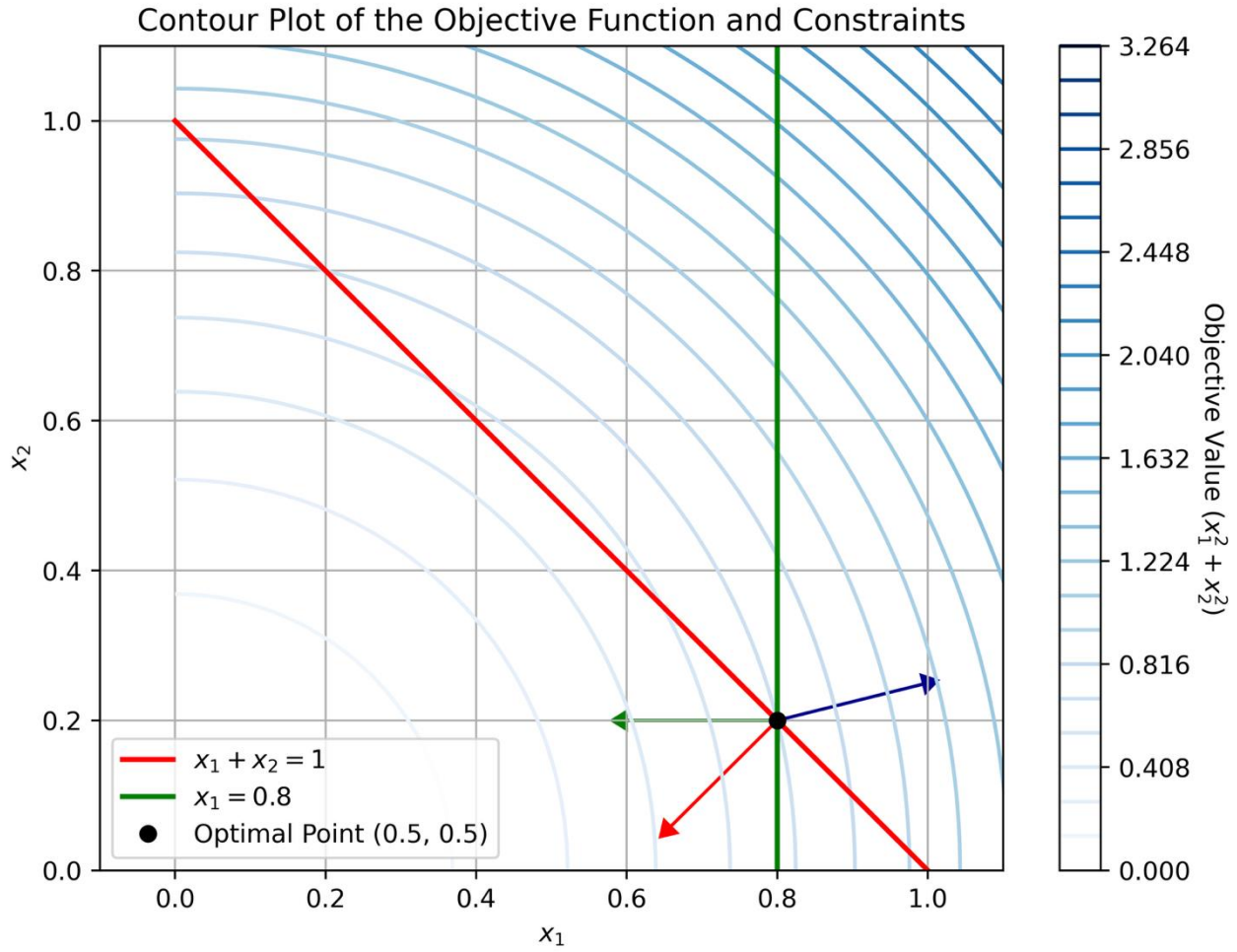
$$\begin{cases} \begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} + \lambda_1 \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} = 0 \\ -x_1^* - x_2^* + 1 \leq 0,\ \lambda_1 \geq 0,\ (-x_1^* - x_2^* + 1) \cdot \lambda_1 = 0 \\ -x_1^* + 0.8 \leq 0,\ \lambda_2 \geq 0,\ (-x_1^* + 0.8) \cdot \lambda_2 = 0 \end{cases}$$

Optimality Condition

$$\begin{aligned} \textbf{minimize} \quad & x_1^2 + x_2^2 \\ \textbf{subject to} \quad & -x_1 - x_2 + 1 \leq 0 \\ & -x_1 + 0.8 \leq 0 \end{aligned}$$

Optimal solution: $[0.8, 0.2]^T$

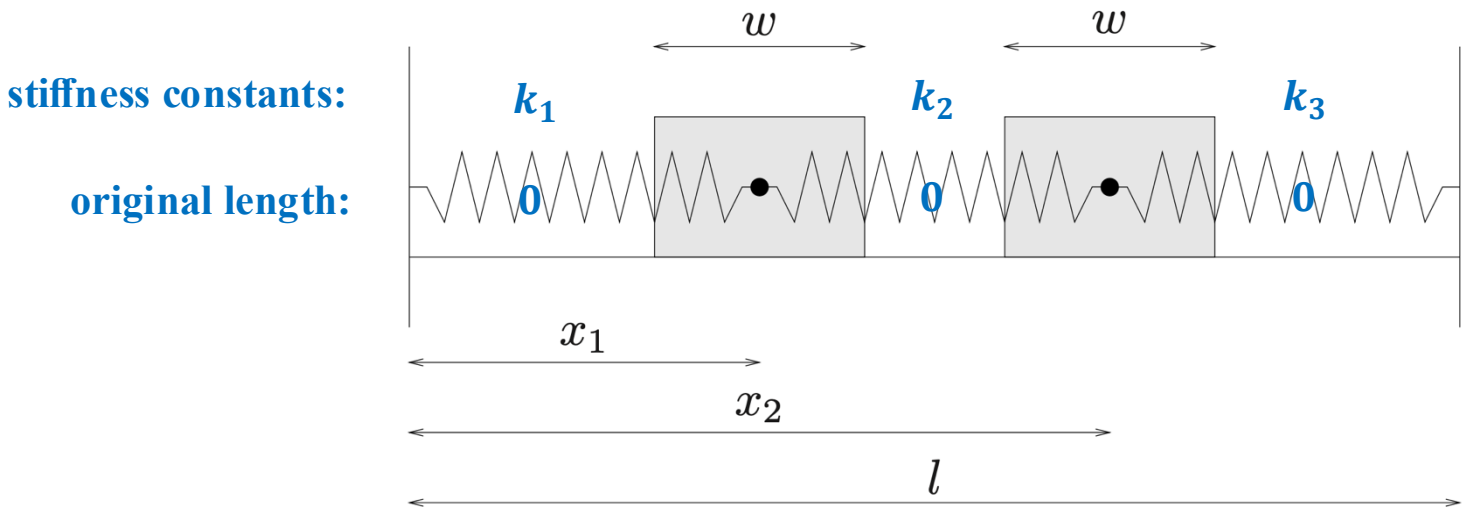
$$\begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} + \lambda_1 \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} = 0$$



Optimality Condition

Exercise

Solve the equilibrium positions of the two blocks in the following system.



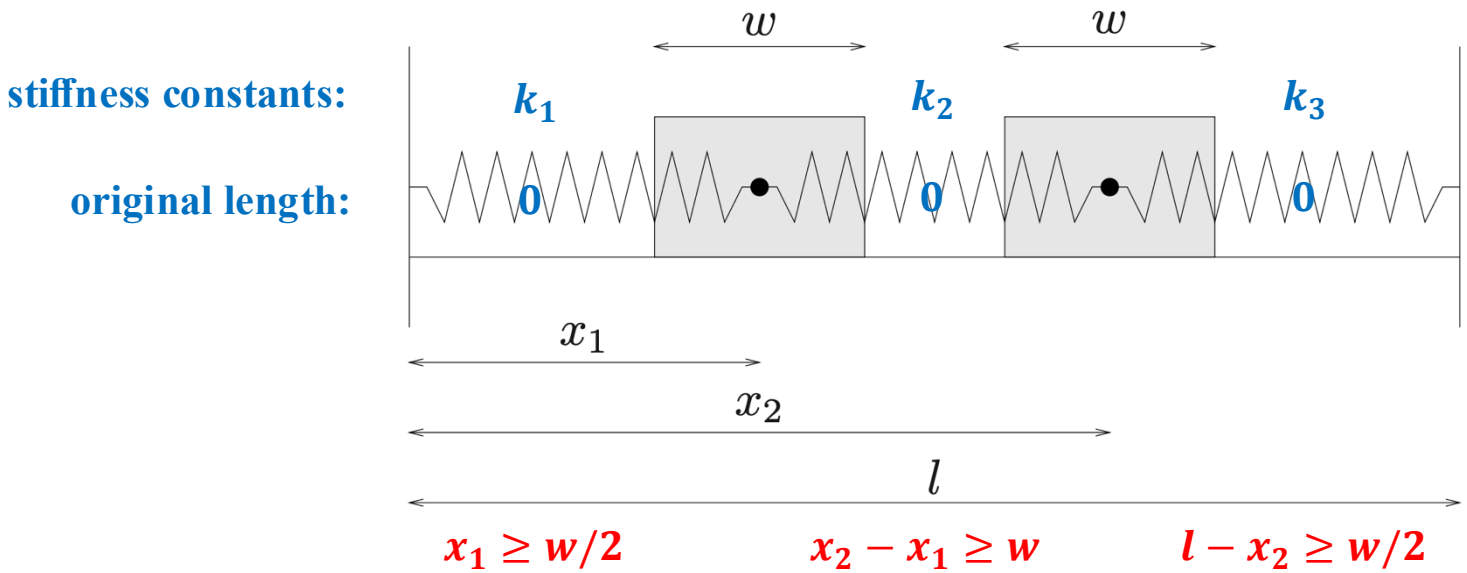
High-school method: Force analysis (受力分析)

Advanced method: The Principle of Minimum Total Potential Energy

Optimality Condition

Exercise

Solve the equilibrium positions of the two blocks in the following system.



Total Potential Energy: $\frac{k_1}{2} \cdot x_1^2 + \frac{k_2}{2} \cdot (x_2 - x_1)^2 + \frac{k_3}{2} \cdot (l - x_2)^2$

Optimality Condition

Exercise

Solve the equilibrium positions of the two blocks in the following system.

$$\text{minimize} \quad \frac{k_1}{2} \cdot x_1^2 + \frac{k_2}{2} \cdot (x_2 - x_1)^2 + \frac{k_3}{2} \cdot (l - x_2)^2$$

$$\text{subject to} \quad -x_1 + w/2 \le 0$$

$$x_1 - x_2 + w \le 0$$

$$x_2 + w/2 - l \le 0$$

$$(-x_1^* + w/2) \cdot \lambda_1 = 0$$

$$(x_1^* - x_2^* + w) \cdot \lambda_2 = 0$$

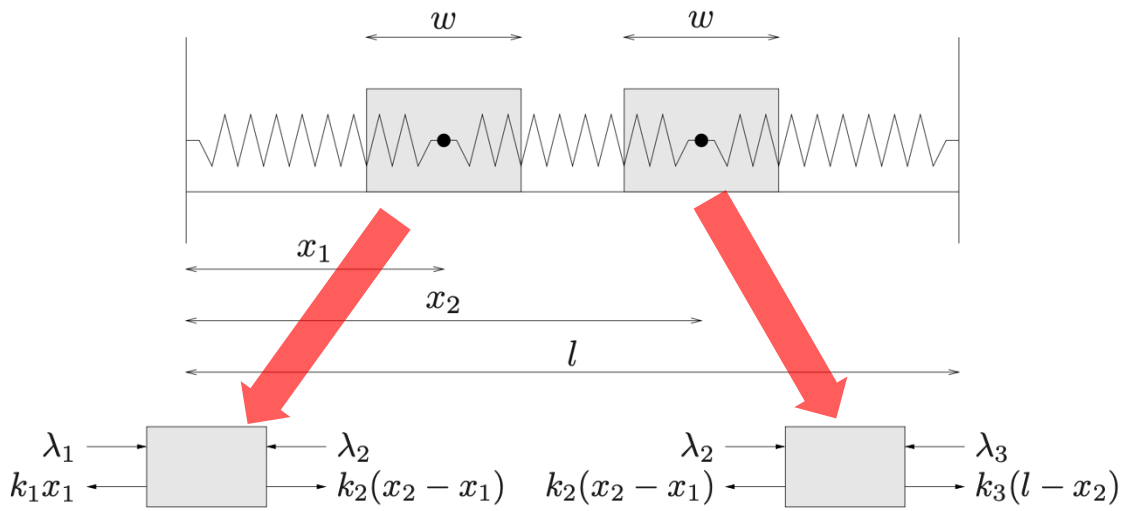
$$(x_2^* + w/2 - l) \cdot \lambda_3 = 0$$

$$\begin{bmatrix} k_1 x_1^* - k_2(x_2^* - x_1^*) \\ k_2(x_2^* - x_1^*) - k_3(l - x_2^*) \end{bmatrix} + \lambda_1 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \lambda_3 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

Optimality Condition

Exercise

Solve the equilibrium positions of the two blocks in the following systems.



$$(-x_1^* + w/2) \cdot \lambda_1 = 0$$

$$(x_1^* - x_2^* + w) \cdot \lambda_2 = 0$$

$$(x_2^* + w/2 - l) \cdot \lambda_3 = 0$$

$$\begin{bmatrix} k_1 x_1^* - k_2(x_2^* - x_1^*) \\ k_2(x_2^* - x_1^*) - k_3(l - x_2^*) \end{bmatrix} + \lambda_1 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \lambda_2 \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \lambda_3 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$\text{minimize} \quad f(x)$$
$$\text{subject to} \quad h_i(x) = 0,\ i = 1, \ldots, p$$

$$\text{minimize} \quad f(x)$$
$$\text{subject to} \quad h_i(x) \le 0,\ i = 1, \ldots, p$$
$$-h_i(x) \le 0,\ i = 1, \ldots, p$$

Try to write down its corresponding KKT condition.

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^+ \cdot \nabla h_i(x^*) - \sum_{i=1}^{p} \lambda_i^- \cdot \nabla h_i(x^*) = 0$$

$$h_i(x^*) \le 0 \qquad \lambda_i^+ \ge 0 \qquad h_i(x^*) \cdot \lambda_i^+ = 0$$

$$-h_i(x^*) \le 0 \qquad \lambda_i^- \ge 0 \qquad -h_i(x^*) \cdot \lambda_i^- = 0 \qquad i = 1, \dots, p$$

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$\nabla f(x^*) + \sum_{i=1}^{p} (\lambda_i^+ - \lambda_i^-) \cdot \nabla h_i(x^*) = 0$$

$$h_i(x^*) \le 0 \qquad \lambda_i^+ \ge 0$$

$$-h_i(x^*) \le 0 \qquad \lambda_i^- \ge 0$$

$$h_i(x^*) \cdot (\lambda_i^+ - \lambda_i^-) = 0$$

$$i = 1, \dots, p$$

Let us define $\lambda = \lambda_i^+ - \lambda_i^-$, then…

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i \cdot \nabla h_i(x^*) = 0$$

$$h_i(x^*) = 0 \qquad i = 1, \ldots, p$$

← **KKT Conditions**

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && h_i(x) = 0, i = 1, \dots, p \\ & && x \geq 0 \qquad x \in \mathbb{R}^n \end{aligned}$$

Try to write down its corresponding KKT conditions.

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i \cdot \nabla h_i(x^*) - \sum_{j=1}^{n} \mu_j \cdot e_j = 0$$

$$h_i(x^*) = 0 \qquad i = 1, \dots, p$$

$$-x_j^* \le 0 \qquad \mu_j \ge 0 \qquad \mu_j \cdot x_j^* = 0 \qquad j = 1, \dots, n$$

$$e_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow j\text{-th element}$$

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}$$

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i \cdot \nabla h_i(x^*) - \sum_{j=1}^{n} \mu_j \cdot e_j = 0$$

$$h_i(x^*) = 0 \qquad i = 1, \dots, p$$

$$-x_j^* \le 0 \qquad \mu_j \ge 0 \qquad \mu_j \cdot x_j^* = 0 \qquad j = 1, \dots, n$$

$$x^* \ge 0 \qquad \mu \ge 0 \qquad \mu^T x^* = 0$$

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

$$h_i(x^*) = 0 \qquad i = 1, \dots, p$$

$$x^* \geq 0 \qquad \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i \cdot \nabla h_i(x^*) \geq 0 \qquad \langle \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i \cdot \nabla h_i(x^*), x^* \rangle = 0$$

KKT Conditions

Optimality Condition

Constrained Convex Programs

Now, consider the following mathematical program,
where $f(x)$ is a differentiable convex function and all $h_i(x)$ are linear functions.

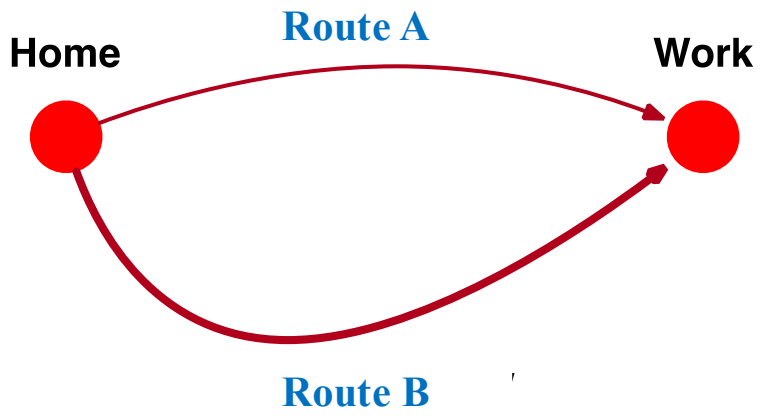
$$h_i(x^*) = 0 \qquad i = 1, \dots, p$$

$$x^* \geq 0 \qquad \nabla f(x^*) - \sum_{i=1}^{p} \lambda_i \cdot \nabla h_i(x^*) \geq 0 \qquad \langle \nabla f(x^*) - \sum_{i=1}^{p} \lambda_i \cdot \nabla h_i(x^*), x^* \rangle = 0$$

This is also correct!

Optimality Condition

Consider a small town with ***d* residents** who drive from "Home" to "Work" every morning.



x_1: The number of travelers choosing Route A

$c_1(x_1)$: The travel time of using Route A

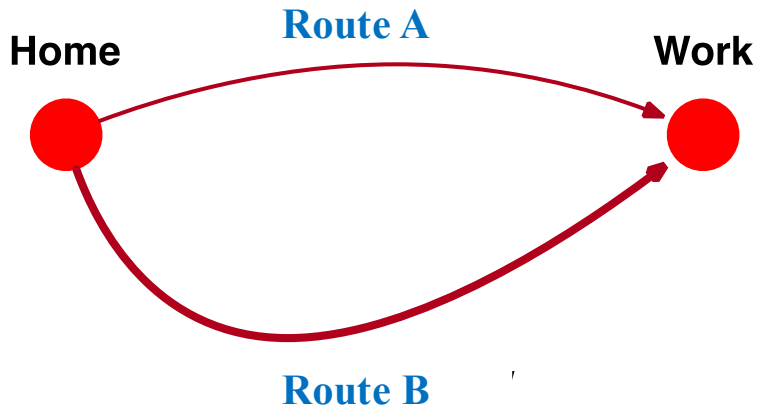
x_2: The number of travelers choosing Route B

$c_2(x_2)$: The travel time of using Route B

How many travelers will select the two routes, respectively?

Optimality Condition

Consider a small town with ***d* residents** who drive from "Home" to "Work" every morning.





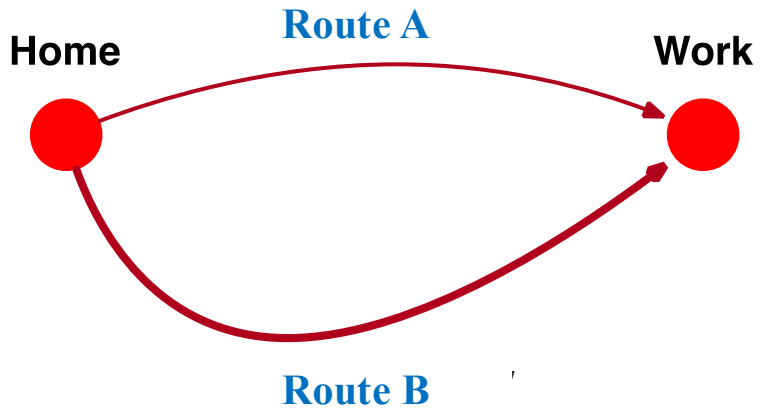
$x_1 = 1$ $c_1 = 2$

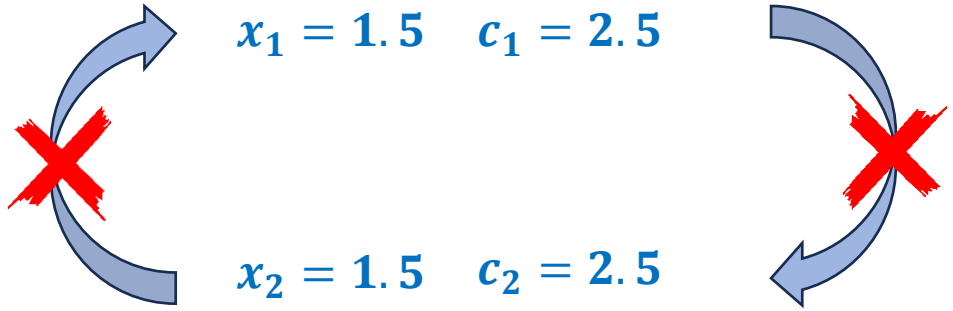
$x_2 = 2$ $c_2 = 3$

Is this a stable outcome?

Optimality Condition

Consider a small town with ***d*** **residents** who drive from "Home" to "Work" every morning.

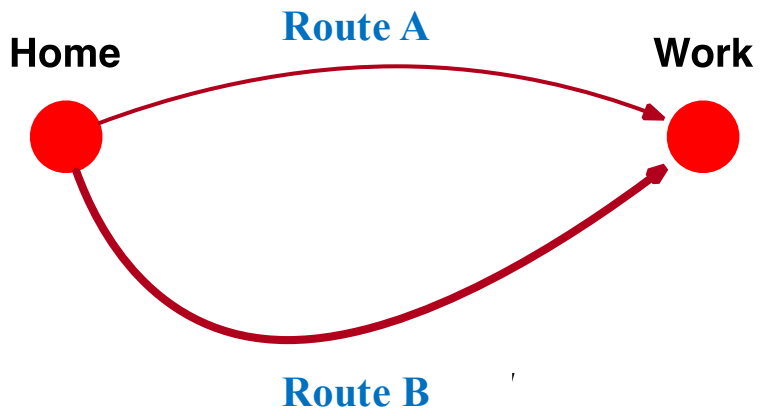


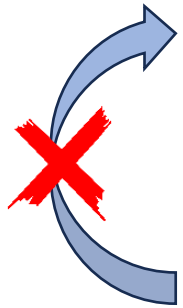


Is this a stable outcome?

Optimality Condition

Consider a small town with ***d* residents** who drive from "Home" to "Work" every morning.





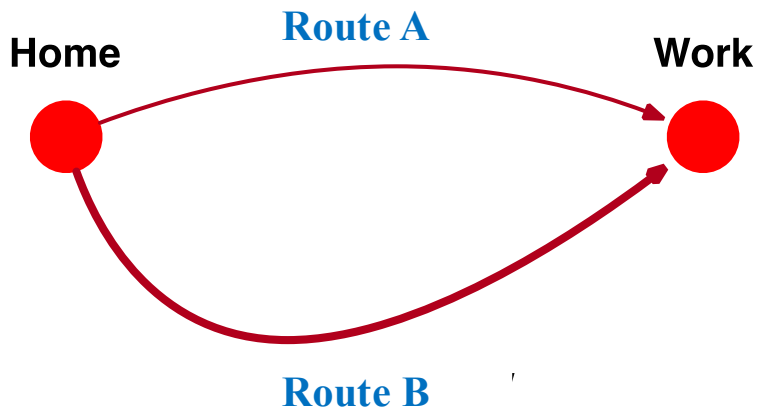
$x_1 = 0$ $c_1 = 3$

$x_2 = 3$ $c_2 = 2$

Is this a stable outcome?

Optimality Condition

Consider a small town with ***d*** **residents** who drive from "Home" to "Work" every morning.



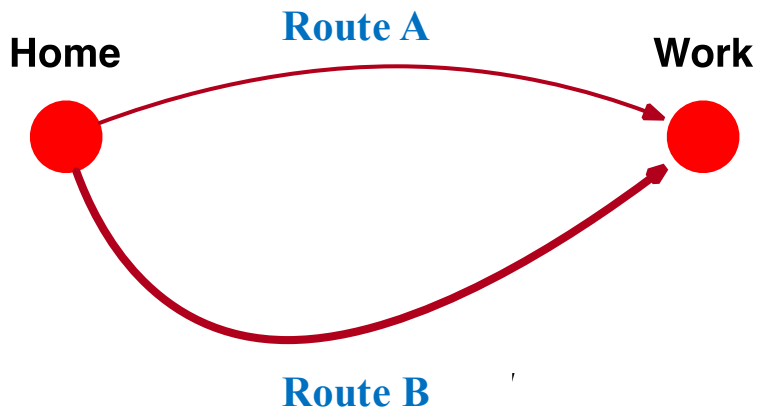
When can the system be stable?

If a route is **not optimal**, then it must **not be used**

If a route is **used**, then it must be **optimal**

Optimality Condition

Consider a small town with d residents who drive from "Home" to "Work" every morning.



Think it as the minimum cost

There exists $\lambda \in \mathbb{R}$ such that

If $c_i(x_i^*) > \lambda$, then we must have $x_i^* = 0$

If $x_i^* > 0$, then we must have $c_i(x_i^*) = \lambda$

$$x_i^* \geq 0 \qquad c_i(x_i^*) - \lambda \geq 0 \qquad x_i^* \cdot (c_i(x_i^*) - \lambda) = 0$$

Optimality Condition

Consider a small town with d residents who drive from "Home" to "Work" every morning.

What if $f(x)$ is a function such that $\nabla f(x) = c(x)$?

$$f(x) = \sum_{i=1,2} \int_0^{x_i} c_i(v)\mathrm{d}v$$

minimize $f(x)$

subject to $x_1 + x_2 = d$

$x \geq 0$

KKT

$$x_1^* + x_2^* = d$$

$$x^* \geq 0 \qquad c(x^*) - \lambda \geq 0 \qquad \langle x^*, c(x^*) - \lambda \rangle = 0$$

$$c(x) = \begin{bmatrix} c_1(x_1) \\ c_2(x_2) \end{bmatrix}$$