Lecture 5: Static Optimization

Guoxuan Ma[1]

UIBE Math Camp, 2022

Introduction

Definition

An optimization problem in $\mathbb{R}^N$ can be defined as follows. Let $f: \mathbb{R}^N \to \mathbb{R}$, and let $D \subset \mathbb{R}^N$. A maximization problem takes the form

$$\max_{x \in X} f(x) \ s.t. \ x \in D$$

where f is called the **objective function**, x is called the **choice variable**, and D is called the **constraint set** or **feasible set**. A point $x \in \mathbb{R}^N$ is said to be **feasible** iff $x \in D$.

The set of **maximizers**, or **maximum points**, of this problem is defined as

$$\arg\max_{x \in X} \{f(x) : x \in D\} := \{x^* \in D : f(x^*) \geq f(x), \ \forall x \in D\}$$

Example of the set of solutions

Example

Let $X = [-1, 1]$. and $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x^2$.
Maximizing f on C has two solutions, $x = -1$ and $x = 1$.
$\arg\max(f(x)|x \in X) = \{-1, 1\}$.

Example

Let $X = \mathbb{R}_+$, and $f : \mathbb{R}_+ \to \mathbb{R}$ be $f(x) = x^2$.
The problem of maximizing f on X has no solution.
$\arg\max(f(x)|x \in X) = \{\emptyset\}$

If the set of maximizers is nonempty, then this problem is said to **have a solution**. In this case, we define the **maximum**, or the **maximum value**, of this problem as $f(x^*)$, where x^* is an arbitrary maximizer, and denote it as $\max_{x \in X}\{f(x) : x \in D\}$.

Existence of a Solution

A sufficient condition for existence of a solution is given by the **Weierstrass Theorem**.

Theorem

If f is continuous and X is closed and bounded (hence compact) and nonempty, then there exist a global maximum and a global minimum.

Example: Consumer's Utility Maximization Problem

$$\max u(x)$$

$$s.t.$$

$$p \cdot x \leq m$$

$$x \geq 0, \; p \in \mathbb{R}^l, \; x \in \mathbb{R}^l$$

The constraint set $X = B(p, m) = \left\{x \in \mathbb{R}^l : p \cdot x \leq m\right\}$ is compact. So if $u(\cdot)$ is continuous, then a solution exists (by Weierstras's Theorem).

Some useful transformation I

Proposition

Let $f: \mathbb{R}^N \to \mathbb{R}$ *and let* $E \subset D \subset \mathbb{R}^N$*. Suppose that* $\forall x \in D, \exists \hat{x} \in E$ *s.t.* $f(\hat{x}) \geq f(x)$*. Consider the following two problems:*

$$\max_{x \in X} f(x) \ s.t. \ x \in D$$

and

$$\max_{x \in X} f(x) \ s.t. \ x \in E$$

The maximizers in the two problems have the following relation

$$\arg\max_{x \in X} \{f(x) : x \in E\} = \left(\arg\max_{x \in X} \{f(x) : x \in D\}\right) \cap E$$

and if one of the two problems has a solution, then the other also has a solution. Furthermore, when the two problems have a solution, they have the same maximum.

Some useful transformation II

Proposition

Let $f: \mathbb{R}^N \to \mathbb{R}$. *Let* $D \subset \mathbb{R}^N$, *and* x_0 *be some arbitrary element of* D, *and define* $E := \{x \in D : f(x) \geq f(x_0)\}$. *Then we have*

$$\arg\max_{x \in X} \{f(x) : x \in E\} = \arg\max_{x \in X} \{f(x) : x \in D\}$$

and the two problems have the same maximum if they have a solution.

Summary

- Variant 1 says that when we choose $x \in D$ to maximize $f(x)$, we can instead focus only on $E \subset D$ (if we can find an alternative $\hat{x} \in E$ that is weakly better than $x \in D$).
- Variant 2 says that if the alternative x_0 is feasible, then we can ignore all alternatives strictly worse than x_0 without loss of optimality.

Illustrative Example

Consider the following maximization problem:

$$\max_{(x_1,x_2)\in\mathbb{R}^2_{++}} \ln x_1 + \ln x_2 \ s.t. \ x_1 + x_2 = 3$$

Notice that the constraint set D in this problem is

$$D := \{(x_1, x_2) \in \mathbb{R}^2_{++}\} : x_1 + x_2 = 3\}$$

which is not compact under the Euclidean distance d_2 (not closed in $(\mathbb{R}^2, d_2)$).
Using Variant 1, since $(1,2) \in D$, and $f(1,2) = \ln 2$, we can define

$$E := \{(x_1, x_2) \in \mathbb{R}^2_{++} : x_1 + x_2 = 3, \ \ln x_1 + \ln x_2 \geq \ln 2\}$$

So $\max_{(x_1,x_2)\in\mathbb{R}^2_{++}} \ln x_1 + \ln x_2 \ s.t. \ (x_1, x_2) \in E$ has the same set of maximizers as the original problem.

Optimization: Necessary Conditions

The theorem provides the necessary first order condition and the necessary second order condition for an interior maximizer.

Theorem

Let X be a set in $\mathbb{R}$, and $D \subset X$. Let $f : X \to \mathbb{R}$, and consider the problem

$$\max_{x \in X} f(x) \ s.t. \ x \in D$$

and let $x^* \in int(D)$ be a maximizer of the problem.

1. *If f is differentiable at x^*, then $f'(x^*) = 0$.*
2. *If f is differentiable in an open ball around x^*, and is twice differentiable at x^*, then $f''(x^*) \leq 0$.*

Multivariate Functions

Theorem

Let X be a set in $\mathbb{R}^n$, and $D \subset X$. Let $f : X \to \mathbb{R}$, and consider the problem

$$\max_{x \in X} f(x) \ s.t. \ x \in D$$

and let $x^* \in int(D)$ be a maximizer of the problem.

1. *If f is differentiable at x^*, then $\nabla f(x^*) = 0$.*
2. *If f is differentiable in an open ball around x^*, and is twice differentiable at x^*, then $H_f(x^*)$ is negative semi-definite.*

More

- (FOC) To maximize f, in practice we usually take partials of f and set them equal to 0,
- (SOC) Negative semi-definite $H_f(x^*)$ may help us to rule out some solutions to FOC but are not maximizers of the problem.

Notice that a negative definite $H_f(x^*)$ is not sufficient for x being a (global) maximizer, since $H_f(x^*)$ only gives us local properties of the function f.

If the objective function is concave, we can attain the global maximizer.

Constrained Optimization: Kuhn-Tucker Theorem

Definition

Let X be an open set in $\mathbb{R}^n$, and let $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^k$, and $h : X \to \mathbb{R}^m$ be C^1 functions. Consider the problem

$$\max_{x \in X} f(x) \ s.t. \ g(x) \geq 0 \ and \ h(x) = 0$$

For a feasible point $\hat{x} \in X$, the inequality constraint $g_j(x) \geq 0$ is said to be **binding at** $\hat{x}$ iff $g_j(\hat{x}) = 0$. We say that the **constraint qualification (CQ) holds** at $\hat{x}$ iff the derivatives of all binding constraints

$$\{\nabla g_j(\hat{x})\}_{\{j: g_j \ binding \ at \ \hat{x}\}} \cup \{\nabla h_l(\hat{x})\}_{l=1}$$

in $\mathbb{R}^n$ are linearly independent; otherwise we say that the **constraint qualification (CQ) fails** at $\hat{x}$.

Lagrangian Function

In practice, we often define the Lagrangian function of the maximization problem as

$$\begin{aligned}\mathscr{L}(x,\lambda,\mu) &= f(x) + \lambda^T g(x) + \mu^T h(x) \\ &= f(x) + \sum_{j=1}^{k} \lambda_j g_j(x) + \sum_{l=1}^{m} \mu_l h_l(x)\end{aligned}$$

and λ_j's and μ's are called the Lagrangian multipliers.

Kuhn-Tucker Theorem

Theorem

(Kuhn-Tucker). Let X be an ***open*** *set in $\mathbb{R}^n$, and let $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^k$, and $h : X \to \mathbb{R}^m$ be C^1 functions. Consider the problem*

$$\max_{x \in X} f(x) \ s.t. \ g(x) \geq 0 \ and \ h(x) = 0$$

If x^* is a maximizer of the problem above, and CQ holds at x^*, then there exists a unique $(\lambda, \mu) \in \mathbb{R}^k_+ \times \mathbb{R}^m_+$ s.t. the following two conditions hold:
(1) First order condition (FOC):

$$\nabla f(x^*) + \lambda^T g'(x^*) + \mu^T h'(x^*) = 0$$

(2) Complementary slackness condition (CSC):

$$h_l(x^*) = 0, \ l \in \{1, ..., m\}$$

$$\lambda_j \geq 0, \ g_j(x^*) \geq 0, \ and \ \lambda_j g_j(x^*) = 0, \ j \in \{1, ..., k\}$$

How to Apply in Practice

In practice, we often write down the following system of conditions

$$\begin{cases} x \in X \\ \frac{\partial f}{\partial x_i}(x) + \sum_{j=1}^{k} \lambda_j \frac{\partial g_j}{\partial x_i}(x) + \sum_{l=1}^{m} \mu_l \frac{\partial h_l}{\partial x_i}(x) = 0 & \forall i = 1, ..., n \\ h_l(x) = 0, & \forall l = 1, ..., m \\ \lambda_j \geq 0,\ g_j(x) \geq 0,\ \textit{and}\ \lambda_j g_j(x) = 0, & \forall j = 1, ..., k \end{cases}$$

Which is sometimes known as the Kuhn-Tucker condition.

Example

Consider the problem

$$\max_{(x_1,x_2)\in\mathbb{R}^2} -x_2$$

$$s.t.$$

$$x_1^2 - x_2^3 = 0$$

Clearly, the unique maximizer $(0,0)$. But if we write down the Lagrangian

$$L(x_1,x_2,\lambda) = -x_2 + \lambda(x_1^2 - x_2^3)$$

and consider the FOC

$$\begin{cases} \frac{\partial L}{\partial x_1}(x_1,x_2,\lambda) = 2\lambda x_1 = 0 \\ \frac{\partial L}{\partial x_2}(x_1,x_2,\lambda) = -1 - 3\lambda x_2^2 = 0 \end{cases}$$

There is no $\lambda \in \mathbb{R}$ s.t. $(0,0,\lambda)$ satisfies the FOC above.

The Right Way...

Solve the following maximization problem

$$\max_{(x_1, x_2 \in \mathbb{R}^2_+)} x_1^{\alpha} x_2^{1-\alpha}$$

$$s.t.$$

$$p_1 x_1 + p_2 x_2 \leq m$$

where $\alpha \in (0,1)$, $p_1, p_2 \in \mathbb{R}_{++}$, and $m \in \mathbb{R}_+$are parameters.
We know that the problem always has a solution by Weierstrass theorem.
But because the power function x^{α} is only defined on $\mathbb{R}_+$ for $\alpha \in (0,1)$
and $\mathbb{R}^2_+$ is not open in $\mathbb{R}^2$,
it does not satisfy the assumption of K-T theorem.

The Right Way...

If $m = 0$, then the only feasible point is $(x_1, x_2) = (0, 0)$, and the problem becomes trivial: it has a unique maximizer $(0,0)$, and the maximum is 0. If $m > 0$, consider the feasible point $\hat{x} = (m/(2p_1), (m/2p_2))$, At this point, the objective takes a strictly positive value. So we can just focus on the domain $\mathbb{R}^2_{++}$.and the new problem can be defined

$$\max_{(x_1, x_2 \in \mathbb{R}^2_{++})} x_1^{\alpha} x_2^{1-\alpha}$$

$$s..t. p_1 x_1 + p_2 x_2 \leq m$$

Write down the Lagrangian

$$L(x_1, x_2, \lambda) = x_1^{\alpha} x_2^{1-\alpha} + \lambda(m - p_1 x_1 - p_2 x_2)$$

and then the K-T condition

$$\begin{cases} x \in \mathbb{R} \\ \alpha x_1^{\alpha-1} x_2^{1-\alpha} - \lambda p_1 = 0 \\ (1-\alpha) x_1^{\alpha} x_2^{-\alpha} - \lambda p_2 = 0 \\ \lambda \geq 0, m - p_1 x_1 - p_2 x_2 \geq 0, and \ \lambda(m - p_1 x_1 - p_2 x_2) = 0 \end{cases}$$

After the algebra operation, we get

$$\frac{p_1 x_1}{p_2 x_2} = \frac{\alpha}{1-\alpha}$$

so we have

$$(x_1, x_2, \lambda) = \left(\frac{\alpha m}{p_1}, \frac{(1-\alpha)m}{p_2}, \frac{\alpha^\alpha (1-\alpha)^{1-\alpha}}{p_1^\alpha p_2^{1-\alpha}} \right)$$

as the unique solution to the K-T condition. Consider the first case (m=0), we know that problem has a unique maximizer $(x_1^*, x_2^*) = (\alpha m/p_1, (1-\alpha)m/p_2)$.

Sufficient condition

The previous theorem offers the necessary condition. Now we learn the sufficient condition

Theorem

Let X be an open and convex set in $\mathbb{R}^n$, and let $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^k$, and $h : X \to \mathbb{R}^m$ be C^1 functions. Consider the problem

$$\max_{x \in X} f(x) \ s.t. \ g(x) \geq 0 \ and \ h(x) = 0$$

If x^* is feasible, and there exists $(\lambda, \mu) \in \mathbb{R}^k_+ \times \mathbb{R}^m$ s.t. the following three conditions hold

(1) FOC:

$$\nabla f(x^*) + \lambda^T g'(x^*) + \mu^T h'(x^*) = 0$$

(2) CSC:

$$\lambda_j \geq 0, g_j(x^*) \geq 0, \ and \ \lambda_j g_j(x^*) = 0$$

for each $j \in \{1, ..., k\}$, and

(3) The Lagrangian $L_{\lambda,\mu} : X \to \mathbb{R}$ defined as

$$L_{\lambda,\mu}(x) := f(x) + \lambda^T g(x) + \mu^T h(x)$$

is a concave function, then x^* is a maximizer of this problem.

Envelope

Let's consider the parameterized optimization problem $P(\alpha)$:

$$\max_{x \in X} f(x, \alpha) \; s.t. \; g(x) \geq 0 \; and \; h(x) = 0$$

where the parameter α is taken from some set A. Assuming the maximization problem has the solution which is defined as $f^*(\alpha)$. it might be interesting to study how the value function $f^*(\alpha)$ changes as the parameter α changes.

Envelope Theorem

Define the value function $f^* : \hat{A} \to \mathbb{R}$ as

$$f^*(\alpha) := \max_{x \in X} \{f(x,\alpha) : g(x,\alpha) \geq 0 \; and \; h(x,\alpha) = 0\}$$

For parameter $\alpha^* \in A$, suppose:
(1) In the problem $P(\alpha^*)$, there is a unique maximizer x^*, and CQ holds at x^*.
(2) There exists $\varepsilon > 0$ and $r > 0$ s.t. $\forall \alpha \in B_\varepsilon(\alpha^*)$, $(\arg\max P(\alpha)) \cap B_r(x^*) \neq \emptyset$. Then the value function $f^*$is differentiable at a^*, and

$$\begin{aligned} f^{*\prime}(\alpha^*) &= \frac{d}{d\alpha} L(x^*, \lambda^*, \mu^*, \alpha)|_{\alpha=\alpha^*} \\ &= \frac{d}{d\alpha} f(x^*, \alpha)|_{\alpha=\alpha^*} + \lambda^{*T} \frac{d}{d\alpha} g(x^*, \alpha)|_{\alpha=\alpha^*} + \mu^{*T} \frac{d}{d\alpha} h(x^*, \alpha)|_{\alpha=\alpha^*} \end{aligned}$$

where λ^* and μ^* are the unique Lagrangian multipliers found by K-T theorem for the problem $P(\alpha^*)$.

Example

Remember the previous example which the solution is

$$(x_1, x_2, \lambda) = \left(\frac{\alpha m}{p_1}, \frac{(1-\alpha)m}{p_2}, \frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{p_1^{\alpha} p_2^{1-\alpha}} \right)$$

The objective function (value function) is

$$v(p_1, p_2 m) = (x_1^*)^{\alpha}(x_2^*)^{1-\alpha} = \frac{m\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{p_1^{\alpha} p_2^{1-\alpha}}$$

Without envelope theorem, if we want to know the impact of $p_1, p_2,$ and$m$, we need to do

$$\frac{\partial v}{\partial p_1} = -\frac{m\alpha^{1+\alpha}(1-\alpha)^{1-\alpha}}{p_1^{1+\alpha} p_2^{1-\alpha}} = -\lambda^* x_1^*$$

$$\frac{\partial v}{\partial p_2} = -\frac{m\alpha^{\alpha}(1-\alpha)^{2-\alpha}}{p_1^{\alpha} p_2^{2-\alpha}} = -\lambda^* x_2^*$$

$$\frac{\partial v}{\partial m} = \frac{m\alpha^{1+\alpha}(1-\alpha)^{1-\alpha}}{p_1^{1+\alpha} p_2^{1-\alpha}} = \lambda^*$$

Using the envelope theorem, we can have

$$\frac{\partial \mathscr{L}}{\partial p_1}(x_1^*, x_2^*, \lambda^*, p_1, p_2, m) = -\lambda^* x_1^*$$
$$\frac{\partial \mathscr{L}}{\partial p_2}(x_1^*, x_2^*, \lambda^*, p_1, p_2, m) = -\lambda^* x_2^*$$
$$\frac{\partial \mathscr{L}}{\partial m}(x_1^*, x_2^*, \lambda^*, p_1, p_2, m) = \lambda^*$$

Interpretation of Lagrangian Multipliers

Consider the parameterized problem $P(a, b)$

$$\max_{x \in X} f(x)$$

s.t.

$$\begin{cases} g(x) + \alpha & \geq 0 \\ h(x) + b & = 0 \end{cases}$$

where $(a, b) \in \mathbb{R}^k \times \mathbb{R}^m$ are parameters. If the problem $P(a, b)$ has a solution, define $f^*(a, b)$ as the maximum value of the problem $P(a, b)$.

A Small Perturbation

- When we move (a, b) around $(a^*, b^*) = (0, 0)$, we are considering perturbations around the original problem $P(0, 0)$.
- A small positive a_j can be viewed as a slight relaxation of the constraint $g_j(x) \geq 0$, which might make the feasible set slightly larger, which in turn might make the maximum value slightly higher.
- we are interested in how such a slight relaxation of the constraint $g_j(x) \geq 0$ will affect the maximum value

Using the Envelope theorem at $(a^*, b^*) = (0,0)$, we have

$$\frac{\partial f^*}{\partial a_j}(0,0) = \lambda_j,\ j = 1, ..., k$$
$$\frac{\partial f^*}{\partial b_l}(0,0) = \mu_l,\ l = 1, ..., m$$

The Lagrangian multiplier λ_j corresponding to the inequality constraint $g_j(x) \geq 0$ measures the marginal increase in the maximum value under a marginal relaxation of the constraint $g_j(x) \geq 0$.

Economic Interpretation

In a firm's maximization problem:

- The objective function: the firms profit function,
- The constraint $g_j(x) \geq 0$: the requirement that total usage of some resource (labor/capital/electricity/...) is weakly less than the total amount of this resource available to the firm.
- The multiplier λ_j : the shadow price of this resource (labor/capital/electricity/...), which measures the marginal increase in profit by marginally increasing the total amount of this resource available to the firm.

Introduction to Dynamic Programming

Consider infinite horizon maximization problem

$$\max_{\{c_k, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t)$$

$$\begin{aligned}
s.t. \quad & \\
& c_t + k_{t+1} \leq f(k_t) \\
& c_t, k_{t+1} \geq 0, \;\; t = 0, 1, ... \\
& k_0 > 0 \; given
\end{aligned}$$

where U and f are strictly increasing. Here the choice variables are a sequence $\{c_t, k_{t+1}\}_{t=0}^{\infty}$.

Define the value function $v : \mathbb{R}_+ \to \mathbb{R}$

$$v(k_0) = \max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t)$$

Since the problem is time-independent, we can also write as

$$\max_{c_0, k_1} U(c_0) + \beta v(k_1)$$

Take the binding constraint, we substitute out c_0 and have

$$v(k_0) = \max_{0 \leq k_1 \leq f(k_0)} U(f(k_0) - k_1) + \beta v(k_1) \quad (*)$$

But we are not quite sure about whether there exists a maximized solution of the $(*)$. The Blackwell's sufficient condition for contraction guarantee the maximization problem holds and there exists a optimal path $\{k_t\}_{t=1}^{\infty}$ lead to the maximal value of $v(k_0)$.
See cake eating problem in the reading material

Blackwell's sufficient condition

Let $X \subset \mathbb{R}^k$ and $B(X)$ a real vector space of bounded functions $f : X \to \mathbb{R}$, with norm defined as $||f|| = \sup_{x \in X} |f(x)|$. Let $T : B(X) \to B(X)$ be an operator satisfying
(1) (monotonicity) If $f, g \in B(X)$ and $f(x)g(x)$ for $\forall x \in X$, then $(Tf)(x) \leq (Tg)(x)$ for $\forall x \in X$
(2)(discounting) There exists some $\beta \in (0, 1)$ s.t.

$$(T(f + a))(x) \leq (Tf)(x) + \beta\alpha, \text{ for } \forall f \in B(X), \alpha \leq 0, \ x \in X$$

where $(f + a)$ is defined as $(f + a)(x) = f(x) + a$.
Then T is a construction with modulus β.