8. Optimization

Daisuke Oyama

Mathematics II

May 13, 2025

Unconstrained Maximization Problem

Let $X \subset \mathbb{R}^N$ be a nonempty set.

Definition 8.1

For a function $f: X \to \mathbb{R}$,

- x̄ ∈ X is a (strict) local maximizer of f if there exists an open neighborhood A ⊂ X of x̄ relative to X such that f(x̄) ≥ f(x) for all x ∈ A (f(x̄) > f(x) for all x ∈ A with x ≠ x̄);
- $\bar{x} \in X$ is a maximizer (or global maximizer) of f if $f(\bar{x}) \ge f(x)$ for all $x \in X$.

(Local and global minimizers are defined analogously.)

First-Order Condition for Optimality

Let $X \subset \mathbb{R}^N$ be a nonempty set.

Proposition 8.1

For $f: X \to \mathbb{R}$, if

• $\bar{x} \in X$ is a local maximizer or local minimizer of f,

• $\bar{x} \in \operatorname{Int} X$, and

• f is differentiable at \bar{x} , then $\nabla f(\bar{x}) = 0$.

Proof

Apply the FOC for the one variable case to $f(x_i, \bar{x}_{-i})$ for each i = 1, ..., N.

Second-Order Condition for Optimality

Let $X \subset \mathbb{R}^N$ be a nonempty set.

Proposition 8.2

For $f: X \to \mathbb{R}$, suppose that $\bar{x} \in \text{Int } X$ and that f is differentiable on Int X and ∇f is differentiable at \bar{x} .

- 1. If \bar{x} is a local maximizer of f, then $D^2 f(\bar{x})$ is negative semi-definite.
- 2. If $\nabla f(\bar{x}) = 0$ and $D^2 f(\bar{x})$ is negative definite, then \bar{x} is a strict local maximizer of f.

Proof

1.

 Fix any z ∈ ℝ^N, z ≠ 0.
 Let h(α) = f(x̄ + αz) - f(x̄) (where α ∈ ℝ is sufficiently close to 0).
 Note that h is differentiable and h' is differentiable at α = 0.

• Recall that
$$h''(\alpha) = z \cdot D^2 f(\bar{x} + \alpha z) z$$
.

• If h''(0) > 0, then $\alpha = 0$ would be a strict local minimizer.

• Thus,
$$h''(0) \leq 0$$
, or $z \cdot D^2 f(\bar{x}) z \leq 0$.

Suppose that $\nabla f(\bar{x}) = 0$ and $D^2 f(\bar{x})$ is negative definite.

Since
$$u \cdot D^2 f(\bar{x})u$$
 is continuos in u and
since $\{u \in \mathbb{R}^N \mid ||u|| = 1\}$ is compact,
it follows from the assumption of negative definiteness and
the Extreme Value Theorem that there is some $\varepsilon > 0$ such
that

$$\frac{1}{2}u \cdot D^2 f(\bar{x})u + \varepsilon < 0 \text{ for all } u \in \mathbb{R}^N \text{ such that } \|u\| = 1.$$

Since ∇f(x̄) = 0, by Taylor's Theorem we can take a sufficiently small δ > 0 such that

$$0 < \|z\| < \delta \Rightarrow \frac{f(\bar{x} + z) - f(\bar{x})}{\|z\|^2} - \frac{\frac{1}{2}z \cdot D^2 f(\bar{x})z}{\|z\|^2} \le \varepsilon.$$

Now take any $x \in B_{\delta}(\bar{x})$, $x \neq \bar{x}$. Then,

$$\frac{f(x) - f(\bar{x})}{\|x - \bar{x}\|^2} \le \frac{1}{2} \frac{x - \bar{x}}{\|x - \bar{x}\|} \cdot D^2 f(\bar{x}) \frac{x - \bar{x}}{\|x - \bar{x}\|} + \varepsilon < 0,$$

where the last inequality follows from the choice of $\varepsilon.$ Thus, $f(x) < f(\bar{x}).$

Concave Functions

Let $X \subset \mathbb{R}^N$ be a nonempty convex set.

Proposition 8.3

For $f: X \to \mathbb{R}$, suppose that $\bar{x} \in \operatorname{Int} X$ and f is differentiable at \bar{x} .

If $\nabla f(\bar{x}) = 0$, then \bar{x} is a global maximizer of f.

If $\nabla f(\bar{x}) = 0$, then \bar{x} is a unique global maximizer of f.

Proof

• Take any
$$x \in X$$
, $x \neq \bar{x}$.

If f is concave, then we have

$$f(x) \le f(\bar{x}) + \nabla f(\bar{x}) \cdot (x - \bar{x}),$$

with a strict inequality if f is strictly concave.

▶ Thus, if $\nabla f(\bar{x}) = 0$, we have $f(x) \leq f(\bar{x})$ if f is concave, and $f(x) < f(\bar{x})$ if f is strictly concave.

Equality Constrained Maximization Problem

Let $X \subset \mathbb{R}^N$ be a nonempty open set, and $f, g_1, \ldots, g_M \colon X \to \mathbb{R}$, where M < N.

Consider the maximization problem:

$$\max_{x} f(x)$$
(P)
s.t. $g_1(x) = 0$
 \vdots
 $g_M(x) = 0.$

Virte
$$g: X \to \mathbb{R}^M$$
, $x \mapsto (g_1(x), \dots, g_M(x))$, and
 $C = \{x \in X \mid g(x) = 0\}.$

▶ $\bar{x} \in C$ is a local (global, resp.) constrained maximizer of (P) if it is a local (global, resp.) maximizer of $f|_C$.

First-Order Condition for Optimality

Proposition 8.4

Suppose that

• $\bar{x} \in C$ is a local constrained maximizer of (P);

• f is differentiable at \bar{x} ;

• g_1, \ldots, g_M are C^1 on a neighborhood of \bar{x} ; and

▶ rank $Dg(\bar{x}) = M$ ("constraint qualification"). Then there exist unique $(\bar{\lambda}_1, \ldots, \bar{\lambda}_M) \in \mathbb{R}^M$ (Lagrange multipliers) such that

$$\nabla f(\bar{x}) = \sum_{m=1}^{M} \bar{\lambda}_m \nabla g_m(\bar{x}).$$

Expression with Lagrangian

• Let
$$L: X \times \mathbb{R}^M \to \mathbb{R}$$
 be defined by

$$L(x,\lambda) = f(x) - \sum_{m=1}^{M} \lambda_m g_m(x).$$

► Then the FOC is:

there exists $\bar{\lambda} \in \mathbb{R}^M$ such that

$$\frac{\partial L}{\partial x_n}(\bar{x},\bar{\lambda}) = 0, \quad n = 1,\dots,N,$$
$$\frac{\partial L}{\partial \lambda_m}(\bar{x},\bar{\lambda}) = 0, \quad m = 1,\dots,M,$$

or

$$\nabla L(\bar{x}, \bar{\lambda}) = 0.$$

Proof

- ► Let $\bar{x} \in C$ be a local constrained maximizer. By assumption $Dq(\bar{x}) \in \mathbb{R}^{M \times N}$ has rank M.
- Without loss of generality, assume that the first M columns of $Dg(\bar{x})$ are linearly independent.

Write x = (p, q), where $p \in \mathbb{R}^M$ and $q \in \mathbb{R}^{N-M}$.

▶ By the Implicit Function Theorem, the equation g(p,q) = 0 is locally solved as $p = \eta(q)$, where

$$D\eta(\bar{q}) = -[D_p g(\bar{p}, \bar{q})]^{-1} D_q g(\bar{p}, \bar{q}).$$

• Consider the unconstrained maximization problem $F(q) = f(\eta(q), q)$, where \bar{q} is a local maximizer.

▶ By the FOC $DF(\bar{q}) = 0$, we have

$$\begin{split} 0 &= D_q f(\eta(q), q)|_{q = \bar{q}} \\ &= D_p f(\bar{x}) D\eta(\bar{q}) + D_q f(\bar{x}) \\ &= -D_p f(\bar{x}) [D_p g(\bar{x})]^{-1} D_q g(\bar{x}) + D_q f(\bar{x}). \end{split}$$

• Let
$$\bar{\lambda}^{\mathrm{T}} = D_p f(\bar{x}) [D_p g(\bar{x})]^{-1}$$
, where $\bar{\lambda} \in \mathbb{R}^M$.

Then we have

$$D_p f(\bar{x}) = \bar{\lambda}^{\mathrm{T}} D_p g(\bar{x}), \quad D_q f(\bar{x}) = \bar{\lambda}^{\mathrm{T}} D_q g(\bar{x}),$$

or

$$\nabla f(\bar{x}) = Dg(\bar{x})^{\mathrm{T}}\bar{\lambda} = \sum_{m=1}^{M} \bar{\lambda}_m \nabla g_m(\bar{x}).$$

Second-Order Condition for Optimality

Proposition 8.5

Suppose that f, g_1, \ldots, g_M are of C^2 class, $\bar{x} \in C$, and rank $Dg(\bar{x}) = M$. Denote $W = \{z \in \mathbb{R}^N \mid Dg(\bar{x})z = 0\}.$

- 1. If \bar{x} is a local constrained maximizer of (P), then $D_x^2 L(\bar{x}, \bar{\lambda})$ is negative semi-definite on W, where $\bar{\lambda} \in \mathbb{R}^M$ is such that $\nabla L(\bar{x}, \bar{\lambda}) = 0$.
- 2. If there exists $\bar{\lambda} \in \mathbb{R}^M$ such that $\nabla L(\bar{x}, \bar{\lambda}) = 0$ and $D_x^2 L(\bar{x}, \bar{\lambda})$ is negative definite on W, then \bar{x} is a strict local constrained maximizer of (P).

Inequality Constrained Maximization Problem Let $X \subset \mathbb{R}^N$ be a nonempty open set, and $f, g_1, \ldots, g_M, h_1, \ldots, h_K \colon X \to \mathbb{R}$, where M < N.

Consider the maximization problem:

 $\max f(x)$ (P) s.t. $q_1(x) = 0$ $q_M(x) = 0$ $h_1(x) < 0$ $h_{K}(x) < 0.$

- Write $C = \{x \in X \mid g(x) = 0, h(x) \le 0\}.$
- ▶ $\bar{x} \in C$ is a local (global, resp.) constrained maximizer of (P) if it is a local (global, resp.) maximizer of $f|_C$.

First-Order Condition for Optimality (KKT Conditions)

For
$$x \in C$$
, write $\mathcal{I}(x) = \{k \mid h_k(x) = 0\}.$

Proposition 8.6

Suppose that

• $\bar{x} \in C$ is a local constrained maximizer of (P);

•
$$f, h_1, \ldots, h_K$$
 are differentiable at \bar{x} ;

• g_1, \ldots, g_M are C^1 on a neighborhood of \bar{x} ; and

▶
$$\nabla g_1(\bar{x}), \ldots, \nabla g_M(\bar{x})$$
 and $\nabla h_k(\bar{x}), k \in \mathcal{I}(\bar{x})$, are linearly independent ("constraint qualification").

Then there exist $\bar{\mu}_1, \ldots, \bar{\mu}_M \in \mathbb{R}$ and $\bar{\lambda}_1, \ldots, \bar{\lambda}_K \in \mathbb{R}$ such that

(i)
$$\nabla f(\bar{x}) = \sum_{m=1}^{M} \bar{\mu}_m \nabla g_m(\bar{x}) + \sum_{k=1}^{K} \bar{\lambda}_k \nabla h_k(\bar{x})$$
, and

(ii) $\bar{\lambda}_k \ge 0$ and $\bar{\lambda}_k h_k(\bar{x}) = 0$ for each $k = 1, \dots, K$.

"\$\overline{\lambda}_k h_k(\$\overline{x}\$) = 0" is called the complementarity condition.
It says: \$\overline{\lambda}_k = 0\$ for all \$k \notice \$\mathcal{I}(\$\overline{x}\$)\$, where \$\mathcal{L}(\$x\$) = {\$k | h_k(\$x\$) = 0\$}.

Example 1 Let $X = \mathbb{R}$. Consider $\max_{x \in [0,1]} f(x)$,

or

$$\max_{x} f(x) s. t. h_{1}(x) = -x \le 0 h_{2}(x) = x - 1 \le 0.$$

• If $\bar{x} \in [0,1]$ is a local constrained maximizer, then clearly we have:

1. if
$$\bar{x} \in (0, 1)$$
, then $f'(\bar{x}) = 0$,

2. if
$$\bar{x} = 0$$
, then $f'(\bar{x}) \le 0$,

3. if
$$\bar{x} = 1$$
, then $f'(\bar{x}) \ge 0$.

Let

$$L(x,\lambda) = f(x) - \lambda_1(-x) - \lambda_2(x-1).$$

The KKT conditions are:

$$L_x(x,\lambda) = f'(x) + \lambda_1 - \lambda_2 = 0 \iff f'(x) = -\lambda_1 + \lambda_2,$$

$$\lambda_1 \ge 0, \ \lambda_1(-x) = 0,$$

$$\lambda_2 \ge 0, \ \lambda_2(x-1) = 0.$$



1. if $-\bar{x} < 0$ and $\bar{x} - 1 < 0$, then $\lambda_1 = \lambda_2 = 0$, so $f'(\bar{x}) = 0$, 2. if $-\bar{x} = 0$ and $\bar{x} - 1 < 0$, then $\lambda_2 = 0$, so $f'(\bar{x}) = -\lambda_1 \le 0$, 3. if $-\bar{x} < 0$ and $\bar{x} - 1 = 0$, then $\lambda_1 = 0$, so $f'(\bar{x}) = \lambda_2 \ge 0$.

- ► To see why we have \u03c6 k ≥ 0, suppose that \u03c5 x satisfies the constraint h_k(x) ≤ 0 with "=" (i.e., h_k(\u03c5) = 0).
- For $z \approx 0$, $f(\bar{x}+z) \approx f(\bar{x}) + f'(\bar{x})z$ and $h_k(\bar{x}+z) \approx h'_k(\bar{x})z$.
- ▶ If $f'(\bar{x}) > 0$, then for small $\varepsilon > 0$, $\bar{x} + \varepsilon$ has to violate the constraint, for which we have to have $h'_k(\bar{x}) \ge 0$.

(Constraint qualification implies that $h_k'(\bar{x}) \neq 0.$)

- If f'(x̄) < 0, then for small ε > 0, x̄ − ε has to violate the constraint, for which we have to have h'_k(x̄) ≤ 0.
- ▶ In these cases, we have $f'(\bar{x}) = \lambda_k h'_k(\bar{x})$ with $\lambda_k > 0$.
- lt is possible that $f'(\bar{x}) = 0$, so it may be the case that $\lambda_k = 0$.

For $p \gg 0$ and w > 0, consider

$$\max_{x} \quad u(x)$$

s.t. $p \cdot x - w \le 0$
 $-x_1 \le 0, \dots, -x_N \le 0.$

▶ The KKT conditions: $\bar{x} \neq 0$,

$$\nabla u(\bar{x}) = \mu p - \sum_{n=1}^{N} \lambda_n e_n,$$

$$\mu \ge 0, \ \mu(p \cdot \bar{x} - w) = 0,$$

$$\lambda_n \ge 0, \ \lambda_n(-\bar{x}_n) = 0 \quad (n = 1, \dots, N).$$

These can be written as

$$\begin{split} &\frac{\partial u}{\partial x_n}(\bar{x}) \leq \mu p_n, \quad \text{with equality if } \bar{x}_n > 0 \quad (n = 1, \dots, N), \\ &\mu \geq 0, \ \mu (p \cdot \bar{x} - w) = 0. \end{split}$$

Let N = 2.

Suppose that
$$\bar{x} = (w/p_1, 0)$$
.

First, we have to have ∂u/∂x₁(x) ≥ 0.
So we have ∂u/∂x₁(x) = λp₁ for some λ ≥ 0.
Thus, we have to have ∂u/∂x₂(x) ≤ λp₂.
(Draw a picture.)

Proof of Proposition 8.6

Case with no equality constraint.

▶ Note that for any
$$z \in \mathbb{R}^N$$
,

$$\begin{split} f(\bar{x} + tz) &= f(\bar{x}) + (\nabla f(\bar{x}) \cdot z)t + o(t), \\ h_k(\bar{x} + tz) &= (\nabla h_k(\bar{x}) \cdot z)t + o(t) \quad \text{for all } k \in \mathcal{I}(\bar{x}). \end{split}$$

- ▶ Since \bar{x} is a local constrained maximizer, there is no $z \in \mathbb{R}^N$ such that $\nabla f(\bar{x}) \cdot z > 0$ and $\nabla h_k(\bar{x}) \cdot z < 0$ for all $k \in \mathcal{I}(\bar{x})$, or $\begin{pmatrix} Df(\bar{x}) \\ -Dh_{\mathcal{I}}(\bar{x}) \end{pmatrix} z \gg 0.$
- ▶ Thus, by Gordan's Theorem, there exist $\lambda_0, \lambda_k \ge 0$, $k \in \mathcal{I}(\bar{x})$, such that

$$\lambda_0 \nabla f(\bar{x}) - \sum_{k \in \mathcal{I}(\bar{x})} \lambda_k \nabla h_k(\bar{x}) = 0, \quad (\lambda_0, \lambda_k)_{k \in \mathcal{I}(\bar{x})} \neq 0.$$

• By the constraint qualification, $\lambda_0 \neq 0$, so normalize $\lambda_0 \equiv 1$.

Proof of Proposition 8.6

Case with inequality and equality constraints.

• We show that there is no $z \in \mathbb{R}^N$ such that $Df(\bar{x})z > 0$, $-Dh_{\mathcal{I}}(\bar{x})z \gg 0$, and $Dg(\bar{x})z = 0$.

• Write x = (p,q), where $p \in \mathbb{R}^M$ and $q \in \mathbb{R}^{N-M}$.

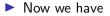
g(p,q)=0 is solved locally around $\bar{x}=(\bar{p},\bar{q})$ as $p=\eta(q)$, where $D\eta(\bar{q})=-[D_pg(\bar{x})]^{-1}D_qg(\bar{x}).$

Suppose that $Dg(\bar{x})z = 0$, or $D_pg(\bar{x})u + D_qg(\bar{x})v = 0$ so that $u = -[D_pg(\bar{x})]^{-1}D_qg(\bar{x})v = D\eta(\bar{q})v$, where z = (u, v).

• Let
$$x(t) = (\eta(\bar{q} + tv), \bar{q} + tv).$$

Then

$$Dx(0) = \begin{pmatrix} D\eta(\bar{q})v\\v \end{pmatrix} = \begin{pmatrix} u\\v \end{pmatrix} = z.$$



$$\begin{split} f(x(t)) &= f(\bar{x}) + (\nabla f(\bar{x}) \cdot Dx(0))t + o(t) \\ &= f(\bar{x}) + (\nabla f(\bar{x}) \cdot z)t + o(t), \\ h_k(\bar{x} + tz) &= (\nabla h_k(\bar{x}) \cdot Dx(0))t + o(t) \\ &= (\nabla h_k(\bar{x}) \cdot z)t + o(t) \quad \text{for all } k \in \mathcal{I}(\bar{x}). \end{split}$$

Since \bar{x} is a local constrained maximizer, we cannot have $\nabla f(\bar{x}) \cdot z > 0$ and $h_k(\bar{x}) \cdot z < 0$ for all $k \in \mathcal{I}(\bar{x})$.

▶ I.e.,
$$\exists z \in \mathbb{R}^N$$
 such that $\begin{pmatrix} Df(\bar{x}) \\ -Dh_{\mathcal{I}}(\bar{x}) \end{pmatrix} z \gg 0$ and $Dg(\bar{x})z = 0.$

▶ Thus, by Motzkin's Theorem, there exist $(\lambda_0, \lambda_I) \geqq 0$ and μ such that

$$\begin{pmatrix} \lambda_0 & \lambda_{\mathcal{I}}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} Df(\bar{x}) \\ -Dh_{\mathcal{I}}(\bar{x}) \end{pmatrix} + \mu^{\mathrm{T}} Dg(\bar{x}) = 0.$$

• By the constraint qualification, $\lambda_0 \neq 0$; so normalize $\lambda_0 \equiv 1$.

Second-Order Condition for Optimality

Proposition 8.7

Suppose that $f, g_1, \ldots, g_M, h_1, \ldots, h_K$ are of C^2 class, $\bar{x} \in C$, and $\nabla g_1(\bar{x}), \ldots, \nabla g_M(\bar{x})$ and $\nabla h_k(\bar{x}), k \in \mathcal{I}$, are linearly independent. If

• there exist $\bar{\mu}_1, \ldots, \bar{\mu}_M \in \mathbb{R}$ and $\bar{\lambda}_1, \ldots, \bar{\lambda}_K \in \mathbb{R}$ such that the KKT conditions hold, and

• $D_x^2 L(\bar{x}, \bar{\lambda})$ is negative definite on W, where

$$W = \{ z \in \mathbb{R}^N \mid \nabla g_m(\bar{x}) \cdot z = 0 \text{ for all } m = 1, \dots, M, \\ \nabla h_k(\bar{x}) \cdot z = 0 \text{ for all } k \in \tilde{\mathcal{I}} \},$$

and $\tilde{\mathcal{I}} = \{k \mid \bar{\lambda}_k > 0\},\$

then \bar{x} is a strict local constrained maximizer of (P).

Quasi-Concavity/Convexity

Proposition 8.8

Suppose that f, h_1, \ldots, h_K are of C^1 class and g_1, \ldots, g_M are affine (i.e., $g_m(x) = a^m \cdot x + b^m$), and $\bar{x} \in C$. Suppose that

1.
$$f(x') > f(x) \Longrightarrow \nabla f(x) \cdot (x' - x) > 0$$
, and

2. for all
$$k = 1, \dots, K$$
,
 $h_k(x') \le h_k(x) \Longrightarrow \nabla h_k(x) \cdot (x' - x) \le 0;$

Then if \bar{x} satisfies the KKT conditions for some $\mu_1, \ldots, \mu_M, \lambda_1, \ldots, \lambda_K \in \mathbb{R}$, then \bar{x} is a global constrained maximizer of (P).

Proof

- Let $\bar{x} \in C$ satisfy the KKT conditions, and take any $x' \in C$ with $x' \neq \bar{x}$.
- If $\lambda_k > 0$, then $h_k(\bar{x}) = 0$. With $h_k(x') \le 0$, we have $h_k(x') \le h_k(\bar{x})$.
- ► Therefore, by Condition 2, we have $\nabla h_k(\bar{x}) \cdot (x' \bar{x}) \leq 0$ whenever $\lambda_k > 0$.
- It follows from the KKT conditions that

$$\nabla f(\bar{x}) \cdot (x' - \bar{x}) = \sum_{k \in \mathcal{N}} \mu_m a^m \cdot (x' - \bar{x}) + \sum_{k \in \mathcal{N}} \lambda_k \nabla h_k(\bar{x}) \cdot (x' - \bar{x})$$
$$\leq 0.$$

• Hence, by Condition 1, we have $f(x') \leq f(\bar{x})$.

Remarks

- Condition 2 $\iff h_k$ is quasi-convex.
- ▶ When Condition 1 holds, *f* is called *pseudo-concave*.
- f: strictly quasi-concave and ∇f(x) ≠ 0 for all x
 ⇒ f: pseudo-concave
 ⇒ f: quasi-concave

Quasi-Concavity

Proposition 8.9

Let $C \subset \mathbb{R}^N$ be a nonempty convex set. Suppose that $f: C \to \mathbb{R}$ is strictly quasi-concave, and consider the maximization problem

 $\max_{x \in C} f(x).$

If $\bar{x} \in C$ is a local maximizer, then it is a unique global maximizer.