Constrained Optimization (I): FOC

Ch.18

mailto:eyeofyou@korea.ac.kr

July 15, 2020

Table of Contents

Terms

General Max(Min)imization Problem

$$\arg\max_{\mathbf{x}} f(\mathbf{x}) \quad s.t. \quad \begin{cases} \mathbf{G}(\mathbf{x}) \leq \bar{\mathbf{b}} \\ \mathbf{H}(\mathbf{x}) = \bar{\mathbf{a}} \end{cases}$$

- f: Object function
- $\mathbf{G}(\mathbf{x}) \leq \bar{\mathbf{b}}$: Inequality Constraints
- $\mathbf{H}(\mathbf{x}) = \bar{\mathbf{a}}$: Equality Constraints

Note: From now on, we define inequality of vector as:

$$G_i(\mathbf{x}) \leq \bar{b}_i \quad \forall i$$

Note2: For minimization problem, use $\arg\min$ instead.

Examples

Utility Maximization Problem

$$\arg\max_{\mathbf{x}} U(\mathbf{x}) \quad s.t. \quad \mathbf{p} \bullet \mathbf{x} \leq Income \wedge \mathbf{x} \geq \mathbf{0}$$

- U: Utility function $\mathbf{x}$: consumption bundle $\mathbf{p}$: price vector

Profit Maximization Problem of a Competitive Firm

$$\arg\max_{\mathbf{x}} \Pi(\mathbf{x})$$

$$\Pi(\mathbf{x}) := \bar{p} f(\mathbf{x}) - \mathbf{w} \bullet \mathbf{x}, \quad \Pi \geq 0 \wedge \mathbf{x} \geq \mathbf{0}$$

- f: production function, p: price of final product
- $\mathbf{x}$: quantity bundle of factors for production
- $\mathbf{w}$: price vector of each factor

FOC of Constrained Max(Min)imization of $f : \mathbb{R}^2 \to \mathbb{R}$

Theorem (18.1)

Let f, h be C^1 functions of $\mathbf{x} \in \mathbb{R}^2$. Suppose $\mathbf{x}^*$ is a solution of the max(min)imization problem

$$\arg\max_{\mathbf{x}} f(\mathbf{x}) \quad s.t. \quad h(\mathbf{x}) = \bar{a}$$

If $\mathbf{x}^*$ is not a critical point of h, Then $\exists(\mathbf{x}^*, \mu^*)$ s.t.

$$L(\mathbf{x}, \mu) := f(\mathbf{x}) + \mu(\bar{a} - h(\mathbf{x})) \quad \wedge \quad DL_{\mathbf{x},\mu} = \mathbf{0}$$

Geometrically, L (Lagrangian function) comes from the fact that if $\mathbf{x}^*$ is a solution of max(min)imization problem, then both gradient vectors of f, h on $\mathbf{x}^*$ should be (1) perpendicular to the level sets of f, h respectively and (2) the level sets of f, h have the same slope at $\mathbf{x}^*$

$$\exists\mu^* \quad s.t. \quad \nabla f(\mathbf{x}^*) = \mu^* \nabla h(\mathbf{x}^*) \quad \wedge \quad \nabla h(\mathbf{x}^*) \neq \mathbf{0}$$

Constraint Qualification

Constraint Qualification (CQ)

If $Dh_{\mathbf{x}}(\mathbf{x}^*) = \mathbf{0}$ ($i.e.$, there is critical point of constraint function), we cannot use Th 18.1. This condition is constraint qualification (CQ). If we have points which can not pass the constraint qualification, we should include these points among our candidates for a solution to the original constrained maximization problem, along with the critical points of L

Nondegenerate CQ (NDCQ)

If there are m equality constraints, NDCQ is $DH_{i\mathbf{x}}(\mathbf{x}^*) \neq \mathbf{0}$ for all $i = 1, \cdots, m$ and this condition should be valid even in its row echelon form (REF). Generally, NDCQ implies

$$rank D\mathbf{H}_{\mathbf{x}}(\mathbf{x}^*) = m$$

Note: NDCQ is a regularity condition: passing NDCQ implies that the constraint set has a well-defined $(n - m)$ dimensional tangent hyperplane

FOCs of General Equality Constraints

We can extend Th 18.1 to general FOCs of constrained max(min)imization problem

General Max(Min)imization problem with Equality Constraints

$$\arg\max_{\mathbf{x}} f(\mathbf{x}) \quad s.t. \quad \mathbf{H}(\mathbf{x}) = \bar{\mathbf{a}}$$

Here equality constraints $\mathbf{H}(\mathbf{x}) = \bar{\mathbf{a}}$ means:

$$\begin{aligned} H_1(\mathbf{x}) &= a_1 \\ &\vdots \\ H_m(\mathbf{x}) &= a_m \end{aligned}$$

FOCs of General Equality Constraints

Theorem (18.2)

Let $f, \mathbf{H}$ *be* C^1 *functions of* n *variables (*$i.e.$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{H} : \mathbb{R}^n \to \mathbb{R}^m$*). Consider the max(min)imization problem with* m *equality constraints:*

$$\arg\max_{\mathbf{x}} f(\mathbf{x}) \quad s.t. \quad \mathbf{H}(\mathbf{x}) = \bar{\mathbf{a}}$$

Suppose (1) $\mathbf{H}(\mathbf{x}^*) = \bar{\mathbf{a}}$, *(2)* $\mathbf{x}^*$ *is a local max (or min) of* f *on* $\mathbf{H}(\mathbf{x}) = \bar{\mathbf{a}}$ *and (3)* $rank D\mathbf{H}_{\mathbf{x}}(\mathbf{x}^*) = m$ *(NDCQ). Then,* $(\mathbf{x}^*, \mu^*) \in R^{n+m}$ *is a critical point of the Lagrangian*

$$L(\mathbf{x}, \mu) = f(\mathbf{x}) + \mu(\bar{\mathbf{a}} - \mathbf{H}(\mathbf{x}))$$

I.e.,

$$DL_{\mathbf{x},\mu}(\mathbf{x}^*, \mu^*) = \mathbf{0}$$

Note: $\mu = (\mu_1, \cdots, \mu_m) \in \mathbb{R}^m$

One Inequality Constraints

Inequality Constraints: Main Concept

Inequality constrained solution = [Equality Constrained solution] (corner solution, binding) or [Unconstrained solution] (internal solution, not binding)

Theorem (18.3)

f, g are C^1 function on $\mathbb{R}^2$ and $\mathbf{x}^*$ max(min)imizes f on the inequality constraint set $g(\mathbf{x}) \leq b$. If $g(\mathbf{x}^*) = b$, and $Dg_{\mathbf{x}}(\mathbf{x}^*) \neq \mathbf{0}$, There is a multiplier λ^* satisfying:

1. $L(\mathbf{x}, \lambda) := f(\mathbf{x}) + \lambda(\bar{b} - g(\mathbf{x}))$
2. $DL_{\mathbf{x},\lambda}(\mathbf{x}^*, \lambda) = \mathbf{0}$
3. $\lambda^*(\bar{b} - g(\mathbf{x}^*)) = 0$
4. $\lambda^* \geq 0$
5. $\bar{b} - g(\mathbf{x}^*) \geq 0$

General Inequality Constraints

Theorem (18.4)

Suppose $f, \mathbf{G}$ *are* C^1 *functions of* n *variables* ($\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^k$)*. Suppose* $\mathbf{x}^* \in \mathbb{R}^n$ *is a local max(min)imizer of* f *on the constraint set defined by the* k *inequalities* $\mathbf{G} \leq \bar{\mathbf{b}}$*. If (1)* k_0 *constraints are binding at* $\mathbf{x}^*$ *and the other* $k - k_0$ *constraints are not binding, and (2)* $rank D\mathbf{G}_{\mathbf{0x}}(\mathbf{x}^*) = k_0$ ($\mathbf{G_0}$*: binding inequality constraints) (NDCQ). Then,* $\exists \lambda^*$ *satisfying:*

1. $L(\mathbf{x}, \lambda) := f(\mathbf{x}) + \lambda(\bar{\mathbf{b}} - \mathbf{G}(\mathbf{x}))$
2. $DL_{\mathbf{x}}(\mathbf{x}^*, \lambda^*) = \mathbf{0}$
3. $\lambda_i^*(\bar{b_i} - G_i(\mathbf{x}^*)) = 0 \quad \forall i = 1, 2, \cdots, k$
4. $\lambda \geq \mathbf{0}$
5. $\bar{\mathbf{b}} - \mathbf{G}(\mathbf{x}^*) \geq \mathbf{0}$

Note: $\lambda := (\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k$. When ith constraint is not binding, $\lambda_i = 0$ (like unconstraint) and when jth constraint is binding, $\lambda_j > 0$ (like equality constraint).

Mixed Constraints

Theorem (18.5-1)

Suppose $f, \mathbf{H}, \mathbf{G}$ *are* C^1 *functions of* n *variables. Suppose* $\mathbf{x} \in \mathbb{R}^n$ *is a local max(min)imizer of* f *on the constraint set defined* m *equalities and* k *inequalities:*

$$\arg\max_{\mathbf{x}} f(\mathbf{x}) \quad s.t. \quad \begin{cases} \mathbf{H}(\mathbf{x}) = \bar{\mathbf{a}} \\ \mathbf{G}(\mathbf{x}) \leq \bar{\mathbf{b}} \end{cases}$$

Assume k_0 *inequality constraints are binding at* $\mathbf{x}^*$ *and the other* $k - k_0$ *inequality constraints are not binding at* $\mathbf{x}^*$*. And suppose that NDCQ is satisfied (*$\mathbf{G_0}$*: binding inequality constraints):*

$$rankD\begin{pmatrix} \mathbf{G_0} \\ \mathbf{H} \end{pmatrix}_{\mathbf{x}}(\mathbf{x}^*) = k_0 + m$$

Mixed Constraints (2)

Theorem (18.5-2)

Then, $\exists \mu^* \in \mathbb{R}^m, \lambda^* \in \mathbb{R}^k$ *satisfying:*

1. $L(\mathbf{x}, \mu, \lambda) := f(\mathbf{x}) + \mu(\bar{\mathbf{a}} - \mathbf{H}(\mathbf{x})) + \lambda(\bar{\mathbf{b}} - \mathbf{G}(\mathbf{x}))$
2. $DL_{\mathbf{x},\mu}(\mathbf{x}^*, \mu^*, \lambda^*) = \mathbf{0}$
3. $\lambda^*(\bar{\mathbf{b}} - \mathbf{G}(\mathbf{x}^*)) = \mathbf{0}$
4. $\lambda^* \geq \mathbf{0}$
5. $\bar{\mathbf{b}} - \mathbf{G}(\mathbf{x}^*) \geq \mathbf{0}$

Note: When minimizing, the only difference is making
$L := f(\mathbf{x}) + \mu(\bar{\mathbf{a}} - \mathbf{H}(\mathbf{x})) - \lambda(\bar{\mathbf{b}} - \mathbf{G}(\mathbf{x}))$ (Equivalent to Th 18.6)

Kuhn-Tucker Formulation

Kuhn-Tucker Formulation

$$\arg\max_{\mathbf{x}} f(\mathbf{x}) \quad s.t. \quad \mathbf{G}(\mathbf{x}) \leq \bar{\mathbf{b}} \quad \wedge \quad \mathbf{x} \geq 0$$

When Considering Kuhn-Tucker Lagrangian

$$\tilde{L} := f(\mathbf{x}) + \lambda(\bar{\mathbf{b}} - \mathbf{G}(\mathbf{x}))$$

FOCs are

- $D\tilde{L}_{\mathbf{x}}(\mathbf{x}^*, \lambda^*) \leq \mathbf{0}$
- $D\tilde{L}_{\lambda}(\mathbf{x}^*, \lambda^*) \geq \mathbf{0}$
- $x_i \frac{\partial \tilde{L}}{\partial x_i}(\mathbf{x}^*, \lambda^*) = 0$ for all i
- $\lambda_i \frac{\partial \tilde{L}}{\partial \lambda_i}(\mathbf{x}^*, \lambda^*) = 0$ for all i

Note: Above lagrangian does not contain n inequality constraints $(\mathbf{x} \geq 0)$