A Cookbook Approach to Constrained Optimization

Jaepil Lee

2022

1 Set up the optimization problem. (WLOG, I assume $\boldsymbol{x} \in U$ where U is open and $U \subseteq \mathbb{R}^n$.)

`MAX` : $\max_{\boldsymbol{x}\in U} f(\boldsymbol{x})$ s.t. $g_i(\boldsymbol{x}) \geq 0$, $h_j(\boldsymbol{x}) = 0 \ \forall i = 1,\dots,n_i, j = 1,\dots,n_e$

`MIN` : $\min_{\boldsymbol{x}\in U} f(\boldsymbol{x})$ s.t. $g_i(\boldsymbol{x}) \geq 0$, $h_j(\boldsymbol{x}) = 0 \ \forall i = 1,\dots,n_i, j = 1,\dots,n_e$

2 Does there exist a solution? Here are some facts you want to use:

A. Weierstrass: $f \in C^0$ on a compact domain $\implies \exists$ max and min on the compact domain.

B. Concave function on a convex set $\implies \exists$ max of the function on the convex set
Convex function on a convex set $\implies \exists$ min of the function on the convex set
NOTE: $f : S \to \mathbb{R}$ defined on a convex set S is ...

  a. concave if $\forall x, x' \in S$ and $\forall \lambda \in (0,1)$, $f((1-\lambda)x + \lambda x') \geq (1-\lambda)f(x) + \lambda f(x')$
  b. convex if $\forall x, x' \in S$ and $\forall \lambda \in (0,1)$, $f((1-\lambda)x + \lambda x') \leq (1-\lambda)f(x) + \lambda f(x')$

NOTE: If $f \in C^2$, then it may be easier to check the Hessian and its Leading Principal Minors (LPM):

  a. Hessian is psd (all LPMs ≥ 0)/nsd (LPMs alternate signs; $\leq 0, \geq 0, \dots$) $\implies f$ is convex/concave.
  b. Hessian is pd (all LPMs > 0)/nd (LPMs alternate signs; $< 0, > 0, \dots$) $\implies f$ is strictly convex/concave.

C. Quasiconcave function on a convex set $\implies \exists$ max of the function on a convex set
Quasiconvex function on a convex set $\implies \exists$ min of the function on a convex set
NOTE: $f : S \to \mathbb{R}$ defined on a convex set S is ...

  a. quasiconcave if $C_a^+ := \{x \in S \mid f(x) \geq a\}$ is convex for every a.
  $\iff \forall x, x' \in S$ and $\forall \lambda \in [0,1]$, $f(x) \geq f(x') \implies f((1-\lambda)x + \lambda x') \geq f(x')$
  $\iff \forall x, x' \in S$ and $\forall \lambda \in [0,1]$, $f((1-\lambda)x + \lambda x') \geq \min\{f(x), f(x')\}$
  b. quasiconvex if $C_a^- := \{x \in S \mid f(x) \leq a\}$ is convex for every a.
  $\iff \forall x, x' \in S$ and $\forall \lambda \in [0,1]$, $f(x) \geq f(x') \implies f((1-\lambda)x + \lambda x') \leq f(x)$
  $\iff \forall x, x' \in S$ and $\forall \lambda \in [0,1]$, $f((1-\lambda)x + \lambda x') \leq \max\{f(x), f(x')\}$

D. Now, is there a solution to the problem?

  a. If YES or NOT SURE (that's okay), go to Step 3.
  b. If NO, stop and conclude 'there is no solution to the problem'.

3 Shortcuts: before we rely on the KKT Theorem (complementary slackness, etc.), let's think for a bit:

A. For `MAX`, is the objective function f (quasi)concave on a convex set?
For `MIN`, is the objective function f (quasi)convex on a convex set?

  a. If YES, then FOC is sufficient to find the solution (NO NEED to check LICQ in Step 5 nor Step 6).
  b. If NO, then we do need to consider all complementary slackness conditions in Step OWEVER...

B. In economics, we still have some shortcuts:

  a. Is the budget constraint binding (due to monotonicity, local non-satiation, etc.)?
    i) If YES, then the inequality constraint becomes equality constraint (from $\boldsymbol{p} \cdot \boldsymbol{x} \leq w$ to $\boldsymbol{p} \cdot \boldsymbol{x} = w$).
    ii) If NO, then check for an interior solution (i.e. a bliss point).
  b. If the budget constraint is binding, do you suspect that the corner points are the solution?
    i) If YES, then claim that the corner solutions are our (local) maximizers (with careful reasoning). Remember to check if other points on the boundary are also (local) solutions under certain conditions. Usually this happens when the goods are perfect substitutes.
    ii) If NO, then analytically solve for the solution.

4 Are the objective function and the constraint functions differentiable wrt $\boldsymbol{x}$?

A. If YES, go to Step 5. Before you go to Step 5, think once more about the behavior of the objective function:
   a. Would the problem have an interior solution (bliss point)?
   b. If you think the solution is on the boundary, where do you think the solution is? Is it a corner solution? Is it located somewhere else on the boundary?

B. If NO, then we should take an 'analytical' approach. The ε-argument is useful here. Some examples are:
   a. min$\{x_1, x_2\} \implies x_1^* = x_2^*$
   b. max$\{x_1, x_2\} \implies (x_1^*, x_2^*) = (\frac{w}{p_1}, 0)$ or $(0, \frac{w}{p_2})$

5 Does the Constraint Qualification hold?

A. Slater condition (for convex problems; ex. f, g are convex and h is affine[1]):
   $\exists\, \boldsymbol{x} \in \mathbb{R}^n$ such that $g_i < 0$, $h_j = 0\ \forall\, i, j$. This is known to be a sufficient condition for the strong duality.

B. Linear Independence Constraint Qualification (LICQ)
   a. Find the gradients of the constraints: Dg_i, Dh_j wrt $\boldsymbol{x}$
   b. Determine all those points in the domain where the constraint qualification FAILS:
      i) Check linear independence of every combination of gradient of constraints. If there are l inequality constraints and k equality constraints, we have 2^l combinations to consider where the combinations include all the k equality constraints.
      ex) Consider l=3, k=2: $g_1 \geq 0$, $g_2 \geq 0$, $g_3 \geq 0$, $h_1 = 0$, $h_2 = 0$. The combinations to check for the linear independence are (notice that we have $\{Dh_1, Dh_2\}$ in all combinations):

$$\begin{aligned}
&\{Dh_1, Dh_2\},\\
&\{Dh_1, Dh_2, Dg_1\}, \{Dh_1, Dh_2, Dg_2\}, \{Dh_1, Dh_2, Dg_3\},\\
&\{Dh_1, Dh_2, Dg_1, Dg_2\}, \{Dh_1, Dh_2, Dg_2, Dg_3\}, \{Dh_1, Dh_2, Dg_3, Dg_1\},\\
&\{Dh_1, Dh_2, Dg_1, Dg_2, Dg_3\}
\end{aligned}$$

      ii) Equivalently, MWG checks linear independence of the matrix of the gradient of all binding constraints. Let $\boldsymbol{E}$ denote the set of equality constraints (note that we have $n_e = 5$ equality constraints!). The matrix of interest is $D_{\mathbf{x}}\boldsymbol{E}$ where:

$$D_{\mathbf{x}}\boldsymbol{E} := \begin{bmatrix} D_{\boldsymbol{x}'} g_1 \\ \vdots \\ D_{\boldsymbol{x}'} g_{n_i} \\ D_{\boldsymbol{x}'} h_1 \\ \vdots \\ D_{\boldsymbol{x}'} h_{n_i} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial g_3}{\partial x_1} & \frac{\partial g_3}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_n} \end{bmatrix}$$

      If $\exists\, \boldsymbol{x}^* \in \mathbb{R}^n$ such that $\boldsymbol{v}' D_{\boldsymbol{x}} \boldsymbol{E}(\boldsymbol{x}^*) = \boldsymbol{0},\ \forall\, \boldsymbol{v} = \boldsymbol{0}$ and $\boldsymbol{v} \in \mathbb{R}^e$, they are the points where LICQ FAILS.
      iii) If there are points where qualification constraint fails, evaluate them and compare with the local maxima/minima we find via KKT theorem later. We do this because the KKT theorem only covers the points where qualification constraint holds.

[1] https://en.wikipedia.org/wiki/Karush%E2%80%93Kuhn%E2%80%93Tucker_conditions

6 KKT Theorem (for simplicity, I denote $h_j \geq 0,\ j \in \{1, 2, \ldots, l\}$ for both equality and inequality constraints.)

A. The Lagrangian function is:
MAX: $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{j=1}^{l} \lambda_j h_j(\boldsymbol{x})$
MIN: $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{j=1}^{l} \lambda_j h_j(\boldsymbol{x})$
Note that the sign in the middle of the Lagrangian changes. This also appears in the theorem below (***).

B. Let $f : U \to \mathbb{R}$ and $h_j : \mathbb{R}^n \to \mathbb{R},\ j = 1, \ldots, l$ be C^1 functions where U is some open subset of $\mathbb{R}^n$.
Let $E \subseteq \{h_1, \ldots, h_l\}$ be the set of effective constraints at $\boldsymbol{x}^*$. Suppose that:

a. $\boldsymbol{x}^*$ is a local maximum/minimum of f on the set $\mathcal{D} = U \cap \{\boldsymbol{x} \in \mathbb{R}^n \,|\, h_j(\boldsymbol{x}) \geq 0,\ \forall j \in \{1, \ldots, l\}\}$ and
b. the vectors of gradients $\{Dh_\tau(\boldsymbol{x}^*) \,|\, \tau \in E\}$ are (linearly) independent.

Then $\exists \lambda_j^* \in \mathbb{R}^n\ \forall j \in \{1, \ldots, l\}$ such that:

a. First Order Condition:
MAX: $D_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}) = Df(\boldsymbol{x}^*) + \sum_{j=1}^{l} \lambda_j h_j(\boldsymbol{x}^*) = \mathbf{0}$ ***
MIN: $D_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}) = Df(\boldsymbol{x}^*) - \sum_{j=1}^{l} \lambda_j h_j(\boldsymbol{x}^*) = \mathbf{0}$ ***
b. Complementary Slackness:
$\lambda_j^* \geq 0$, $h_j(\boldsymbol{x}^*) \geq 0$, and $\lambda_j^* h_j(\boldsymbol{x}^*) = 0$, $\forall j \in \{1, \ldots, l\}$.

7 Compare the points where the constraint qualification fails and the (local) optimizers from the KKT Theorem.

A. Did you show that the maximum/minimum exists from Step 2?

a. If YES, then go to Step 7B.
b. If NOT SURE, then go to Step 7C.

B. Did you find any candidate points from Steps 5 and 6?

a. If YES, evaluate the objective function on those points. The one that gives you the largest/smallest value is your solution (I hope there are only a few candidate points...).
b. If NO, something is wrong. Go back to Step 3 and think about whether there is any economic intuition you are missing. You should also go to Steps 5 and 6 to find mistakes.

C. Did you find any candidate points from Steps 5 and 6?

a. If YES, you may have a global solution among the candidate points OR all of them could just be local solutions. Give yourself a reason why they are or they are not global solutions.
b. If NO, then it is possible that there is no solution. Why does the function have no solution? Is the function unbounded? If you have no particular reason for this, we cannot rule out that you missed the candidate points. Go back to Steps 5 and 6 and check once more just to be safe.

Reference

- Chiang, Alpha C. Fundamental Methods of Mathematical Economics. 1984.
- Mas-Colell, Andreu, Michael Dennis Whinston, and Jerry R. Green. Microeconomic theory. Vol. 1. New York: Oxford university press, 1995.
- Nava, F. EC400 Mathematics for Microeconomics, LSE Course pack, 2016.
- Osborne, Martin J. Mathematical Methods for Economics Theory. Retrieved from `https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/toc` on September 6, 2019.
- Sundaram, Rangarajan K. A First Course in Optimization Theory. Cambridge university press, 1996.