Kuhn-Tucker Theorem, Stackelberg Games, and An Extended Mathematical Programming (EMP) Framework in Trade (Part A)

Anton Yang

Lecture developed for graduate international trade curriculum
Links to download codes using the Baron solver and the EMP

Purdue University

Exercises

- We study a simple application of EMP framework, e.g., Ferris *et al.* (Computers and Chemical Engineering, 2009).
- In so doing, we apply the Kuhn-Tucker Theorem in a bilevel Stackelberg game.
- We solve a two-layer model computationally.
- Then we reformulate this model using a mathematical program with complementarity constraints (MPCC).
- Finally, we solve this model again directly using the EMP.

Task

1. We extend this modeling and computational framework in trade with multiple optimizers absent an assurance that the problem is a convex global optimization program;

2. We need to ask: whether and when the EMP can be a better approach than the MPEC (discussed later) in solving a large and complex general equilibrium model.

Kuhn-Tucker Theorem and an LP Example

- Kuhn-Tucker theorem or Karush–Kuhn–Tucker (KKT) theorem is that if regularity conditions hold for a LP/NLP then the theorem provides first-order necessary conditions for a solution in this program to be optimal.

- We will see that the necessary and sufficient optimality conditions for the following *canonical* LP program

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.:} \quad & Ax \geq b, \quad x \geq 0. \end{aligned}$$

  (think that firms minimize some costs, subject to some linear technology and non-negativity constraints)

Kuhn-Tucker Theorem and an LP Example

- The complementarity relationships written in "$\perp$" are:

$$
\begin{aligned}
0 \leq c - A^T\lambda \quad &\perp \quad x \geq 0, \\
0 \leq Ax - b \quad &\perp \quad \lambda \geq 0.
\end{aligned}
$$

- We used symbol "$\perp$" or the "perp" in a previous lecture (i.e., solution to the structural Hicksian import demand).

- One way to interpret this is that the product of the non-zero element(s) on each side of "$\perp$" is zero, or has an either or relationship.

- See also Choi (Computational Economics, 2015), p.308.

Kuhn-Tucker Theorem and an LP Example

- Recall that the relevant Lagrange function is

$$\mathcal{L} = c^T x + \lambda_1(-Ax + b) + \lambda_2(-x).$$

- The KKT conditions are

$$\begin{aligned}
&\nabla \mathcal{L}(x) = c - A\lambda_1 - \lambda_2 = 0; \quad (\text{stationary}) \\
&g_1 = -AX + b \leq 0, \quad g_2 = -x \leq 0; \quad (\text{primal feasibility}) \\
&\lambda_1 \geq 0, \quad \lambda_2 \geq 0; \quad (\text{dual feasibility}) \\
&\lambda_1(Ax - b) = 0, \quad \lambda_2(x) = 0; \quad (\text{complementary slackness})
\end{aligned}$$

- which is always sufficient, necessary under strong duality (Slater's condition) $\longrightarrow$ assurance of a convex program?

A Bilevel Example

- Upper level (leader) and lower level (follower): the upper level leader's decision is made upon the optimal solution to the lower level follower as follows:

$$\min_{x} \quad F_1 = -100x - 1000y_1$$

where y_1, y_2 solve

$$\begin{aligned} \min_{y_1,y_2} \quad & F_2 = -y_1 - y_2 \\ s.t. \quad & -x - y_1 + y_2 \geq -1 \\ & -2y_1 - y_2 \geq -1 \\ & x, y_1, y_2 \geq 0 \end{aligned}$$

- This simplest form of bilevel programming can be viewed as a *subset* of the more realistic Stackelberg game, where there is one leader who faces many followers (who theoretically behave in a Nash equilibrium).

A Bilevel Example

- One way to solve this problem is to transform the bilevel program to a single-level problem.
- Starting from the lower level, the Lagrange function is

$$\mathcal{L} = -y_1 - y_2 - \lambda_1(-x - y_1 + y_2 + 1) - \lambda_2(-2y_1 - y_2 + 1)$$

- Note that the standard form of $\mathbf{g} \geq 0$ for a minimization program involves signs that are negative.

A Bilevel Example

- By the Kuhn-Tucker theorem:

$$
\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial y_1} = -1 + \lambda_1 + 2\lambda_2 \geq 0, \quad y_1 \geq 0 \\
&y_1(-1 + \lambda_1 + 2\lambda_2) = 0 \\
&\frac{\partial \mathcal{L}}{\partial y_2} = -1 - \lambda_1 + \lambda_2 \geq 0, \quad y_2 \geq 0 \\
&y_2(-1 - \lambda_1 + \lambda_2) = 0 \\
&\frac{\partial \mathcal{L}}{\partial \lambda_1} = x + y_1 - y_2 - 1 \leq 0, \quad \lambda_1 \geq 0 \\
&\lambda_1(-x - y_1 + y_2 + 1) = 0 \\
&\frac{\partial \mathcal{L}}{\partial \lambda_2} = 2y_1 + y_2 - 1 \leq 0, \quad \lambda_2 \geq 0 \\
&\lambda_2^t(-2y_1 - y_2 + 1) = 0
\end{aligned}
$$

A Bilevel Example

- The single-level problem now becomes a leader's problem

$$
\begin{aligned}
\min_{x, y_1, y_2, \lambda_1, \lambda_2} \quad & F_1 = -100x - 1000y_1 \\
& s.t. \quad -1 + \lambda_1 + 2\lambda_2 \geq 0 \\
& y_1(-1 + \lambda_1 + 2\lambda_2) = 0 \\
& -1 - \lambda_1 + \lambda_2 \geq 0 \\
& y_2(-1 - \lambda_1 + \lambda_2) = 0 \\
& -x - y_1 + y_2 + 1 \geq 0 \\
& \lambda_1^t(-x - y_1 + y_2 + 1) = 0 \\
& -2y_1 - y_2 + 1 \geq 0 \\
& \lambda_2(-2y_1 - y_2 + 1) = 0 \\
& y_1, y_2, \lambda_1, \lambda_2 \geq 0
\end{aligned}
$$

NLP and MPEC

- So this program can be solved in GAMS using NLP, e.g., Branch-And-Reduce Optimization Navigator (BARON). For now, we can just keep in mind that BARON can deal with nonconvex objectives and constraints. See Baron code.
- One can also choose to reformulate the problem as an MPCC using "perp"'s (the complementarity relationships) and lower level's optimality conditions (Ferris *et al., 2009*).
- It can be solved using nonlinear programming with equilibrium constraints (NLPEC) which solves Mathematical Programming with Equilibrium Constraints (MPEC) and Mixed Complementarity Problems (MCP) with the reformulated complementarity relationships.

MPEC Applications in Trade and EMP

- MPEC is an ideal program to solve optimizations of engineering and complex economic problems such as general equilibrium gravity models of trade, models of spatial economy (see Balistreri and Hillberry, 2007; Balistreri, Hillberry and Rutherford, 2011; Tan, 2012; Yang, 2019).

- There are several disadvantages of solving an NLPEC, e.g.,

  (1) The solutions or sufficient solutions in the optimality conditions associated with the MCP are not guaranteed;

  (2) It is possible that the MPCC will only find local solutions but not global solutions (Ferris *et al.*).

EMP

- The EMP, on the other hand, is a more advanced algebraic modeling language which facilitates the reformulation of the models into a new problem (that can be used) to tackle structural identification problems; it allows a mature solver algorithm and automatically generates an MCP.

- See a recent *spatial* households problem: Ferris, Rutherford and van Nieuwkoop, 2019.

- And since an MCP is automatically formulated from the NLP/LP, we only need to write out the original bilevel problem and solve it directly via the EMP annotations.

- See sample EMP code.