IEOR 290A – Lecture 37
Estimating an Individual Utility

1 Utility Maximizing Agent

Recall the following abstract model: Suppose that an agent makes decisions by solving the following optimization problem:

$$x_i^* = \arg\max\{J(x, u_i) \mid x \in \mathcal{X}(u_i)\},$$

where $u_i \in \mathbb{R}^q$ are inputs, $x_i^* \in \mathbb{R}^d$ are decisions, $J(x, u_i)$ is the utility function of the agent, and $\mathcal{X}(u_i)$ is a bounded set (that depends on u_i). In this model, we observe (u_i, x_i^*) for $i = 1, \ldots, n$ and would like to infer the function $J(x, u_i)$.

To make this model more concrete, we will specify a specific instantiation of this problem. In particular, suppose that

- The constraint set is described by linear equality and inequality constraints:

  $$\mathcal{X}(u) = \{x : Ax + Bu_i \leq c, Fx + Gu_i = h\},$$

  where (A, b) and (F, h) are suitably defined matrices and vectors.

- Assume that we have a parametrization of the utility function, that is we have $\phi(x, u; \beta)$ and a bounded set Γ such that there exists $\beta_0 \in \Gamma$ with $J(x, u) = \phi(x, u; \beta_0)$.

Though these two conditions make the problem more specific, we will still impose additional conditions on the model formulation to make the problem computationally tractable.

2 Key Technical Difficulty

Recall the feasibility problem formulation of the inverse decision-making problem for this single utility maximizing agent model:

$$\begin{aligned}
\hat{\beta} = \arg\min_{\beta}\ & 0 \\
\text{s.t. } & x_i^* \in \arg\max_x\{\phi(x, u_i; \beta) \mid Ax + Bu_i \leq c, Fx + Gu_i = h\} \\
& \beta \in \Gamma.
\end{aligned}$$

This feasibility problem is difficult to solve because it has an atypical constraint: The constraint that x_i^* be the minimizer to some optimization problem cannot be directly handled by nonlinear programming techniques. There are two reasons that this constraint presents challenges:

1. Depending on the value of β there may be zero, one, or multiple maximizers. This means that in general we must treat the function

$$P(u_i, \beta) = \arg\max_x \{\phi(x, u_i; \beta) \mid Ax + Bu_i \leq c, Fx + Gu_i = h\}$$

as a multi-valued function.

2. The function $P(u_i, \beta)$ has a complex form, because it is defined as a set of maximizers. This means that in general we cannot even hope for continuity of $P(u_i, \beta)$ (cf. the Berge Maximum Theorem, which says that for continuous ϕ we can only expect upper-hemicontinuity of $P(u_i, \beta)$), much less differentiablity.

3 Tractable Formulation

Since $P(u_i, \beta)$ is a multi-valued function, we can make the problem more tractable by imposing additional conditions on our model so that instead $P(u_i, \beta)$ is a single-valued (and hence continuous by the Berge Maximum Theorem) function. In particular, suppose that for all fixed values of $\beta \in \Gamma$ the function $\phi(x, u_i; \beta)$ is strictly concave in (x, u_i). Then the corresponding optimization problem has a single maximizer, and so this additional condition fixes our first difficulty.

The second difficulty regarding the complex form of $P(u_i, \beta)$ still remains. However, since our constraints are linear, we have linearity constraint qualification, and so the unique maximizer $x_i^* = P(u_i, \beta)$ satisfies the KKT conditions: There exist row-vectors λ_i and μ_i such that

$$\begin{aligned}
-\nabla_x \phi(x_i^*, u_i; \beta) + \lambda_i A + \mu_i F &= 0 \\
Ax_i^* + Bu_i &\leq c \\
Fx_i^* + Gu_i &= h \\
\lambda_i^j &\geq 0 \\
\lambda_i^j &= 0 \text{ if } A_j x_i^* + B_j u_i < c_j,
\end{aligned}$$

where A_j, B_j, c_j denote the j-th row of A, B, c respectively. As a result, we can now pose our feasibility problem as

$$\begin{aligned}
\hat{\beta} = \arg\min_\beta \ & 0 \\
\text{s.t. } & -\nabla_x \phi(x_i^*, u_i; \beta) + \lambda_i A + \mu_i F = 0 \\
& \lambda_i^j \geq 0 \\
& \lambda_i^j = 0 \text{ if } A_j x_i^* + B_j u_i < c_j \\
& \beta \in \Gamma.
\end{aligned}$$

Note that because (u_i, x_i^*) are measured, they are constant in our feasibility formulation and in the KKT conditions. Therefore, the conditional statement "if $A_j x_i^* + B_j u_i < c_j$"

is computed before we solve the feasibility problem. In other words, we decide to either include or exclude the constraint $\lambda_i^j = 0$ in our feasibility problem based on a precomputed conditional.

This problem can still be difficult to solve, because this reformulated problem may not be convex. Consider the constraint

$$-\nabla_x \phi(x_i^*, u_i; \beta) + \lambda_i A + \mu_i F = 0,$$

and note that it is an equality constraint. However, a standard result is that an equality constraint $Q(\beta)$ is convex if and only Q is an affine function (meaning that it can be written as $Q = M\beta + k$ where M is a matrix and k is a constant vector). As a result, our feasibility problem to estimate the parameters β of our utility function is convex if and only if $Q(\beta) = -\nabla_x \phi(x_i^*, u_i; \beta)$ is an affine function. Stated in another way, our formulation is convex if and only if the gradient of ϕ with respect to x is affine in β when the gradient is evaluated at x_i^* and u_i.