‘MATH+ECON+CODE’ MASTERCLASS ON MATCHING MODELS, OPTIMAL TRANSPORT AND APPLICATIONS

Alfred Galichon (New York University)

Spring 2018
Day 4, January 18 2018: “Multinomial choice”
Block 12. Dynamic discrete choice

Learning objectives: block 12

- Finite-horizon Rust's model
- identification and estimation
- normalization issues

- Rust (1987), "Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. *Econometrica*.
- Chiong, Galichon and Shum (2016). Duality in discrete choice models. *Quantitative economics*.

Section 1

Theory

- Recall the dynamic programming model seen in block 3. The setting is the same: there are n_x units (buses) in state x at the initial period $(t = 1)$; at each period, one must choose for each unit some alternative $y \in \mathcal{Y}$; the probability of transiting to state x' at period $t + t$ conditional on being in state x and choosing alternative y at time t is $P^t_{x'|xy}$.
- The difference with the setting seen in block 3 is that, following Rust, the utility associated with choosing y in state x at t is no longer deterministic, but includes an additional random term $\varepsilon_y \sim \mathbf{P}_{xt}$, so it is

$$u^t_{xy} + \varepsilon_y.$$

The stochastic structure is such that $x_t, (x_t, \varepsilon), (x_t, y_t)$ is a Markov chain – which rules out persistent shocks, i.e. there cannot be correlaton between ε_t and ε_{t+1} conditional on (x_{t+1}, y_{t+1}).

- As a result of the random utility term, Bellman's equation becomes

$$
\begin{aligned}
U_x^t &= \mathbb{E}_{\mathbf{P}_{xt}}\left[\max_{y\in\mathcal{Y}}\left\{u_{xy}^t+\sum_{x'}U_{x'}^{t+1}P_{x'|xy}+\varepsilon_y\right\}\right]\\
&= G_{xt}(u_{x.}^t+\sum_{x'}U_{x'}^{t+1}P_{x'|x.}).
\end{aligned}
$$

- Set $W_{xt}\left(U\right)=G_{x(t-1)}(u_{x.}^{t-1}+\sum_{x'}U_{x'}^tP_{x'|x.})$ for $1<t<T$, the equation becomes

$$
U_x^{t-1}=W_{xt}\left(U^t\right).
$$

- Note that

$$
\frac{\partial W_{xt}\left(U\right)}{\partial U_{x'}}=\sum_y P_{x'|xy}\sigma_{x(t-1),y}
$$

is the conditional probability of a transition to x' given being at x at time $t-1$, denoted $\mu_{x'|x}^{t-t}$.

- In the logit case, one has

$$U_x^t = \log \sum_{y \in \mathcal{Y}} \exp \left(u_{xy}^t + \sum_{x'} U_{x'}^{t+1} P_{x'|xy} \right)$$

- Setting $V_x^t = \exp\left(U_x^t\right)$ and $v_{xy}^t = \exp\left(v_{xy}^t\right)$, this becomes an algebraic expression

$$V_x^t = \sum_{y \in \mathcal{Y}} v_{xy}^t \prod_{x' \in \mathcal{X}} \left(V_x^t\right)^{P_{x'|xy}}.$$

- The dual problem can be expressed as:

$$\begin{aligned} &\min_{U_x^t,} \sum_{x\in\mathcal{X}} n_x U_x^1 \\ s.t.\ &U_x^{t-1} = W_{xt}(U^t)\ 1 < t \leq T \\ &U_x^T = G_{xT}(u_{x.}^T) \end{aligned} \tag{1}$$

- In the logit case, with $V_x^t = \exp\left(U_x^t\right)$ and $v_{xy}^t = \exp\left(v_{xy}^t\right)$, one has

$$\begin{aligned} &\min_{U_x^t,\ t\in\mathcal{T},\ x\in\mathcal{X}} \sum_{x\in\mathcal{X}} n_x \log V_x^1 \\ s.t.\ &V_x^t = \sum_{y\in\mathcal{Y}} v_{xy}^t \prod_{x'\in\mathcal{X}} \left(V_x^t\right)^{P_{x'|xy}},\ t < T \\ &V_x^T = \sum_{y\in\mathcal{Y}} v_{xy}^t. \end{aligned}$$

- Set n_x^t the Lagrange multipliers associated with the constraints. First order conditions in the dual problem yield

$$n_x = n_x^1$$
$$n_x^t = \sum_{x'} \frac{\partial W_{x(t-1)}}{\partial U_{x'}^{t-1}} n_{x'}^{t-1}, 1 < t \leq T$$

- The second line are Kolmogorov-forward equations (forward propagation of mass)

$$\sum_{x'} \mu_{x'|x}^{t-t} n_{x'}^{t-1} = n_x^t.$$

- Let $W_t\left(U^t; n^{t-1}\right) = \sum_x n_x^{t-1} W_{xt}(U_x^t)$, and let $W_t^*\left(n^t; n^{t-1}\right)$ be its Legendre transform with respect its first variable.
- **Theorem**: the value of the primal problem is

$$\max_{n^t} \left\{ \sum_{x\in\mathcal{X}} n_x^T G_{xt}(u_{x.}^T) - \sum_{\substack{x\in\mathcal{X} \\ 1<t\leq T}} W_t^*\left(n^t; n^{t-1}\right) \right\}$$

$$s.t.\ n_x^1 = n_x$$

- Start from the dual

$$\begin{aligned} \min_{U_x^t} \sum_{x \in \mathcal{X}} n_x U_x^1 & \\ s.t.\ U_x^{t-1} = W_{xt}(U^t) & \ 1 < t \leq T \\ U_x^T = G_{xT}(u_{x.}^T) & \end{aligned} \tag{2}$$

- Write the saddlepoint formulation

$$\min_{U^t} \max_{n^t} \left\{ \begin{array}{c} \sum_{x \in \mathcal{X}} n_x U_x^1 - \sum_{x, 1 \leq t \leq T} n_x^t U_x^t \\ + \sum_{x, 1 < t \leq T} n_x^{t-1} W_{xt}(U^t) \\ + \sum_x n_x^T G_{xT}(u_{x.}^T) \end{array} \right\}$$

- Saddlepoint rewrites

$$\max_{n^t} \min_{U^t} \left\{ \begin{array}{c} \sum_x n_x^T G_{xt}(u_{x.}^T) + \sum_x \left(n_x - n_x^1 \right) U_x^1 \\ + \sum_{1<t\leq T} n_x^{t-1} W_{xt}(U_x^t) - n_x^t U_x^t \end{array} \right\}$$

- Recall that $W_t\left(U^t; n^{t-1}\right) = \sum_x n_x^{t-1} W_{xt}(U_x^t)$, and $W_t^*\left(n^t; n^{t-1}\right)$ be its Legendre transform with respect its first variable, one has

$$\max_{n_x^t, t\geq 1} \left\{ \sum_{t<T} n_x^T G_{xt}(u_{x.}^T) - \sum_{1<t\leq T} W_t^*\left(n^t; n^{t-1}\right) \right\}$$
$$s.t.\ n_x^1 = n_x$$

QED.

- Recall

$$\max_{n_x^t, t \geq 1} \left\{ \sum_{t<T} n_x^T G_{xt}(u_{x.}^T) - \sum_{1<t\leq T} W_t^* \left(n^t; n^{t-1} \right) \right\}$$
$$s.t. \; n_x^1 = n_x$$

- For $1 \leq t < T$, one has

$$\frac{\partial W_t^*}{\partial n_x^t} \left(n^t; n^{t-1} \right) + \frac{\partial W_{t+1}^*}{\partial n_x^t} \left(n^{t+1}; n^t \right) = 0,$$

and note that

$$\frac{\partial W_t^*}{\partial n_x^t} = U_x^t \text{ and } \frac{\partial W_{t+1}^*}{\partial n_x^t} \left(n^{t+1}; n^t \right) = -W_{xt+1}(U^{t+1})$$

hence the first order condition recovers the Bellman equation.

- One has

$$W_t\left(U; n^{t-1}\right) = \sum n_x^{t-1} \log \sum_{y \in \mathcal{Y}} \exp\left(u_{xy}^{t-1} + \sum_{x'} U_{x'}^t P_{x'|xy}\right)$$

and thus

$$\begin{aligned} & W_t\left(n^t; n^{t-1}\right) \\ & = \max_U \left\{ \sum_x n_x^t U_x - \sum_x n_x^{t-1} \log \sum_{y \in \mathcal{Y}} \exp\left(u_{xy}^{t-1} + \sum_{x'} U_{x'}^t P_{x'|xy}\right)\right\} \end{aligned}$$

- Sadly, no closed-form formula.

- Rust studies the infinite-horizon version of the problem, in which case u_{xy}^t does not depend on t, and, if $\beta > 0$ is a discount factor, then the intertemporal utility is given by the set of equations

$$U_x = W_x(\beta U),$$

where $W_x(\beta U) = \mathbb{E}\left[\max_y \left\{u_{xy} + \sum_{x'} \beta U_{x'} P_{x'|xy} + \varepsilon_y\right\}\right]$.

- It's possible to show that $(U_x) \to (W_x(\beta U))$ is a contraction mapping, so the above equation has a unique solution.

- Let
$$w_{xy} = u_{xy} + \sum_{x'} \beta U_{x'} P_{x'|xy}$$
We have
$$\begin{cases} w_{xy} = \ln \pi_{y|x} + a_x \\ U_x = \log \sum_y \exp w_{xy} \end{cases}$$

- If we choose the normalization $u_{x0} = 0$, we have

$$w_{x0} = \sum_{x'} \beta U_{x'} P_{x'|x0}$$

and

$$w_{xy} - w_{x0} = \ln \frac{\pi_{y|x}}{\pi_{0|x}}$$

- Hence

$$U_x = w_{x0} - \log \pi_{0|x}$$

- Thus

$$U_x = \sum_{x'} \beta U_{x'} P_{x'|x0} - \log \pi_{0|x} \tag{3}$$

- Let L be the column vector whose general term is $\left(\log \pi_{y|0}\right)_{x\in\mathcal{X}}$, let U be the column vector whose general term is $(U_x)_{x\in\mathcal{X}}$, and let Π be the $|X| \times |X|$ matrix whose general term is

$$\Pi_{xx'} = Pr\left(x_{t+1} = x' | x_t = x, y = 0\right).$$

Equation (3), rewritten in matrix notation, is

$$L = (\beta\Pi - I)\, U$$

- It is possible to show that for $\beta < 1$, matrix $I - \beta\Pi^0$ is invertible. Thus Equation (3) becomes

$$U = (\beta\Pi - I)^{-1} L. \tag{4}$$

- Therefore U_x is identified from data.

- It follows that all the remaining quantities are also identified by

$$
\begin{aligned}
w_{x0} &= U_x + \log \pi_{0|x} \\
w_{xy} &= w_{x0} + \log \frac{\pi_{y|x}}{\pi_{0|x}} \\
u_{xy} &= w_{xy} - \beta \mathbb{E}\left[U_{x'} | x, y\right]
\end{aligned}
$$