Efficient counterfactuals in a semiparametric multinomial choice model: a note on Allen and Rehbeck (2019) and Chiong et al. (2017)

Grigory Franguridi*

September 3, 2020

Abstract

I develop a computationally attractive procedure of calculating sharp bounds on counterfactual demand for a perturbed utility model of Allen and Rehbeck (2019) using the framework of Chiong et al. (2017). The inequality system defining the counterfactual bounds is indexed by cycles in the space of observed markets $1, \ldots, M$. Obtaining the "fully efficient" inequality system exploiting cycles of all possible lengths $K = 1, \ldots, M$ can be reduced to finding (the length of) the shortest path between every pair of vertices in a complete bidirected weighted graph on M vertices – the problem that can be solved using Floyd–Warshall algorithm with computational complexity $O\left(M^3\right)$, which takes only seconds to run even for thousands of markets. Monte Carlo simulations show that the gain from using the fully efficient inequalities exists but is very small.

1 The Algorithm

The family of inequalities for a linear program in Chiong et al. (2017) take the form

$$\begin{aligned}\left(\boldsymbol{\delta}^{l_1} - \boldsymbol{\delta}^{M+1}\right)' \mathbf{s}^{M+1} \leq -1'(D_1 \circ S_1)1 &= \sum_{k=1}^{K-1} \left(\boldsymbol{\delta}^{l_k} - \boldsymbol{\delta}^{l_{k+1}}\right)' \mathbf{s}^{l_k} \\ &= \sum_{k=1}^{K-2} \left(\boldsymbol{\delta}^{l_k} - \boldsymbol{\delta}^{l_{k+1}}\right)' \mathbf{s}^{l_k} + \left(\boldsymbol{\delta}^{l_{K-1}} - \boldsymbol{\delta}^{M+1}\right)' \mathbf{s}^{l_{K-1}} \qquad (1)\end{aligned}$$

for every cycle $l_1, \ldots, l_K, l_1$ from $\mathcal{M} = \{1, \ldots, M\}$ with $l_K = M+1$. Here $\boldsymbol{\delta}^m \in \mathbb{R}^J$ and $\mathbf{s}^m \in \Delta_J$ are given mean utility vector and market share vector for market $m \in \mathcal{M}$, $\boldsymbol{\delta}^{M+1}$ is a given mean utility for the counterfactual market $M+1$ and $\mathbf{s}^{M+1}$ is a variable of choice. Denote the inequality (1) by $\text{ineq}(l_1, \ldots, l_{K-1})$.

The number of possible cycles in a complete graph is exponential in M and is of order 10^{20} already for $M = 22$[1]. Therefore, a simple enumeration is computationally infeasible and Chiong et al. (2017) resort to using a (tiny) subset of inequalities (1) for cycles of length $K = 2$. I will show that enumeration is not necessary for this problem and that the full system of inequalities is easy to construct using graph optimization.

Define G to be a complete bidirected weighted graph on vertices $\mathcal{M}$ with weight $w_{ij} = \left(\boldsymbol{\delta}^i - \boldsymbol{\delta}^j\right)' \mathbf{s}^i$ for edge (i, j). Since the mean utilities and shares for markets $\mathcal{M}$ are identified via cyclic monotonicity, it holds that

$$\sum_{k=1}^{K} \left(\boldsymbol{\delta}^{l_k} - \boldsymbol{\delta}^{l_{k+1}}\right)' \mathbf{s}^{l_k} \geq 0$$

for every path $l_1, \ldots, l_K, l_{K+1} = l_1$ from $\mathcal{M}$ and therefore G has no negative cycles.

*University of Southern California, franguri@usc.edu

[1] See sequence A119913 in the Online Encyclopedia of Integer Sequences

Note that left-hand side of the inequality (1) depends on $\boldsymbol{\delta}^{l_1}$ and so the inequalities for different values of l_1 are generally non-parallel. Hence, for each l_1, we can minimize the right-hand side over all possible paths $l_1, \ldots, l_{K-1}$ starting from l_1 to get the sharpest inequality.

This minimization problem resembles the problem of finding the shortest path from vertex l_1 in the graph G, but is not equivalent to it because of the presence of the second term in (1), which corresponds to the last edge in the path.

I suggest to iterate through $M(M-1)$ ordered pairs (l_1, l_{K-1}) of different vertices from $\{1, \ldots, M\}$ and, for each such pair, find the shortest path that connects l_1 and l_{K-1} in graph G. This can be accomplished using the Floyd–Warshall algorithm, which has polynomial computational complexity $O\left(M^3\right)$ for graphs with no negative cycles such as G. The length of the shortest path between l_1 and l_{K-1} will correspond to the sharpest of the family of inequalities

$$\mathcal{F}_{l_1, l_{K-1}} = \{\text{ineq}(l_1, \ldots, l_{K-1}) : \ l_2, \ldots, l_{K-2} \in \mathcal{M} \text{ s.t. } l_1, \ldots, l_{K-1} \text{ is a path from } l_1 \text{ to } l_{K-1}\},$$

from which we can then pick the sharpest one over l_{K-1}.

I suggest the following **algorithmic implementation**:

- Run the Floyd–Warshall algorithm on graph G, obtaining a matrix D_G of shortest path lengths between every pair of vertices. Set $(D_G)_{ii} = 0$ for all $i \in \mathcal{M}$.
- For every $l_1 \in \mathcal{M}$
  - For every $l \in \mathcal{M}$, set $\text{rhs}(l_1, l) = (D_G)_{l_1, l} + \left(\boldsymbol{\delta}^l - \boldsymbol{\delta}^{M+1}\right)' \mathbf{s}^l$
  - set $l^* = \arg\min_{l \in \mathcal{M}} \text{rhs}(l_1, l)$
  - save the inequality $\text{ineq}^*_{l_1}$ with the right hand side $\text{rhs}(l_1, l^*)$
- The resulting system of M inequalities $\text{ineq}^*_1, \ldots, \text{ineq}^*_M$ is equivalent to the system (1) with all cycles exhausted.

2 Monte Carlo Simulation

I extend a Monte Carlo exercise of Chiong et al. (2017) and compare performance of their original procedure using only 2-cycles (denoted CHS) with the suggested procedure exhausting all possible cycles (denoted All cyc.).

I use two model specifications: one is the multinomial logit as in Chiong et al. (2017) and the other is multinomial probit with the same DGP parameters as the logit and with the covariance matrix of the error term

$$\Sigma = \begin{pmatrix} 1 & -0.7 & 0.3 \\ -0.7 & 1 & 0.3 \\ 0.3 & 0.3 & 1 \end{pmatrix}$$

A counterfactual of interest is the increase of price of good j by 1% (denoted $p_j \uparrow$), for $j = 1, 2, 3$. As can be seen from Tables 1 and 2, the fully efficient procedure always dominates the CHS procedure both in terms of widths of the bounds and their standard errors. Although the gain is minimal and is unlikely to be relevant in practice, I would still suggest using the fully efficient procedure since it bears no extra computational costs (average computational time is less than 20 seconds for $M = 1000$ and is less than 3 seconds for $M \in \{200, 500\}$).

		$M=200$			$M=500$			$M=1000$		
		s_1	s_2	s_3	s_1	s_2	s_3	s_1	s_2	s_3
$p_1\uparrow$	CHS	0.0478	0.0835	0.0932	0.0074	0.0373	0.0367	0.0070	0.0376	0.0376
		(0.0136)	(0.0267)	(0.0290)	(0.0022)	(0.0129)	(0.0125)	(0.0030)	(0.0117)	(0.0118)
	All cyc.	0.0454	0.0794	0.0878	0.0071	0.0351	0.0345	0.0066	0.0355	0.0355
		(0.0126)	(0.0243)	(0.0255)	(0.0021)	(0.0112)	(0.0108)	(0.0026)	(0.0103)	(0.0104)
$p_2\uparrow$	CHS	0.0890	0.0575	0.0978	0.0131	0.0245	0.0289	0.0101	0.0255	0.0289
		0.0286	0.0217	0.0306	0.0046	0.0118	0.0105	0.0031	0.0075	0.0082
	All cyc.	0.0837	0.0551	0.0915	0.0123	0.0230	0.0270	0.0096	0.0247	0.0278
		(0.0256)	(0.0197)	(0.0271)	(0.0040)	(0.0098)	(0.0090)	(0.0028)	(0.0069)	(0.0075)
$p_3\uparrow$	CHS	0.0844	0.0833	0.0626	0.0125	0.0314	0.0268	0.0097	0.0263	0.0225
		(0.0306)	(0.0279)	(0.0254)	(0.0041)	(0.0085)	(0.0073)	(0.0032)	(0.0092)	(0.0086)
	All cyc.	0.0799	0.0790	0.0597	0.0119	0.0299	0.0256	0.0092	0.0250	0.0214
		(0.0274)	(0.0256)	(0.0226)	(0.0036)	(0.0077)	(0.0066)	(0.0029)	(0.0083)	(0.0077)

Table 1: Widths of bounds on counterfactual market shares and their standard errors. Bounds always cover true (**logit**) counterfactuals. Number of simulations is 100.

		$M=200$			$M=500$			$M=1000$		
		s_1	s_2	s_3	s_1	s_2	s_3	s_1	s_2	s_3
$p_1\uparrow$	CHS	0.0050	0.0051	0.0003	0.0167	0.0516	0.0523	0.0006	0.0060	0.0061
		(0.0015)	(0.0015)	(0.0002)	(0.0056)	(0.0161)	(0.0156)	(0.0001)	(0.0018)	(0.0018)
	All cyc.	0.0048	0.0049	0.0003	0.0160	0.0487	0.0496	0.0006	0.0057	0.0058
		(0.0013)	(0.0013)	(0.0002)	(0.0052)	(0.0140)	(0.0139)	(0.0001)	(0.0016)	(0.0016)
$p_2\uparrow$	CHS	0.0036	0.0035	0.0003	0.0271	0.0333	0.0434	0.0012	0.0048	0.0051
		(0.0017)	(0.0017)	(0.0002)	(0.0082)	(0.0104)	(0.0133)	(0.0004)	(0.0012)	(0.0012)
	All cyc.	0.0035	0.0033	0.0003	0.0256	0.0318	0.0412	0.0011	0.0046	0.0049
		(0.0016)	(0.0015)	(0.0002)	(0.0075)	(0.0095)	(0.0121)	(0.0003)	(0.0011)	(0.0011)
$p_3\uparrow$	CHS	0.0051	0.0051	0.0000	0.0265	0.0413	0.0332	0.0013	0.0063	0.0063
		(0.0021)	(0.0021)	(0.0000)	(0.0095)	(0.0137)	(0.0128)	(0.0004)	(0.0017)	(0.0017)
	All cyc.	0.0049	0.0049	0.0000	0.0248	0.0389	0.0315	0.0012	0.0059	0.0059
		(0.0020)	(0.0020)	(0.0000)	(0.0084)	(0.0122)	(0.0112)	(0.0004)	(0.0015)	(0.0016)

Table 2: Widths of bounds on counterfactual market shares and their standard errors. Bounds always cover true (**probit**) counterfactuals. Number of simulations is 100.

References

Allen, R. and Rehbeck, J. (2019). Identification with additively separable heterogeneity. *Econometrica*.

Chiong, K., Hsieh, Y.-W., and Shum, M. (2017). Counterfactual estimation in semiparametric discrete-choice models.