1 The choo and siow model, matching estimation and inverse optimal transport

The Choo-Siow model.

Heterosexual marriage market men's characteristics $x \in X$ women's $y \in Y$.

n_x men of type x and m_y women of type y

If x and y match, then they get respectively

$$\alpha_{xy} + \tau_{xy} + \varepsilon_y \text{ and } \gamma_{xy} - \tau_{xy} + \eta_x$$

where τ_{xy} is a transfer from y to x. If partners remain unmatched, they get respectively

$$\varepsilon_0 \text{ and } \eta_0.$$

Assume $(\varepsilon_y)_{y \in Y \cup \{0\}}$ is iid Gumbel and $(\eta_x)_{x \in X \cup \{0\}}$ is the same. In this logit framework, we have

$$\frac{\mu_{xy}}{n_x} = \frac{\exp(\alpha_{xy} + \tau_{xy})}{\sum_{y'} \cdots} = \exp(\alpha_{xy} + \tau_{xy} - u_x)$$

where $u_x = \log\left(1 + \sum_{y \in Y} \exp(\alpha_{xy} + \tau_{xy})\right)$.

$$\frac{\mu_{x0}}{n_x} = \frac{1}{\sum_{y'} \cdots} = \exp(-u_x).$$

Thus we have

$$\frac{\mu_{xy}}{n_x} = \frac{\mu_{x0}}{n_x} \exp(\alpha_{xy} + \tau_{xy})$$

that

$$\mu_{xy} = \mu_{x0} \exp(\alpha_{xy} + \tau_{xy}).$$

Similarly on the other side of the market we have

$$\mu_{xy} = \mu_{0y} \exp\left(\gamma_{xy} - \tau_{xy}\right).$$

Multiplying term by term to eliminate τ, we get

$$\begin{aligned} \mu_{xy}^2 &= \mu_{x0}\mu_{0y} \exp\left(\alpha_{xy} + \gamma_{xy}\right) \\ &= \mu_{x0}\mu_{0y} \exp(\Phi_{xy}) \end{aligned}$$

Thus

$$\mu_{xy} = \sqrt{\mu_{x0}\mu_{0y}} \exp\left(\frac{\Phi_{xy}}{2}\right).$$

This is a matching function. μ_{x0} and μ_{0y} are determined by

$$\begin{aligned} n_x &= \sum_{y \in Y} \mu_{xy} + \mu_{x0} \\ m_y &= \sum_{x \in X} \mu_{xy} + \mu_{0y} \end{aligned}$$

that is

$$\begin{aligned}
n_x &= \sum_{y\in Y}\sqrt{\mu_{x0}\mu_{0y}}\exp\left(\frac{\Phi_{xy}}{2}\right)+\mu_{x0}\\
m_y &= \sum_{x\in X}\sqrt{\mu_{x0}\mu_{0y}}\exp\left(\frac{\Phi_{xy}}{2}\right)+\mu_{0y}
\end{aligned}$$

Note that $\mu_{x0}=n_x\exp(-u_x)$ and $\mu_{0y}=m_y\exp(-v_y)$ and therefore

$$\begin{aligned}
n_x &= \sum_{y\in Y}\sqrt{n_x m_y}\exp\left(\frac{\Phi_{xy}-u_x-v_y}{2}\right)+n_x\exp(-u_x)\\
m_y &= \sum_{x\in X}\sqrt{n_x m_y}\exp\left(\frac{\Phi_{xy}-u_x-v_y}{2}\right)+m_y\exp(-v_y)
\end{aligned}$$

My question to you: of what minimization problem are these the first order conditions?

$$\min_{u_x,v_y}\sum_x n_x u_x+\sum_y m_y v_y+2\sum_{xy}\sqrt{n_x m_y}\exp\left(\frac{\Phi_{xy}-u_x-v_y}{2}\right)+\sum_x n_x\exp(-u_x)+\sum_y m_y\exp(-v_y)$$

Note that at equilibrium, the quantity

$$\begin{aligned}
& 2\sum_{xy}\sqrt{n_x m_y}\exp\left(\frac{\Phi_{xy}-u_x-v_y}{2}\right)+\sum_x n_x\exp(-u_x)+\sum_y m_y\exp(-v_y)\\
= & 2\sum_{xy}\mu_{xy}+\sum_x\mu_{x0}+\sum_y\mu_{0y}
\end{aligned}$$

is nothing else than the sum of men and women, that is $\sum_x n_x+\sum_y m_y$.

2 Back to a model with heterogeneity in preferences in finite population

Assume that one inidividual man $i\in I$ has observable characteristics $x_i\in X$ and one inidividual woman $j\in J$ has observable characteristics $y_j\in Y$

If i and j match, then their joint utility is

$$\phi_{ij}=\Phi_{x_iy_j}+\varepsilon_{iy_j}+\eta_{x_ij}$$

and if they remain unmatched, they get respectively

$$\varepsilon_{i0} \text{ and } \eta_{0j}.$$

Dual of the optimal matching problem

$$\begin{aligned}
\min_{u_i, v_j} \quad & \sum_i u_i + \sum_j v_j \\
s.t. \quad & u_i + v_j \geq \phi_{ij} \\
& u_i \geq \varepsilon_{i0} \\
& v_j \geq \eta_{0j}
\end{aligned}$$

Solve this knowing $\phi_{ij} = \Phi_{x_i y_j} + \varepsilon_{i y_j} + \eta_{x_i j}$. We have

$$u_i + v_j \geq \Phi_{x_i y_j} + \varepsilon_{i y_j} + \eta_{x_i j}$$

For all i such that $x_i = x$ and j such that $y_j = y$, we have

$$u_i + v_j \geq \Phi_{xy} + \varepsilon_{iy} + \eta_{xj}$$

that is

$$u_i - \varepsilon_{iy} + v_j - \eta_{xj} \geq \Phi_{xy}$$

that

$$\min_{i: x_i = x} \left\{ u_i - \varepsilon_{iy} \right\} + \min_{j: y_j = y} \left\{ v_j - \eta_{xj} \right\} \geq \Phi_{xy}$$

thus if we set

$$\begin{aligned}
U_{xy} &= \min_{i: x_i = x} \left\{ u_i - \varepsilon_{iy} \right\} \\
V_{xy} &= \min_{j: y_j = y} \left\{ v_j - \eta_{xj} \right\}
\end{aligned}$$

then the constraint rewrite

$$U_{xy} + V_{xy} \geq \Phi_{xy}$$

and

$$\begin{aligned}
u_i &\geq U_{xy} + \varepsilon_{iy} \\
v_j &\geq V_{xy} + \eta_{xj}
\end{aligned}$$

The dual problem therefore equivalently rewrites as

$$\begin{aligned}
\min_{u_i, v_j, U_{xy}, V_{xy}} \quad & \sum_i u_i + \sum_j v_j \\
s.t. \quad & U_{xy} + V_{xy} \geq \Phi_{xy} \\
& u_i \geq U_{xy} + \varepsilon_{iy} \\
& v_j \geq V_{xy} + \eta_{xj} \\
& u_i \geq \varepsilon_{i0} \\
& v_j \geq \eta_{0j}
\end{aligned}$$

But $u_i \geq U_{xy} + \varepsilon_{iy}$ and $u_i \geq \varepsilon_{i0}$ can be reexpressed as

$$u_i \geq \max_y \{U_{xy} + \varepsilon_{iy}, \varepsilon_{i0}\}$$

but at an optimum, this should hold as an equality. Thus the problem rewrites as

$$\begin{aligned} \min_{U_{xy}, V_{xy}} \quad & \sum_{i \in I} \max_y \{U_{xy} + \varepsilon_{iy}, \varepsilon_{i0}\} + \sum_{j \in J} \max_x \{V_{xy} + \eta_{xj}, \eta_{0j}\} \\ s.t. \quad & U_{xy} + V_{xy} = \Phi_{xy} \end{aligned}$$

This can be rewritten as

$$\begin{aligned} \min_{U_{xy}, V_{xy}} \quad & \sum_{x \in X} n_x \frac{1}{n_x} \sum_{i: x_i = x} \max_y \{U_{xy} + \varepsilon_{iy}, \varepsilon_{i0}\} + \sum_{y \in Y} m_y \frac{1}{m_y} \sum_{j: y_j = y} \max_x \{V_{xy} + \eta_{xj}, \eta_{0j}\} \\ s.t. \quad & U_{xy} + V_{xy} = \Phi_{xy} \end{aligned}$$

Let's now assume a large market.

$$\begin{aligned} \frac{1}{n_x} \sum_{i: x_i = x} \max_y \{U_{xy} + \varepsilon_{iy}, \varepsilon_{i0}\} \quad & \to \quad E\left[\max_y \{U_{xy} + \varepsilon_y, \varepsilon_0\}\right] \\ \frac{1}{m_y} \sum_{j: y_j = y} \max_x \{V_{xy} + \eta_{xj}, \eta_{0j}\} \quad & \to \quad E\left[\max_x \{V_{xy} + \eta_x, \eta_0\}\right] \end{aligned}$$

Let's now make the logit assumption

$$\begin{aligned} E\left[\max_y \{U_{xy} + \varepsilon_y, \varepsilon_0\}\right] &= \log\left(1 + \sum_y \exp U_{xy}\right) \\ E\left[\max_x \{V_{xy} + \eta_x, \eta_0\}\right] &= \log\left(1 + \sum_x \exp V_{xy}\right) \end{aligned}$$

Thus we have

$$\begin{aligned} \min_{U_{xy}, V_{xy}} \quad & \sum_{x \in X} n_x \log\left(1 + \sum_y \exp U_{xy}\right) + \sum_{y \in Y} m_y \log\left(1 + \sum_x \exp V_{xy}\right) \\ s.t. \quad & U_{xy} + V_{xy} = \Phi_{xy} \end{aligned}$$

which rewrites into

$$\min_{U_{xy}} \sum_{x \in X} n_x \log\left(1 + \sum_y \exp U_{xy}\right) + \sum_{y \in Y} m_y \log\left(1 + \sum_x \exp\left(\Phi_{xy} - U_{xy}\right)\right)$$

By first order conditions, we have

$$\mu_{xy} = n_x \frac{\exp U_{xy}}{1 + \sum_y \exp U_{xy}} = m_y \frac{\exp \Phi_{xy} - V_{xy}}{1 + \sum_x \exp\left(\Phi_{xy} - U_{xy}\right)}$$

and

$$\mu_{x0} = \frac{n_x}{1 + \sum_y \exp U_{xy}} \text{ and } \mu_{0y} = \frac{m_y}{1 + \sum_x \exp\left(\Phi_{xy} - U_{xy}\right)}$$

therefore

$$\mu_{xy}^2 = \mu_{x0}\mu_{0y} \exp\left(\Phi_{xy}\right)$$

and we recover the Choo-Siow model.