Empirical IO: Homework 1

Cameron Scalera, Yoel Feinberg, Antoine Chapel

February 28, 2025

Problem 1

Part (a)

The individual choice probability for individual h to choose option j at period t is given by:

$$P_{hjt} = \frac{\exp(\beta_j^h + \eta^h x_{jt}^h)}{1 + \sum_{j=1}^{J} \exp(\beta_j^h + \eta^h x_{jt}^h)}$$

For each individual h, we observe a sequence of choices $y_{ht} = (y_{h1}, ..., y_{hT})$. The likelihood function for an individual's sequence of choices is given by:

The individual-level log-likelihood function is given by:

$$\begin{aligned}
\mathcal{L}_h(\Theta^h) = \log\left(L_h(\Theta^h)\right) &= \sum_{t=1}^{T} \sum_{j \in J} y_{jt}^h \log(P_{hjt}) \\
&= \sum_{t=1}^{T} \sum_{j=1}^{J} y_{jt}^h \log\left(\frac{\exp(\beta_j^h + \eta^h x_{jt}^h)}{1 + \sum_{k=1}^{J} \exp(\beta_k^h + \eta^h x_{kt}^h)}\right)
\end{aligned}$$

Where y_{jt}^h is to be correctly interpreted as an indicator for which alternative individual h has chosen. Given that the choice set does not change over time, the denominator of the probability remains a simple expression.

The score function for individual h is given by:

$$\nabla_{\Theta^h}\mathcal{L}_h(\Theta^h) = \begin{bmatrix} \frac{\partial \mathcal{L}_h(\Theta^h)}{\partial \beta_1^h} \\ \vdots \\ \frac{\partial \mathcal{L}_h(\Theta^h)}{\partial \beta_J^h} \\ \frac{\partial \mathcal{L}_h(\Theta^h)}{\partial \eta^h} \end{bmatrix}$$

$$\frac{\partial \mathcal{L}}{\partial \beta_k^h} = \sum_j \sum_t y_{jt}^h \frac{\partial}{\partial \beta_k^h} \log(P_{hjt})$$

For easy notation, I denote $v_k = \beta_k^h + \eta^h x_{kt}^h$

$$\begin{aligned} \frac{\partial \log(P_{hjt})}{\partial \beta_k^h} &= \frac{1+\sum_k e^{v_k}}{e^{v_j}} \frac{-e^{v_j} e^{v_k}}{(1+\sum_k e^{v_k})^2} \\ &= -P_{hkt} \end{aligned}$$

$$\begin{aligned} \frac{\partial \log(P_{hjt})}{\partial \beta_j^h} &= \frac{1+\sum_k e^{v_k}}{e^{v_j}} \frac{e^{v_j}(1+\sum_k e^{v_k}) - e^{v_j} e^{v_j}}{(1+\sum_k e^{v_k})^2} \\ &= (1-P_{hjt}) \end{aligned}$$

$$\begin{aligned} \frac{\partial \log(P_{hjt})}{\partial \eta^h} &= \frac{\partial}{\partial \eta_h} v_j - \log(1+\sum_k e^{v_k}) \\ &= x_{jt}^h - \frac{1}{1+\sum_k e^{v_k}} \sum_k e^{v_k} x_{kt}^h \\ &= x_{jt}^h - \sum_k P_{kt}^h x_{kt}^h \end{aligned}$$

So the score vector is given by:

$$\nabla_{\Theta^h}\mathcal{L}_h(\Theta^h) = \begin{bmatrix} \sum_{t=1}^T \left(y_{1t}^h (1-P_{h1t}) + \sum_{j\neq 1} y_{jt}^h(-P_{h1t}) \right) \\ \vdots \\ \sum_{t=1}^T \left(y_{Jt}^h (1-P_{hJt}) + \sum_{j\neq J} y_{jt}^h(-P_{hJt}) \right) \\ \sum_{t=1}^T \sum_{j=1}^J y_{jt}^h \left(x_{jt}^h - \sum_{k=1}^J P_{hkt} x_{kt}^h \right) \end{bmatrix}$$