Elección Discreta: Teoría, estimación y métodos númericos

Tarea 4: Demanda de productos heterogéneos (MLE)*

Germán Augusto Campos Ortíz** Juan José Merino Zarco***

24 de abril de 2022

1. Pregunta 1

Reportar los valores estimados para los parámetros que caracterízan la distribución de preferencias en la población, esto es $\bar{\beta}, \bar{\alpha}, \hat{\sigma}$. Recuerden probar varios puntos iniciales y algoritmos de optimización y reportar los resultados. Reporten además el número de evaluaciones de la función objetivo para cada algoritmo.

Solución:

Cuadro 1: Resumen de estimaciones

Metodo	Tiempo (s)	Optimizacion	Iteraciones	Valor óptimo	Punto inicial
Nelder-Mead	29.20033097	False	22117	7274.864	0
BFGS	20.76631618	False	100	6754.766	0
L-BFGS-B	8.188747644	True	72	6754.766	0
Nelder-Mead	22.52596116	True	72	6754.766	1
BFGS	14.29874778	True	72	6754.766	1
L-BFGS-B	6.436999083	True	72	6754.766	1
Nelder-Mead	22.75346065	True	72	6754.766	-1
BFGS	33.49950838	True	72	6754.766	-1
L-BFGS-B	6.993002653	True	72	6754.766	-1

*El Colegio de México 2020-2022. Profesor: Edwin Muñoz Rodríguez

**El Colegio de México, gacampos@colmex.mx
***El Colegio de México, jmerino@colmex.mx

En el cuadro No. 1 se presenta un resumen de los diversos métodos utilizados para evaluar la función, junto con 3 distintos puntos de inicio. El punto inicial indicado en el cuadro corresponde al valor inicial para todas las variables. A su vez, en este cuadro se presentan el resultado de la optimización, en donde True indica que la función convergió a un punto óptimo y False indica el caso contrario. También se muestra el valor óptimo de la función, el número de iteraciones que le tomó a cada algoritmo y el tiempo en segundos.

El método “Nelder Mead” y el método “BFGS” no lograron convergir y encontrar un óptimo en el punto inicial en donde todas las variables son iguales a cero. Sin embargo, todos los demás modelos bajo los distintos puntos iniciales si convergieron con el mismo número de iteraciones y llegando a un valor óptimo muy similar en todas las especificaciones.

Dado lo anterior, y considerando que el valor óptimo de la función es prácticamente idéntico para todas las combinaciones que si convergieron, se decidió presentar solamente un conjunto de estimaciones, el cual corresponde al método “L-BFGS-B” con el punto inicial en cero para todas las variables. Esta decisión está sustentada en el hecho de que las estimaciones no cambian de forma significativa ante los distintos modelos que convergieron, lo que también hace posible que los resultados sean confiables.

Con la optimización exitosa se estimaron los valores de $\hat{\sigma}$ y $\hat{\delta}$, y posteriormente mediante el uso de variables instrumentales por el método de momentos, llegamos a calcular los estimadores de los parámetros $\bar{\beta}, \bar{\alpha}$. En el Cuadro No. 2 presentamos los resultados para $\hat{\sigma}$.

Cuadro 2: Estimaciones para el vector de parametros σ

Parámetro	**Estimación**
sigma_air	1.227
sigma_hp2wt	2.925
sigma_mpd	2.321
sigma_car_size	-0.486
sigma_price	2.763

Por su parte, en el Cuadro 3 se presenta las estimaciones $\hat{\beta}$ y $\hat{\alpha}$ usando variables instrumentales por medio del método de momentos. Los instrumentos utilizados tienen las terminaciones de ive e ivi, en donde ive son instrumentos de atributos pertenecientes a otros firmas, e ivi, son aquellos instrumentos que se extraen de los atributos de las propias firmas.

Así mismo, en el Cuadro 3 se presentan los resultados de varias combinaciones de instrumentos. En el Modelo 1, por ejemplo, solo se utiliza el instrumento relacionado con el atributo de aire acondicionado de los carros pertenecientes a las otras firmas, mientras que el modelo 2 incluye un instrumento más, y de esta forma se interpreta el cuadro. En el modelo 4 se incluyen todos los instrumentos de atributos de otras firmas (ive) y en el modelo 5 todos los instrumentos pero ahora pertenecientes a los atributos de las propias firmas (ivi) . Por ultimo, el modelo 6 incluye un promedio aritmético de cada instrumento, logrando así estimar un modelo con todos los instrumentos.

Cuadro 3: Estimaciones de $\hat{\beta}$ y $\hat{\alpha}$

	Modelo 1	**Modelo 2**	**Modelo 3**	**Modelo 4**	**Modelo 5**	**Modelo 6**
hp2wt	0.0342	0.2671	0.2988	0.2982	0.2274	0.5396
air	0.0801	-0.0755	-0.0361	-0.0370	-0.1246	0.2621
mpd	-0.0146	-0.1176	-0.1242	-0.1241	-0.1093	-0.1742
car_size	1.1395	0.7884	0.8855	0.8834	0.6672	1.6210
price	-0.6210	-0.7444	-0.8553	-0.8529	-0.6061	-1.6945
Instrumentos						
air_ive	Sí	Sí	Sí	Sí		Sí
mpd_ive		Sí	Sí	Sí		Sí
hp2wt_ive			Sí	Sí		Sí
car_size_ive				Sí		Sí
air_ivi					Sí	Sí
mpd_ivi					Sí	Sí
hp2wt_ivi					Sí	Sí
car_size_ivi					Sí	Sí

En general, las estimaciones de los coeficientes son robustos ante las diferentes combinaciones de instrumentos, dado que su signo no cambia (a excepción de la variable air) y el tamaño de los coeficientes, aunque cambian, las variaciones son generalmente reducidas.

Los atributos hp2wt (relación potencia a peso) y car_size (tamaño del carro) tienen un efecto positivo robusto, siendo el atributo del tamaño del carro el de mayor impacto, mientras que los atributos air (aire acondicionado) y mpd (millas por dólar) afectan de forma negativa la elección de los individuos, aunque en el caso del aire acondicionado, cuando se estima el modelo 6, combinando todos los instrumentos, el atributo del aire acondicionado pasa a tener un efecto positivo.

Finalmente, en cuanto al precio, su impacto en las elecciones de los individuos es negativo, y el tamaño del efecto es mucho más grande cuando se estima el modelo 6, combinando todos los instrumentos.

Centrándonos en el modelo 6, y eligiéndolo como el mejor modelo (dado que tiene una combinación en promedio de todos los instrumentos), todos los coeficientes parecen tener el signo lógico y consistente. En resumen, de acuerdo a las estimaciones, los consumidores tienden a preferir un carro de mayor potencia, que tenga aire acondicionado, y que sea un carro de mayor dimensión en tamaño. A su vez, las estimaciones indican que los consumidores en promedio suelen preferir un carro con un consumo de combustible no tan económico (dado que los consumidores parecen tener una preferencia negativa hacía los carros que recorran más millas por cada dólar). Por ultimo, el precio del carro afecta fuerte y negativamente las decisiones de los consumidores.

2. Pregunta 2

Esbocen, pero no escriban el código, el procedimiento que usarían para encontrar los errores estándar de los estimadores reportados en el punto 2 y analicen su viabilidad computacional.

Solución:

En primer lugar, los errores estándar del estimador $\hat{\sigma}$ se podrían encontrar de la forma común como hemos hecho antes en las tareas pasadas, usando las características del Hessiano con las propiedades de teoría asintótica, por lo cual, no habría mayor inconveniente para estimar estos errores estándar.

Por su parte, los errores estandar de $\bar{\beta}$ y $\bar{\alpha}$ no se pueden obtener directamente, ya que estos coeficientes se obtuvieron mediante el método de los momentos a partir de de la $\hat{\bar{\delta}}$ y de los instrumentos. Por lo cual, podríamos recurrir a dos métodos que en teoría son similares, los cuales son el método delta y el método bootstrap. Estos métodos nos permitirán re estimar los errores estándar y aproximarnos a los verdaderos valores con intervalos de confianza aproximados mediante una generalización del teorema del límite central (en el caso de Delta) y mediante muestras aleatorias repetidas de gran tamaño (mediante Bootstrap).

En cuanto a la viabilidad computacional, los errores estándar de $\hat{\sigma}$ no serán difícil de estimar computacionalmente, dado que basta con obtener los resultados óptimos para posteriormente encontrar los errores estándar. Con respecto al método delta, su demanda computacional no es de gran tamaño y se podría calcular sin tener que recurrir a una gran cantidad de tiempo. Por ultimo, el problema computacional quizá lo podríamos encontrar en Bootstrap, dado que el número de muestras repetidas que serían suficientes para confiar en los errores estándar serían por lo menos 500, y en bases de datos de gran tamaño esto podría ser un problema exponencial. E incluso, la dificultad sería mayor si para Bootstrap se eligiera el método de Jackknife. Sin embargo, dado el tamaño de nuestra base de datos, el costo computacional no representaría mayor problema en términos computacionales.

3. Pregunta 3

Describan las estrategias que usaron para mejorar el costo computacional y cómo midieron la mejora.

Solución:

La principal estrategia para mejorar disminuir el tiempo computacional fue utilizar la librería numba, la cual nos permitió reducir la evaluación a segundos. La principal ventaja de numba es que después de compilar la función la primera vez las siguientes evaluaciones de la función son, con mucha diferencia, mas rápidas. La primera evaluación del mixed logit duro 9.7539 segundos, mientras que la segunda evaluación duro 0.009 segundos. Creamos el loop usando loops sin numba, y el código duro 0.02999 segundos, comparando el tiempo con la segunda evaluación usando numba, la diferencia es notable.

Adicionalmente, señalamos los siguientes tres aspectos que disminuyeron el costo computacional. Primero, es que vectorizamos algunas de las operaciones, en particular vectorizamos la creación de la exponencial de las utilidades de cada producto para cada nodo, que por experiencia en las tareas previas acelera el tiempo de estimación. El segundo aspecto, fue el empleo del comando “prange” perteneciente a la librería de numba que paraleliza las operaciones que realiza. Por último, fuimos cuidadosos en no realizar operaciones innecesarias que pudieran aumentar el costo computacional. Dado lo anterior, notamos que las probabilidades de cada producto son iguales para todos los que realizaron la misma compra, por lo que calculamos las probabilidades para cada producto y después multiplicamos esa probabilidad por el número de veces que cada producto fue comprado.