

Máximo Ripani

Nro registro: 893368

Ajustamiento

Trabajo práctico 3

1. Introducción

En el siguiente informe se realizó un estudio con el método de ajustamiento de Whittaker. Recordamos que el recién nombrado es un método mecánico, es decir que el ajustamiento a los datos se hace a travez de diferencias finitas. El mismo utiliza la regularidad y la fidelidad como medidas de bondad del ajuste. Aquello que esta midiendo la regularidad, es la suavidad de la curva de ajustamiento. Esto se logra con diferencia finitas ‘pequeñas’ en valor absoluto. Por otro lado la fidelidad, busca representar el sesgo matemático. El sesgo será ‘bajo’ cuando la distancia entre los valores observados y aquellos ajustados sea ‘pequeña’. A continuación mostramos la formula tanto de fidelidad (F) como de regularidad (R):

$$F = \sum_{x=1}^n w_x (v_x - u_x)^2 \quad \text{tal que } F \rightarrow 0 \quad R = \sum_{x=1}^m (\Delta^z v_x)^2 \quad \text{tal que } R \rightarrow 0$$

2. Primer acercamiento

Se hizo un primer acercamiento al estudio de ajustamiento con Whittaker a travez de R. Se utilizó un programa provisto por la cátedra con algunas modificaciones que consideré necesarias. En principio el problema que se buscaba resolver era minimizar la matriz M. Recordando que la matriz M es una combinación lineal de las medidas de bondad del ajuste utilizadas (fidelidad y regularidad):

$$M = F + h \cdot R$$

Vemos que en la combinación lineal a minimizar aparece h que vendría a ser el coeficiente que pondera el peso relativo que se le da a la regularidad por encima de la fidelidad. Vemos que para h ‘altos’ la regularidad gana importancia y viceversa para h ‘pequeños’. Para facilitar el calculo con varios valores a la vez se propone hacer la operación de forma matricial. A continuación se presenta la expresión como la combinación de dos formas cuadráticas:

$$M = (V - U)^T W (V - U) + h (K_z V)^T I (K_z V)$$

Profesor: J.E. Fabris

Donde U es el vector de valores observados ($n \times 1$); V es el vector de valores ajustados ($n \times 1$); W es una matriz diagonal con ponderadores; k_z es la matriz que expresa las diferencias finitas $((n - z) \times n)$.

Luego el programa de ajustamiento de Whittaker se encargaría de minimizar dicha combinación lineal. Para ello es necesario aclarar que para asegurarse el calculo del mínimo correspondería realizar las condiciones de primer y segundo orden. Se verificó el correcto funcionamiento del programa con datos del estudio de Nocon y Scott. Una vez verificado, se hizo el ajustamiento con datos generados por un programa provisto por la cátedra. Se sugirió utilizar $z = 3$ con $h_1 = 1$; $h_2 = 97$; $h_3 = 1160$. Los resultados se encuentran en la Tabla 1 del Anexo.

Como medida estandarizada para la comparación del ajustamiento se utilizó la Raíz del Error Cuadrático Medio que por sus siglas en inglés es RMSE:

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (y_t^f - y_t^a)^2}{T}}$$

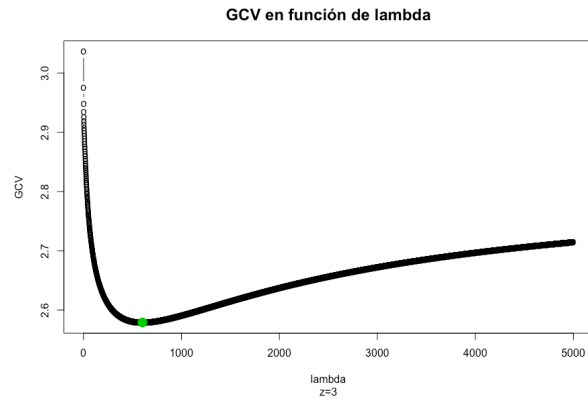
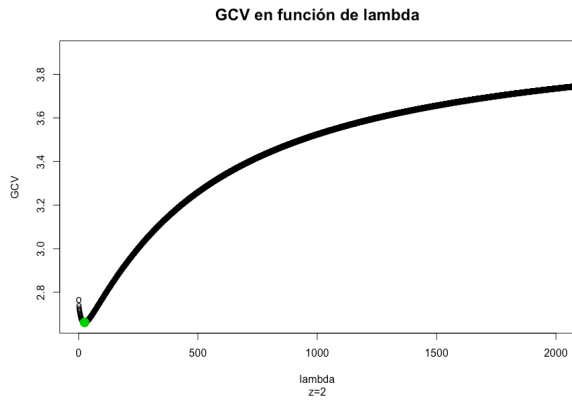
Para los h_i sugeridos por la cátedra el RMSE obtenido fue de 1.457412. De por sí solo este valor no nos es de mucha utilidad pero en el siguiente apartado lo compararemos con otro ajustamiento realizado a partir de la obtención de los h_i óptimos.

3. Más en profundidad

En este apartado se adentro más dentro del método de ajustamiento de Whittaker con el fin de obtener los h_i óptimos (λ según la notación de la cátedra). Para ello se utilizó el Coeficiente de Validación Cruzada que por su siglas en inglés es GCV:

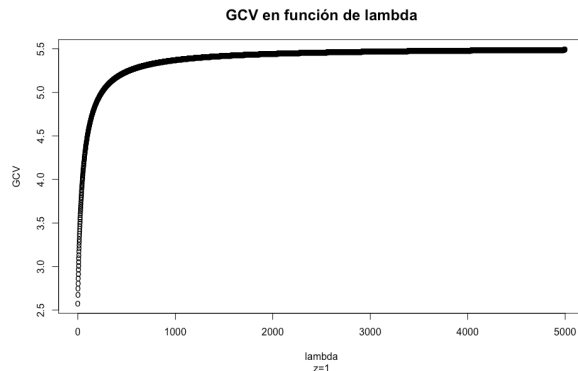
$$GCV = \frac{(V - U)^T W (V - U)}{n \cdot (1 - \frac{tr(H)}{n})}$$

Se calculó el GCV para valores de h_i dentro del intervalo $[1 : 5000]$ y se graficó los resultados en función de h para $z = 1$; $z = 2$; $z = 3$. Aquellos h_i que presenten el mínimo GCV serán los óptimos. Con los gráficos podemos analizar la convergencia del calculo, donde se pudo apreciar que para algunos z la serie converge mientras que para otros no (al menos dentro del intervalo). En la Tabla 2 del Anexo se puede ver el resultado del h óptimo para cada uno de los valores de z posibles.



Como podemos apreciar con sus gráficos para $z = 2$ y $z = 3$ el cálculo del h óptimo converge y se obtiene que $h_2^* = 25$ y $h_3^* = 602$. Por otro lado vemos el caso de $z=1$ donde el cálculo de h_1^* no converge dentro del intervalo. Por lo que se utilizó como $h_1^* = 1$

Finalmente se procedió a realizar el ajustamiento con Whittaker con los h_i^* encontrados. Los resultados se pueden ver en la Tabla 3 del Anexo. Como era de esperarse el ajustamiento tuvo una mejora. El nuevo RMSE para el ajustamiento de Whittaker con los h_i^* fue 1.378651. Como vemos al comparar los RMSE ($1.378651 < 1.457412$) el nuevo es menor al primero, recordamos que un menor RMSE indica una menor distancia entre los valores observados y aquellos ajustados.



Referencias

- *Alicja S. Nocon & William F. Scott (2012) An extension of the Whittaker–Henderson method of graduation, Scandinavian Actuarial Journal*
- *Pablo Caviezel, Principios de Ajustamiento, Análisis Numérico, Facultad de Ciencias Económicas de Buenos Aires*

Anexo

Ajustamiento Whittaker h1=1;h2=97;h3=1160

	x	u	m	vi
1	1	98.4	9.54451281111687	99.4315777274171
2	2	100.1	7.49892506981269	99.9529373874733
3	3	101.5	13.9830230525695	100.462201311856
4	4	99.9	9.58636710420251	100.958733955728
5	5	101.9	14.4607227109373	101.441504975159
6	6	101.2	13.5659355600365	101.908091467952
7	7	99.4	11.3997305766679	102.354426430238
8	8	103	13.1408149749041	102.773867894384
9	9	104.9	9.8237357637845	103.156875444568
10	10	105.2	10.9866811707616	103.493290825405
11	11	103.3	8.18462045863271	103.774349196738
12	12	103.5	5.22424997529015	103.993411612225
13	13	103.5	7.55527646746486	104.14535113454
14	14	106.2	13.3701069932431	104.22644427946
15	15	103	9.11271362332627	104.234653764646
16	16	105.7	7.57889332948253	104.170522060461
17	17	105.9	9.04372212709859	104.036131567523
18	18	104.4	14.4026865228079	103.835199073478
19	19	100.7	7.17533058021218	103.572395739829
20	20	101.5	10.039050988853	103.251391237252

Lambda óptimos ∀ z

	z	lambda	GCV
1	1	1	2.57859733102419
2	2	25	2.66176645239435
3	3	602	2.57903558867769
4	4	5000	2.64181719986281
5	5	5000	2.91694990960491
6	6	5000	3.20951773182263
7	7	27	3.38742380676558
8	8	164	3.37529611407438
9	9	5000	3.46530627005995
10	10	5000	4.18655017131628
11	11	3	4.99352910118752
12	12	4968	5.40171053392067
13	13	1	6.90307895839459
14	14	3127	3.75348139520376
15	15	4814	4.59963832356702
16	16	4149	7.05796116709029
17	17	1483	10.8454189127101
18	18	578	6.2779639062856
19	19	382	22.030645290892

1

Ajustamiento Whittaker h1=1; h2=25; h3=602

	x	u	m	vi
1	1	98.4	9.54451281111687	99.3124565287741
2	2	100.1	7.49892506981269	99.8270489382901
3	3	101.5	13.9830230525695	100.337298632123
4	4	99.9	9.58636710420251	100.843577199222
5	5	101.9	14.4607227109373	101.346871775306
6	6	101.2	13.5659355600365	101.846435026187
7	7	99.4	11.3997305766679	102.338829155435
8	8	103	13.1408149749041	102.815572377518
9	9	104.9	9.8237357637845	103.262004350104
10	10	105.2	10.9866811707616	103.661332505372
11	11	103.3	8.18462045863271	103.998248717057
12	12	103.5	5.22424997529015	104.260086675548
13	13	103.5	7.55527646746486	104.435462042399
14	14	106.2	13.3701069932431	104.514013673508
15	15	103	9.11271362332627	104.48697236771
16	16	105.7	7.57889332948253	104.349098189359
17	17	105.9	9.04372212709859	104.097298626057
18	18	104.4	14.4026865228079	103.731708793047
19	19	100.7	7.17533058021218	103.255441142386
20	20	101.5	10.039050988853	102.672014425551