Modelos De Puntuación Autoregresivos Generalizados

Generalized Autoregressive Score

(GAS)

Mario A. Guerra U.

Abstract

En la actualidad, las series de tiempo se encuentran en la gran mayoría de los análisis de datos, por lo que se propone implementar una de las muchas metodologías denominadas modelos de puntuacin autorregresivos generalizados (GAS), que implementan un mecanismo que es responsable de actualizar los parámetros a medida que pasa el tiempo. Con el puntaje de escalada de la funcin de probabilidad, este mecanismo es más confiable cuando se trata de estimar parámetros que son variantes a lo largo del tiempo. Asimismo, el modelo GAS también cubre otros modelos generalizados conocidos, como La heterosticidad condicional autorregresiva, la duración condicional autorregresiva, la La intensidad condicional autorregresiva y los modelos de conteo de Poisson con media que varía con el tiempo. Cabe agregar que esta metodología puede llevar a la formulación de nuevos modelos basados en observaciones.

Palabras clave: GAS.

1 Introducción.

En muchas configuraciones dentro de la serie de tiempo, la selección de parámetros de un modelo dado que varían con el tiempo es importante para capturar el comportamiento dinámico de procesos de series de tiempo univariados y multivariados. Las series de tiempo con parámetros que varían en el tiempo han sido categorizadas por Cox (1981) en dos clases de modelos: Modelos conducidos por observación y Modelos controlados por parámetros .

El enfoque que se le otorga los modelos impulsados por observacion , la variacion de parametros en el tiempo se introduce al permitir que los parametros sean funciones de variables dependientes .Asi como variables exógenas contemporneas y retrasadas.(Las variables exógenas estn determinadas fuera del modelo. El valor de las variables exógenas está determinado por factores o variables que no son incluídas en el modelo que se est utilizando.),tambien se tiene encuenta que aunque los parametros son estocasticos , son perfectamente predecibles gracias a la informacion pasada dada.Este tipo de enfoque explica por que los modelos orientados a la observacion se han hecho populares en la estadistica aplicada y la econometria . Un ejemplo tipico de estos modelos son los de heteroscedaticidad condicional autoregresvia generalizada (GARCH) modelos de Engle(1982).

En los modelos controlados por los parámetros, los parámetros son procesos

estocásticos con su propia fuente de error. Además, se aclara que, dadas las observaciones pasadas y concurrentes, los parámetros no son perfectamente predecibles, los ejemplos típicos de modelos controlados por parámetros son el modelo estocástico de volatilidad. , esto se ve de manera ms detallada en Shephard (2005), los modelos de intensidad estocástica de Bauwens, Hautsch (2006) Koopman, Lucas y Monteiro (2008).
La estimación suele ser mas engrorosa para estos modelos debido a que las funciones de probabilidad asociados no estan disponibles en forma cerrada.cabe añadir que La funcin de probabilidad requiere la evaluación de una integral de alta dimensión basada en simulacioø'n.
Por consiguiente cuando nos refiramos a un modelo basado en la observacion que se fundamenta en la funcion de puntuacion , lo denotaremos como modelo de puntuacion autoregresivo generalizado (GAS).El modelo GAS tiene ciertas ventajas sobre otros modelos basados en la observacion , se pueden considerar extensiones asimétricas , memoria larga y otras dinámicas mas complicadas sin introducir complejidades adicionales.Dado que el modelo GAS se basa en las puntuaciones , la extructura de densidad es distinta en estos casos y se dan a notar todas las observaciones en la densidad en lugar de los las observaciones medias ademas tambien tienden a sobresalir los momentos superiores.

2 Objetivos.

La principal contribucion de este escrito es dar luces sobre el desarrollo de una metodologia que esta en relacion directa con parametros que varian en el tiempo basados en la funcion de densidad del modelo predictivo en el instante t.

Se argumentara que la funcion de puntuacion es una opcion eficaz para introducir un mecanismo de conduccion para parámetros que varian en el tiempo.

Considerar cuando que cuando al escalar la funcion de puntuacion de manera apropiada ,se pueden recuperar los modelos estandar de observacion como lo son los modelos GARCH,ACD,y ACI.

Mostrar una aplicación de tal modelo y generar una discusion acerca de los resultados impresos bajo esta metodologia

3 Marco teórico

En esta sección se formulará una serie de modelos cuya característica es que los parámetros varían en el tiempo según la observación. Además, se considera la estimación de máxima verosimilitud y la especificación del modelo.

3.1 Especificación del modelo y propiedades.

Sea y_t el vector $N \times 1$ que denota la variable dependiente de interes, f_t el vector de parametros variando en el tiempo, x_t un vector de variables exógenas (Covariables), todas en el tiempo t y θ un vector que tiene parametros que se mantienen fijós por asi decirlo.Se define $Y^t = y_1, ...y_t$, $F^t = f_0, f_1, ...f_t$ y $X^t = x_1, x_2, ...x_t$.La información disponible contenida en el tiempo t consiste en $[f_t; F_t]$ donde

$$F_t = [Y^{t-1}, F^{t-1}, X^t]$$

Se asume que y_t es generado por una densidad de observación:

$$y_t \sim p(y_t \mid f_t, F_t; \theta) \tag{1}$$

Luego se asume que el mecanismo para actualizar los parámetros f_t que varian en el tiempo, es dado por la ecuacion de actualización que tiene una forma autoregresiva y se escribe de la siguiente manera:

$$f_{t+1} = \omega + \sum_{i=1}^{p} A_i s_{t-i+1} + \sum_{j=1}^{q} B_J s_{t-j+1} \tag{2}$$

Donde ω es un vector de constantes , las matrices de coeficientes A_i y B_j tienen dimensiones adecuadas para $i = 1, ..., p$ y $j = 1, ...q$, mientras que s_t es una función apropiada de datos pasados , y se define de la siguiente manera con la siguiente notación, $s_t = s_t(y_t, f_t, F_t; \theta)$. Los coeficientes desconocidos en (2) son funciones de θ es decir que $\omega = \omega(\theta)$; $A_i = A_i(\theta)$, $B_j = B_j(\theta)$, para $i = 1, ...p$ y $j = 1, ...q$. La principal funcion de este escrito es lograr la eleccion particular para el mecanismo de accionamiento s_t que es aplicado directamente en una amplia clase de observaciones , densidades y modelos no lineales.

El enfoque en este caso se basa en la densidad de observación de la ecuación (1) para un parámetro dado f_t .Cuando una observación y_t se realiza, se actualiza, f_t, que varía en el tiempo al siguiente perido $t + 1$ usando (2) con:

$$s_t = S_t \cdot \bigtriangledown_t; \bigtriangledown_t = \frac{\partial ln(p(y_t \mid f_t, F_t; \theta))}{\partial f_t}; S_t = S(t, f_t, F_t; \theta) \tag{3}$$

Donde $S(.)$ es una función de matriz. Dada la dependencia del mecanismo de accionamiento en (2) de la del vector de puntajes escalado en (3) dejamos que las ecuaciones $(1), (2)y(3)$ definan el puntaje autoregresivo generalizado , modelo con ordenes correcponidentes p y q pedidos , podemos abreviar el modelo resultante bajo la siguiente notacion $GAS(p, q)$.

El uso a la hora de actualizar f_t es intuitivo .Esto define una direccion de ascenso más pronunciada para mejorar el ajuste local del modeloen terminos de probabilidad o de densidad en el instante t dada la orientacion posicion del parametro f_t.Esto proporciona la dirección natural para actualizar dicho parametro .Además, la puntuación depende de la densidad completa y no solo de los momentos de primer y segundo orden de las observaciones y_t. Esto diferencia la metodologia GAS de la mayoria de los otros enfoques impulsados por los enfoques de observación . Al utilizar la funcion de densidad en su completitud ,el modelo GAS introduce nuevas transformaciones de los datos que pueden utlizarse para actualizar el parámetro que varia en el tiempo f_t.

A través de la eleccion de la matriz de escala S_t,el modelo GAS permite una felixbilidad adicional en la forma como se emplea la actualizacion de f_t. Es importante tener en cuenta que cada opción diferente para el f_t la matriz escala S_t da como resultado un modelo GAS diferente en cada caso.Las propiedades estadísticas y empíricas de cada uno de estos modelos pueden ser diferentes y requieren un an'alisis por separado.

En muchas situaciones, es natural considerar una forma de escala que dependa de la varianza de la partidura. Por ejemplo definamos la matriz de escala como

$$S_t = I_{t|t-1}^{-1}; I_{t|t-1} = E_{t-1}[\bigtriangledown_t \bigtriangledown_t'] \quad (4)$$

;

Donde E_{t-1} denota un esperado con respecto a $p(y_t, f_t, F_t; \theta)$. Para esta elección de S_t,el modelo GAS abarca el conocido modelo GARCH de Engle basado en observaciones (1982) y Bollerslev (1986), el modelo ACD de Engle y Russell (1998), y el modelo ACI de Russell (2001), así como la mayoría de los modelos de conteo de Poisson considerados por Davis et al. (2003).Otra posibilidad que consideramos en este documento es el modelo GAS con matriz de escalado.

$$S_t = J_{t|t-1}; J_{t|t-1}' J_{t|t-1} = J_{t|t-1}^{-1} \quad (5)$$

donde S_t se define como la matriz de la raíz cuadrada de la matriz de información ,inversa para (1) con respecto a f_t.Una ventaja de esta elección específica para S_t es que las propiedades estadísticas del modelo GAS correspondiente se vuelven más manejables.Esto se deduce del hecho de que para $S_t = J_{t|(t-1)}$, el paso s_t de GAS tiene una variación de unidades constante.

3.2 Caso especial de modelos GAS.

Se proporcionara un ejemplo sencillo que muestra cómo hacer funcionar el modelo GAS .El ejemplo también revela que la metodologia GAS abarca una gran cantidad de modelos disponibles basados en la observación presentados en literaturas referentes a series temporales para una adecuada elección de la matriz de escala S_t.En este caso solo nos referimos al caso en que el modelo GAS coincide con el modelo GARCH.

Ejemplo : Modelo GARCH , se considera el modelo basico $y_t = \sigma_t \epsilon_t$ donde el ruido gaussiano ϵ_t tiene media cero y varianza σ_t^2 variando en el tiempo.Esto es un ejercicio basico para mostrar que el modelo GAS$(1,1)$ con $S_t = I_{t|t-1}^{-1}$ y $f_t = \sigma_t^2$ se reduce a la siguiente ecuacion:

$$f_{t+1} = \omega + A_1(y_t^2 - f_t) + B_1 f_f \tag{6}$$

Que es equivalente a tener el modelo GARCH(1,1) y seguido de la siguiente forma:

$$f_{t+1} = \alpha_0 + \alpha_1(y_t^2 - f_t) + \beta_1 f_t; f_t = \sigma_t^2 \tag{7}$$

donde los coeficientes $\alpha_0 = \omega$, $\alpha_1 = A1$ y $\beta_1 = B_1 - A_1$ son desconocidos y requieren ciertas condiciones para la estacionariedad, ver Bollerslev (1986). Sin embargo, si asumimos que ϵ_t sigue una distribución t de Student con ν grados de libertad y variación de unidades, la especificación GAS$(1,1)$ para la varianza condicional conduce a la ecuación de actualización:

$$f_{t+1} = \omega + A_1 \cdot (1 + 3\nu^{-1}) \cdot (\frac{(1 + \nu^{-1})}{(1 - 2\nu^{-1})(1 + \nu^{-1} y_t/(1 - 2\nu^{-1}))f_t}) \cdot (y_t^2 - f_t) \tag{8}$$

En el caso en que $\nu^{-1} = 0$ a distribución t de Student se reduce a la distribución y actualización gaussianas.El denominador del segundo término en el lado derecho de (8) provoca un aumento más moderado de la varianza para una gran realización de y_t siempre y cuando ν sea finito.La intuición es clara: si los errores son modelados por una distribucion fat-tailed (Una distribución de cola gruesa es una distribución de probabilidad que muestra una gran asimetra o kurtosis, relativa a la de una distribución normal o una distribución exponencial.) una gran realizacion de y_t no requiere un aumento sustancial en la varianza.

3.3 Estimación de máxima verosimilitud.

Una propiedad conveniente de los modelos guiados por observación es la forma relativamente simple de estimar parmetros por máxima verosimilitud (ML).Esta caracter'istica también se aplica al modelo GAS. Para una serie temporal observada $y_1, ..., y_n$ y al adoptar la descomposicin estándar del error de predicción, podemos expresar el problema de maximización como:

$$\hat{\theta} = argmax \sum_{i=1}^{n} \ell_i \tag{9}$$

Donde $\ell = lnp(y_t \mid f_t, F_t, \theta)$ para una realizacion de y_t.Es posible formular recursiones para calcular el gradiente de la probabilidad con respecto al vector de parámetros estáticos. Las recursiones de gradiente para el modelo GARCH han sido desarrolladas por Fiorentini et al. (1996). En el caso de la especificacin GAS (1, 1), el gradiente se calcula a travs de la regla de la cadena, es decir:

$$\frac{\partial \ell}{\partial \theta} = \frac{\partial ln(p_t)}{\partial \theta} + \frac{\partial ln(p_t)}{\partial f_t} \cdot \frac{\partial f_t}{\partial \theta} \tag{10}$$

Donde $p_t = p(y_t \mid f_t, F_t; \theta)$ y

$$\frac{\partial f_t}{\partial \theta'} = \frac{\partial \omega}{\partial \theta'} + A_1 \frac{\partial s_{t-1}}{\partial \theta'} + B_1 \frac{\partial f_{t-1}}{\partial \theta'} + \left(s'_{t-1} \otimes \mathrm{I}\right) \frac{\partial \overrightarrow{A}_1}{\partial \theta'} + \left(f'_{t-1} \otimes \mathrm{I}\right) \frac{\partial \overrightarrow{B}_1}{\partial \theta'}$$

$$\frac{\partial s_{t-1}}{\partial \theta'} = S_{t-1} \frac{\partial \nabla_{t-1}}{\partial \theta'} + \left(\nabla'_{t-1} \otimes \mathrm{I}\right) \frac{\partial \overrightarrow{S}_{t-1}}{\partial \theta'}$$

Los derivados de probabilidad logarítmica pueden calcularse simultáneamente con los parámetros que varian en el tiempo f_t.

Proponemos calcular los errores estándar y los valores de t para los parámetros estimados basados en el Hessiano inverso de la logarítmica probabilidad evaluada en el óptimo. En particular, si θ reúne todos los parámetros estáticos del modelo, suponemos que, en condiciones de regularidad adecuadas, como las de White (1994) y Wooldridge (1994), el estimador de probabilidad máxima $\hat{\theta}$ de θ es consistente y satisface:

$$\sqrt{n}(\hat{\theta} - \theta)\vec{d}N(0, H^{-1}) \tag{11}$$

Donde $H = \lim_{n\to\infty}(E[(\frac{\partial \ell}{\partial \theta})(\frac{\partial \ell}{\partial \theta})'])/n$; y $\ell = \sum_{i=1}^{n} \ell_t$.Una prueba formal de estos resultados para la clase general de modelos GAS está más all del alcance del presente documento ;Los resultados han sido establecidos para subclases especficas de modelos GAS. Por ejemplo, Davis et al. (2005) demostraron la consistencia y la normalidad asinttica del estimador ML para los modelos de conteo de Poisson de primer orden. Straumann y Mikosch (2006) proporcionan un conjunto de condiciones para la consistencia y la normalidad asinttica del modelo GARCH Gaussiano y para especificaciones más generales de GARCH.

4 Conclusiones.

Se logro mostrar una teoria alterna a las vistas con anterioridad referentes a series temporales en las que predominaba el hecho de que el parámetro variaba en el tiempo , ademas se dejo visto un caso en el que el modelo abarcarba un caso particular de modelos como el GARCH ,Aclarando ademas que es una de las tantes posibles particularidades del modelo GAS.

5 Referencias bibliográficas

Creal, D., Koopman, S. J. A. N. (2013). GENERALIZED AUTOREGRESSIVE SCORE MODELS WITH APPLICATIONS, 795(January 2012), 777-795. https://doi.org/10.1002/jae.1279