التقديرات الحصينة في نموذج الانحدار الذاتي
من خلال طريقتي التنقية الملائمة Adaptive Filtering والمربعات الصغرى Least Squares

أ.د. كنعان عبد اللطيف عبد الرزاق

الكلية التقنية الأدارية

أ.م.د. نزار مصطفى

الكلية التقنية الأدارية

الخلاصة

يعتبر أسلوب تحليل السلاسل الزمنية واحد من الأساليب الإحصائية المهمة في بناء نماذج التنبؤ حيث بحثت العديد من طرق التقدير في هذا المجال .

إن هدف البحث هو الوقوف على حساسية تقديرات طريقتي التنقية الملائمة والمربعات الصغرى من خلال اعتماد توزيعات مختلفة للخطأ وحجوم عينات متنوعة ومعالم مختلفة كما إن توليد البيانات وبناء النماذج ومقارنة النتائج قد تم من خلال استخدام أسلوب المحاكاة وباعتماد برامج تم تصميمها وكتابتها من قبل الباحثين وهي متوفرة لديهم.

١.المقدمة وهدف البحث

من الطرق الإحصائية المهمة التي يمكن استخدامها في التنبؤ هي السلاسل الزمنية حيث هنالك العديد من طرق تحليل السلاسل الزمنية، في هذا البحث نتعامل مع طريقة التنقية الملائمة Adaptive Filtering وطريقة المربعات الصغرى.

إن هدف البحث هو الوقوف على حصانة تلك الطرق وذلك عندما يكون هناك اختلاف في توزيع الخطأ لنماذج الانحدار الذاتي

٢. الجانب النظري

٢٠١طريقة التنقية الملائمة Adaptive Filtering

يتم من خلال هذه الطريقة تجزئة المركبة العشوائيةالى مركبات اخرى بهدف الوصول الى النموذج الأكثرملائمة" وتعتمد على معالم المعادلة للمشاهدة السابقة حيث انها تفترض قيم ابتدائية لمعالم معادلة التنبؤ وبعد ذلك يتم التنبؤ بالمشاهدات المستقبلية ثم يتم احتساب الفرق بين القيم الفعلية والقيم المتنبأ بها لإيجاد قيمة الخطأ والذي يتم استخدامه في معادلة خاصة لتحديد قيم جديدة لمعالم المعادلة ويمكن تكرار هذه الطريقة عدة مرات حتى يتم الوصول إلى افضل قيم للمعالم وبالاعتماد على المؤشر الإحصائي المسمى متوسط مربعات الخطأ (Mean squares error) (MSE) ومن ذلك يمكننا ملاحظة التغير في المعالم بعد فترة من الزمن ومن هنا

نلاحظ وجود فرق جوهري بين طريقة (Box- Jenkins) وطريقة (A-F) حيث إن معالم طريقة بوكس وجنكيز ثابتة بينما تكون المعالم متغيرة في طريقة (A-F) لكل مشاهدة جديدة،لذا تكون الخاصية المميزة لهذه الطريقة هي إن لها القدرة على ضبط معالم النماذج تلقائياً بالاعتماد على طريقة الانحدار السريع Steepest-deseent وهي طريقة مثلى لتحديد قيم المعالم.

ولتوضيح هذه الطريقة نأخذ نموذج الانحدار الذاتي Autoregressive model التالي

$$X_{t+1} = W_t^T X_t + e_{t+1} \ldots\ldots\ldots\ldots(1)$$

حيث W_t هي معالم النموذج

$$W_t = (W_{1t}, W_{2t}, \ldots W_{pt})$$

X_t هي قيم المشاهدات

$$X_t = (X_t, X_{t-1}, \ldots . X_{t-p+1}$$

$$e_{t+1} = X_{t+1} - W_t^T X_t \ldots\ldots\ldots.(2)$$

حيث e_{t+1} هي قيم الخطأ

وان طريقة الانحدار السريع Steepest -deseent هي كالأتي

$$W_{t+1} = W_t - \alpha_t g(w_t) \ldots\ldots\ldots.(3)$$

حيث W_{t+1} معالم جديدة

W_t معالم قديمة

$g(w_t)$ هي قيمة الميل ويساوي المشتقة الأولى لمربع الخطأ بالنسبة إلى قيم المعالم أي إن

$$g(w_t) = \left\lfloor \frac{\partial e_{t+1}^2}{\partial w_t} \ldots\ldots \frac{\partial e_{t+1}^2}{\partial w_p} \right\rfloor$$

$$g(w_t) = \frac{\left\{\partial e_{t+1}^2\right\}}{\partial w_1} = -2 e_{t+1} X_t$$

حيث (α) هو ثابت معلوم ويرمز له عادة بالحرف (k) لذا تصبح معادلة الانحدار السريع كالأتي

$$W_{t+1} = W_t \mp 2k e_{t+1} X_{t \ldots\ldots\ldots\ldots(4)}$$

وعلى ضوء هذه العلاقة نستطيع ايجاد قيمة MSE لكل معلمة جديدة وتصبح قيم المعالم مثلى عندما نحصل على اقل قيمة لمقدار MSE .

ان قيمة K في المعادلة (4) تلعب دوراً مهماً في سرعة الوصول الى اقل قيمة ل(MSE) اي عند اختيارقيمة K صغيرة جداً سوف يكون مقدار التغير في قيم المعالم صغير جداً وبالتالي تكون سرعة الوصول الى اقل قيمة ل (MSE) بطيئة اما عند اختيار قيمة K كبيرة فأن تغير قيم المعالم يصبح كبير وبالتالي قد يتجاوز قيمة ل(MSE) الصغرى او يكون متذبذباً حولها ولايستطيع الوصول اليها . لذا فان قيمة K تتحدد بالاعتماد على طبيعة البيانات وحجمها وكذلك على درجة النموذج المستخدم للتنبؤ .

٢٠٢طريقة المربعات الصغرى Least squares method

تعتمد طريقة المربعات الصغرى على جعل مجموع مربعات الخطأ اقل مايمكن

أي عندما يكون لدينا نموذج الانحدار الذاتي التالي

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} ..... \phi_p X_{t-p} + e_t \quad ....(5)$$

$$\sum e_t^2 = \sum (X_t - \phi_1 X_{t-1} ..... \phi_p X_{t-p})^2 \quad ....(6)$$

ولكي يتم الحصول على اقل مجموع مربعا ت للخطأ يجب اخذ المشتقة الأولى للمعادلة بالنسبة إلى قيمة ϕ_i

مثلا عند النموذج AR(1) سوف تكون قيمة ϕ **كالأتي**

$$\phi^{\wedge} = \frac{\sum x_t x_{t-1}}{\sum x_{t-1}^2} \quad ....(7)$$

٢٠٣المحاكاة Simulation

تعتبر المحاكاة الأسلوب المناسب الذي يمكن استخدامه للمقارنة بين طريقة التنقية الملائمة -A) F) مع طريقة المربعات الصغرى Least squares وذلك بالاعتماد على المؤشر الإحصائي MSE ومن خلال دراسة تأثير عدم تحقق شرط التوزيع الطبيعي للخطأ العشوائي و تأثير العينات الصغيرة والكبيرة و تجانس الخطأ العشوائي .

وقد تناول البحث توزيعات متعددة للخطأ منها:

١- التوزيع الطبيعي Normal distribution

يعتبر التوزيع الطبيعي من اهم التوزيعات المستمرة واكثرها استخداماً ويقال للمتغير x بان له توزيع طبيعي اذا كانت دالة الكثافة الاحتمالية هي:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(x-\mu)^2 \;\; / 2\sigma^2 \qquad \begin{array}{l} \infty < x, \mu < \infty \\ \sigma^2 > 0 \end{array} \qquad ....(6)$$

وباستخدام أسلوب Box-muller نحصل على ارقام عشوائية لها توزيع طبيعي قياسي من الأرقام العشوائية ذات التوزيع المنتظم وكالأتي

$$X = (-2\log(u_1)^{\frac{1}{2}} \cos(2\pi u_2))$$

حيث u_1, u_2 أرقام عشوائية ذات توزيع منتظم

ويمكن الحصول على متغير ذات توزيع طبيعي بمتوسط μ وتباين σ^2 باستخدام الصيغة التالية $Y = \mu + \sigma X$

٢- توزيع الجيب العكسي Arcsine distribution

يقال للمتغير الذي يملك دالة الكثافة الاحتمالية التالية

$$f(x) = \frac{1}{\pi\sqrt{2-x^2}} \qquad \sqrt{2} < x < \sqrt{2}$$

بأنه يتبع توزيع (Arcsine) بمتوسط صفر وتباين واحد

وبأستخدام التحويل العكسي نحصل على

$$x = \sqrt{2}\sin((u-0.5)\pi$$

حيث u رقم عشوائي متولد ذات توزيع منتظم

٣- توزيع اللوجستك Logistic Distribution

المتغير العشوائي X الذي له دالة الكثافة الاحتمالية التالية :

$$f(x)=\frac{Exp(1-(x-\mu)/\beta)}{\beta(1+Exp(-(x-\mu)/\beta))^2} \qquad \infty<x,\mu<\infty$$

$$0<\beta$$

وباستخدام أسلوب التحويل نحصل على

$$X=\mu-(LOG(1-U)-LOG(U))$$

حيث u رقم عشوائي ذات توزيع منتظم

الجانب التطبيقي

لقد تم التركيز على نموذج (AR) الآتي للاعتماد عليه في البحث وهو

$$X_t=\phi X_{t-1}+e_t$$

وتم كتابة البرامج الخاصة من قبل الباحثين لعدم توفر التطبيق الجاهز الذي يستخدم تطبيق تجارب المحاكاة .

وقد تم استخدام أسلوب (A-F)وأسلوب المربعات الصغرى للتقديرولتوزيعات الخطأ الثلاثة وهي التوزيع الطبيعي وتوزيع الجيب العكسي وتوزيع اللوجستك اخذين بنظر الاعتبار حجوم عينات مختلفة (150,100,50) وقيم للتباين (3,6,9) وقيم أولية للمؤشرات هي (-0.75,-0.25,0.25,0.75.)

وقد تم احتساب قيم MSE للمقارنة بين النماذج المحتملة وعلى أساس (250) تجربة مكررة ، وبذلك يكون عدد التجارب المستخدمة لكل توزيع من توازيع الخطأ الثلاثة ومجموع التجارب هي 54000 تجربة والنتائج موضحة ضمن الجداول (1,2,3).

الاستنتاجات والتوصيات

من خلال تجارب المحاكاة تم التوصل إلى الأتي:

١- لوحظ إن هناك تأثير لحجم العينة على النتائج التي تم الوصول إليها ولمختلف توزيعات الخطأ .

٢- إن زيادة تباين الخطأ يؤدي إلى زيادة قيم MSE .

٣- يلاحظ إن قيم MSE لاتتاثر عند اختلاف قيم المعالم .

٤- إن أفضل النتائج تم الحصول عليها هي في حالة التوزيع الطبيعي ولكلتا الطريقتين .

٥- أتضح من خلال هذا البحث إن طريقة التنقية الملائمة Adaptive Filtering كانت افضل من طريقة المربعات الصغرى وللتوزيعات الثلاثة وحسب قيم MSE الموضحة في الجداول (1,2,3) وهذا يعني إن طريقة (A-F) هي اكثر حساسية من طريقة المربعات الصغرى .

٦- يوصي الباحثان بتعميم التجربة على نماذج أخرى وتوزيع وطرق أخرى للتقدير للمقارنة وبيان العمومية ودقة النتائج المستخرجة.

ABSTRACT

Time series analysis considered one of statistics methods for build the forecasting model , and the main purpose of this research is to find the difference between the estimation of Adaptive filtering and least squares method ,based on different sample size ,error and different parameter .The simulation procedures is used to generating data and models to compare between these methods .

REFERENCES:

1.Ahtola,J.And Tiao G.C(1984) “Parameter Inference for a nearly No stationary first-order Autoregressive Model “,Biometrika.vol 71 ,n0.2,pp.263-272

2. Arerill M .Law,W. David kelton, "Simulation modeling and Analysis ",Third edition (2000) .

3. Box,G.E.P and Jenkins,G.M,(1976) Time series Analysis Forecasting and control ,Sanfrancisco , Holden- Day U.S.A.

4. Douglas C.m& Controas, Auto on forecasting with Adaptive filtering O.P.Q, vol 28, No, 1,pp,(87-91),1972.

5. Makridakis, S.d wheel wright, S.C.," forecasting Methods and Application ", John Wiley, New York. (1978)

6 . Nakayama,M,h; Fas simulation methods for highly dependable systems, proc 1994,

7.Orcutt,G.H. and Winokur, (1969) First order Autoregressive ,inference ,Estimation and prediction, Econometrica, vol 37, n01,pp1-14 .

8. Terasvivt,T.(1994) "Specification ,estimation and evaluation of smooth transition Autoregressive models ", JASA,89, 208-218 .