تصفية السلاسل الزمنية من التغيرات العشوائية

د. الحسين سالم كعيبة
شعبة الإحصاء – جامعة مصراتة

الملخص

اشتملت هذه الورقة على موضوع تصفية (Smoothing) السلاسل الزمنية من التغيرات العشوائية ، وتكمن أهمية هذا الموضوع بشكل خاص عند دراسة السلاسل الزمنية المعبرة عن الأنشطة الاقتصادية في البلدان النامية حيث يكون المنحنى الفعلي ذو تأرجحات و تذبذبات حادة وقوية مما تسبب في بعض الأحيان إلى إخفاء المنحنى الأساسي ، الأمر الذي يدفعنا إلى عملية التصفية للتخلص من هذه التذبذبات و التأرجحات الغير منتظمة و إظهار المنحنى الأساسي . وقد بحثنا عدة طرق لهذه التصفية .

المقدمة

أن المدخل المتبع لتحليل السلسلة الزمنية يتم في تحديد العوامل الرئيسية التي تؤثر على قيم السلسلة و مثل هذه العوامل غالبا ما تكون غير معلومة و تؤثر باتجاهات مختلفة . وقد صنفت العوامل المؤثرة إلى أربع أنواع وهي :

- العوامل الاتجاهية
- العوامل الدورية
- العوامل الموسمية
- العوامل العشوائية

ومن الواضح أن أيا من هذه العوامل يمكن أن تقع في أي وقت وليس هناك موعد محدد لوقوعها و هي قوى تحدد الحركة العاملة للسلسلة الزمنية بتقلباتها المختلفة سوى إن كانت هذه التقلبات تظهر بشكل دوري شبه منتظم في فترة طولها بين سنة و ثلاثة عشر سنة أو في فترة أقل من سنة أو قوى أخرى متفرقة سواء إن كانت هامة أو قليلة الأهمية. ويمكن التعبير عن قيم السلسلة الزمنية وفق التقسيم السابق بشكل مجموع أو ضرب عناصرها ، بالشكل التالي :

$$y_t = T_t + C_t + S_t + \varepsilon_t$$

أو

$$y_t = T_t * C_t * S_t * \varepsilon_t$$

حيث

y_t = القيمة الفعلية لحد السلسلة الموافق للزمن .

T_t = الجزء من قيمة حد السلسلة الزمنية الموافق للزمن الناتج عن تأثير العوامل الاتجاهية .

C_t = الجزء من قيمة حد السلسلة الزمنية الموافق للزمن الناتج عن تأثير العوامل الدورية.

S_t = الجزء من قيمة حد السلسلة الزمنية الموافق للزمن (t) الناتج عن تأثير العوامل الموسمية.

ε_t = الجزء من قيمة حد السلسة الزمنية الموافق للزمن (t) الناتج عن تأثير العوامل العشوائية.

إن تحليل هذه القوى لمعرفة درجة تأثيرها ونوعها وما تجلبه من أثار على حركات السلسلة الزمنية يعتبر المدخل لتحليل السلاسل الزمنية كما أشرنا سابقا ، ويساعدنا على معرفة الأجزاء المكونة لهذه السلسلة وبالتالي معرفة الوزن العام لكل جزء من هذه من الأجزاء في تكوين حدود السلسلة الزمنية.

ويلاحظ في أغلب الأحيان لدى دراسة السلاسل الزمنية عن المؤشرات الاقتصادية للبلدان النامية بأن مساهمة الجزء العشوائي في تكوين السلسلة الزمنية يلعب دورا ملموسا وهذا ما يعبر عنه بتأرجحات و الانحرافات غير المنتظمة ، بالإضافة إلى ذلك هناك عامل آخر يلعب دورا في التذبذبات الحادة القوية ألا وهو عدم الدقة بالمعلومات الإحصائية في البلدان النامية وهذا ناتج عن عوامل عدة نذكر منها على سبيل المثال وليس الحصر : ضعف الكادر الإحصائي وعدم تفاعل المؤسسات بتزويد المراكز الإحصائية بالمعلومات الدقيقة الخاصة بها وذلك بشكل دوري منتظم .

إن التمثيل البياني للسلاسل الزمنية التي تعكس تطور المؤشرات الاقتصادية مع الزمن في البلدان النامية يظهر بوضوح تلك التذبذبات و غالبا يكون خط المنحنى الفعلي معقدا لما يحويه من تذبذبات مختلفة تتم بوتائر وبلحظات زمنية متعددة وتستمر لفترات من الزمن متفاوتة وهذه التذبذبات غير المنتظمة ليست ذات أهمية لإظهار وتحديد المنحنى الأساسي بل على العكس إن وجودها يخفي المنحنى الأساسي وبالتالي هذا يعيق عملية إظهاره وتحديده ، إضافة إلى وجود هذه التذبذبات و الانحرافات غير المنتظمة تزيد التشتت وتضعف الدقة بتحديد المعالم في النموذج الرياضي للمنحنى الأساسي وبالتالي دقة التنبؤ .

إن عملية التصفية تتم بطرق مختلفة و تعتمد أساساً على استخدام طرق وأساليب متنوعة بهدف تخليص السلاسل الزمنية من التموجات والتذبذبات غير المنتظمة والتي من جراء استخدامها يصبح المنحنى الممثل للسلسلة اقل تموجا .

وفيما يلي أهم الطرق التي تستخدم لعملية التصفية

1- طريقة الأوساط المتحركة

تعتبر إحدى الطرق الشهيرة و أبسطها لتسوية السلاسل الزمنية ومن جراء تطبيقها يمكن التخلص لحد ما من التأرجحات غير المنتظمة و الحصول على سلسلة جديدة تعكس تأثير العوامل الأساسية الثلاثة (قوى الاتجاه العام – القوى الموسمية – القوة الدورية). وهنا نميز نوعين من الأوساط:-

1.1 – الأوساط المتحركة الخطية أو البسيطة.

وهذه الطريقة تعتمد على استبدال القيم الفعلية للسلسة الزمنية بمتوسطها الحسابية البسيطة، وان عدد قيم السلسة التي تؤخذ في حساب الوسط المتحرك تتعلق بطول الفترة التي تتخذ أساساً في حساب المتوسطات المتحركة (فترة الوسط المتحرك).

إن عدد حدود السلسة الزمنية التي تشملها فترة الوسط المتحرك يمكن أن يكون فردياً أو زوجياً. عندما يكون عدد الحدود فردياً ($m=2p+1$ ، حيث p عدد صحيح موجب) فغن قيمة الوسط المتحرك تقع في منتصف الفترة $\overline{y}_{p+1}$. أما عندما يكون عدد الحدود زوجياً فإن

قيمة الوسط المتحرك تقع بين أقرب قيمتين إلى منتصف فترة الوسط، وهنا يمكن من جديد أخذ الوسط الحسابي للوسطين المتحركين للحصول على قيمة تقابل تاريخ معين.

إن استخدام الوسط المتحرك أكثر سهولة، إذا كان عدد الحدود التي تشملها فترة الوسط المتحرك فردياً أي m=2p+1 . وتعطي حدود المتوسطات المتحركة بالعلاقة الرياضية التالية:

$$l=p+1,\ldots..,n-p$$

$$\bar{y}_l = \frac{1}{2p+1}\sum_{t=l-p}^{l+p} y_t \quad ، \qquad (1.1.1)$$

حيث أن:-

$\bar{y}_t$: قيمة الوسط المتحرك.

y_t : القيمة الفعلية لحد السلسة في اللحظة (t).

P : عين من طول فترة الوسط المتحرك (m)، بحيث $p = \frac{m-1}{2}$

1.2 – الأوساط المتحركة غير الخطية أو المرجحة.

في حالات عدة إن التسوية باستخدام الوسط المتحرك البسيط تبدو قوية، بحيث تختفي الجزيئات الهامة والخاصة الضرورية من أجل التحليل الاقتصادي، بالإضافة إلى أنه عند استخدامنا الوسط المتحرك البسيط نعتبر جميع حدود السلسة الزمنية متكافئة (لها نفس الوزن)، أن

أننا نعتبر منحنى تطور الظاهرة ضمن فترة الوسط المتحرك البسيط يمثل بمعادلة خط مستقيم $y_t = a + bt$ وعندئذ $\bar{y}_t = a$.

لذلك إذا كان معلوماً إن تطور الظاهرة ضمن فترة الوسط المتحرك تمثل بمنحنى غير خطي. فعندئذ لابد من اتباع أسلوب أدق وهو أسلوب الأوساط المتحركة غير الخطية أو المرجحة، وهي تعتمد على نفس فكرة ألا وساط المتحركة البسيطة، بحيث يعطي لكل حد من حدود السلسة الزمنية ضمن فترة الوسط المتحرك أوزاناً تتوقف على بعد الحد عن مركز مجال التسوية.

فإذا كانت قيم السلسة الزمنية الفعلية داخل كل فترة توصف بدالة كثيرة الحدود من الدرجة (k=2,3,....) k:

$$\bar{y}_t = a + \sum_{i=1}^{k} b_i t^i \qquad (1.2.1)$$

حيث (t) ترتيب الحد ضمن فترة الوسط المتحرك، بحيث t=...,-2,-1,0,1,1,2,..... وتعيين المعالم (Parameters) a, b_i يتم وفق طريقة المربعات الصغرى (OLS)، فمثلاً عندما تكون K=2 فإن الدالة تكون بالصورة التالية:-

$$y_t = a + b_1 t + b_2 t^2 \qquad (1.2.2)$$

وقيمة حد التسوية في النموذج المختار هي المعلمة a ، وهي عبارة عن الوسط الحسابي المرجح لحدود السلسة الفعلية الواقعة في

فترة الوسط المتحرك الفروض وبالإمكان الحصول على قيمة a وذلك باستخدام طريقة المربعات الصغرى (OLS) على النموذج المختـــار، حيث a تساوي:-

$$\alpha = \frac{\sum_{-p}^{+p} y_t \sum_{-p}^{+p} t^4 - \sum_{-p}^{+p} t^2 y_t \sum_{-p}^{+p} t^2}{n\sum_{-p}^{+p} t^4 - \left(\sum_{-p}^{+p} t^2\right)^2} \qquad (1.2.3)$$

ولحساب الأوساط المتحركة وفق هذه الصيغة مـــن أجـــل (m) عـــدد فردي نورد الأمثلة التالية:-

m=5

$$\bar{y}_t = \frac{1}{35}(-3y_{t-2} + 12y_{t-1} + 17y_t + 12y_{t+1} - 3y_{t+2})$$

m=7 (1.2.4)

$$\bar{y}_t = \frac{1}{35}(-2y_{t-3} + 3y_{t-2} + 6y_{t-1} + 7y_t - 3y_{t+2} - 2y_{t+3})$$

من الصيغتين السابقتين يمكن ملاحظة أن الأوزان متناظرة بالنسبة للحد المركزي y_t. وبالإمكان الحصول على m=9.

إن من عيوب طريقة المتوسطات المتحركة أنها تفقد بعض القيم في بداية السلسلة ونهايتها، بالإضافة إلى أنها لا نستطيع تمثيلها بدالة.

2. طريقة الأوساط الآسية.

رأينا في طريقة الأوساط المتحركة لتسوية الزمنية، أن حـــدود السلسلة ضمن فترة الوسط المتحرك أعطيت أوزاناً متساوية (الوســـط

المتحرك البسيط) أو أوزاناً متساوية بالتناظر (الوسط المتحرك غير الخطي). فأحيانا لتصفية السلسلة الزمنية من تأثير العوامل العشوائية نتبع طريقة أخرى وهي طريقة الأوساط الآسية، وهذه الطريقة تفترض أن الوسط المناسب يتوقف إلى حد بعيد على القيمة الحالية.

وبشكل أضعف على القيمة السابقة. أي كلما كانت المشاهدات أقدم كلما كان تأثيرها على الوسط أضعف. وبهذا الشكل تأثير المشاهدات السابقة على الوسط يتناقص كلما ابتعدت عن الزمن التي يحسب من أجلها الوسط. وبالتالي لحساب الوسط الآسي يتبع نظام خاص للأوزان يتوقف على حداثة المعطيات. ويعطي الوسط الآسي بالعلاقة:

$$\tilde{y}_t = \alpha y_t + (1-\alpha)\tilde{y}_{t-1} \qquad (2.1)$$

حيث

$\tilde{y}_t$: الوسط الآسي في الزمن t

α :المعامل الذي يصف وزن الملاحظة الحالية عند حساب الوسط الآسي (معامل التسوية)، بحيث $0 < \alpha \leq 1$.

نلاحظ بناء على العلاقة (2.1) أن الوسط الآسي في الزمن (t) عبارة عن تركيب خطي في مقدارين: القيمة الحقيقية لحد السلسلة الموافقة للزمن (t) والوسط الآسي المحسوب من أجل الفترة السابقة.

لابد من الإشارة إلى أن عملية تصفية السلاسل الزمنية من تأثير العوامل العشوائية يمكن إجرائها على مرحلتين أو ثلاث. إن التطبيق

الصحيح باستخدام الطريقة المناسبة لكل تصفية عند معالجة السلاسل الزمنية تمكننا من تصفية السلسلة الزمنية من الانحرافات العشوائية المختلفة، وبالتالي الحصول على سلسلة زمنية تعكس تأثير العوامل المنتظمة المختلفة وأخيراً فإن عملية تحديد نوع المنحنى الأساسي للتطور تصبح سهلة إلى حد ما.

المراجع:-

1. تحليل السلاسل الزمنية والتنبؤ (دراسة تطبيقية على محصول القمح بالجماهيرية العظمى)، رسالة الإجازة العالية للطالب الحسين سالم كعيبه، إشراف الدكتور : عبد الحفيظ مصطفى، لسنة 1994ف.

2. Estimations the parameters of non – linear tine – series in random encironment.

3. رسالة الإجازة الدقيقة للطالب الحسين سالم كعيبه، إشراف Professor V.V Anisimov لسنة 1999ف.