AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University

المقارنة بين طريقتي انحدار اساس الشريحة و الشريحة التكعيبية في تقدير نموذج المعلمات المتغيرة زمني

نور عبد الكريم فياض(1) ا.م.د وفاء جعفر حسين(2)

nab393820@gmail.com wjaffer@uowasit.edu.iq

Abstract

This research involves the study of varying coefficient model and time-varying coefficient models, as they have attracted great interest in recent years. These methods were used as two methods, the Regression basis spline method and the cubic spline method. The simulation method is used for comparison. The nonparametric method shows the preference of the cubic spline method for penalized constraints by comparing three mathematical functions to represent variable parameters, different sample sizes, and different levels of model standard deviation.

Keyword : time-varying coefficient models , cubic spline ، Basic Functions ، smoothing parameter ،Regression basis spline

المستخلص

يتناول البحث دراسة نموذج المعلمات المتغيرة وكذلك نموذج المعلمات المتغيرة زمنيا لما لهما من اهتمام ملحوظ في السنوات الاخيرة وتعد هذه النماذج ذات كفاءة عالية في التحليل الاحصائي نتيجة لعدم فرض قيود على النموذج وبهذا يهدف البحث الى تقدير نموذج المعلمات المتغيرة زمنيا وكذلك باستخدام طرائق التقدير اللامعلمية، واجراء المقارنة بين تلك الطرائق إذ تم استخدام طريقتين هي طريقة انحدار اساس الشريحة و طريقة الشريحة التكعيبية، وقد تم استخدام اسلوب المحاكاة لأجل المقارنة بين الطرائق اللامعلمية بمقارنة ثلاث دوال رياضية لتمثيل المعلمات المتغيرة وحجوم عينات مختلفة ومستويات متباينة من الانحراف المعياري للأنموذج ، اذ تبين افضلية طريقة الشريحة التكعيبية لوجود حد الجزاء.

المصطلحات الرئيسية للبحث : نموذج المعلمات المتغيرة زمنيا ، الشريحة التكعيبية ، دوال الاساس ، معلمة التمهيد ، انحدار اساس الشريحة

1. المقدمة

لتحليل الظواهر الاقتصادية والمالية والبيئية يتطلب ذلك بناء نموذج مناسب والطريقة الاحصائية لبناء نموذج هي تحليل الانحدار للتنبؤ بمتوسط متغير عشوائي اعتمادا على قيم متغير عشوائي واحد او اكثر فهو يبحث على ايجاد العلاقة بين المتغيرات التوضيحية ومتغير الاستجابة حيث عملية بناء نموذج تستند الى ظروف والعوامل المحيطة بالظاهرة بهيئة صيغة رياضية وقد قام الباحثون في بناء تلك الصيغة بشكل مبسط بحيث يمثل الظاهرة ويستند على العلاقات بين عواملها وكذلك ايجاد افضل الطرائق الرياضية لحل نموذجها ولدراسة اي ظاهرة او البحث عنها تكون في كيفية الاعتماد على التحليل المناسب الذي يساعدنا في اتخاذ العديد من القرارات التي تؤدي لحل هذا الظاهرة وفهم الظاهرة يمثل المرحلة الاولى التي تساعدنا في جمع المعلومات ذات الصلة بها وتمثيلها بطريقة ذات معنى لكي يتسنى لنا الكشف عن العلاقات من بين المعلومات المختلفة فضلاً عن إنَّ النماذج المعلمية تعد أحد أهم أدوات تحليل البيانات ، لأنها فعالة ويمكن تفسيرها بسهولة لكن في بعض الاحيان لا يتحقق احد الفروض الخاصة بل نموذج المعلمي الذي ينص على خطية العلاقة بين المتغيرات التوضيحية ومتغير الاستجابة وبسبب عدم تحقق هذا الافتراض سوف يؤدي الى مقدرات متحيزة وغير واقعية . لذلك لجأ الباحثون الى نماذج اخرى هي نماذج اللامعلمية التي تتميز بعدم وجود اي قيود عند صياغة النماذج لكن هذا النماذج ايضا تعاني من مشاكل ومن بين هذا المشاكل مشكلة البعدية و المقدرات تكون ذات تباينات كبيرة ولتفادي جميع المشاكل السابقة تم استخدام انموذجات المعلمات المتغيرة في تمثيل الظاهرة تحت الدراسة الذي يعد من الأنموذجات شبه المعلمية التي تعد بديلاً عن الانموذجات المعلمية التي تهمل الطبيعة الديناميكية اذ تسمح للمعلمات أن تتغير اعتمادا على المتغيرات التوضيحية ومن ثــمَّ فإنها تتمكن من استكشاف الانماط الديناميكية التي تتميز بها اغلب البيانات.

1. الجانب النظري

(1-2) نموذج المعلمات المتغيرة

ان صيغة أنموذج الانحدار الخطي المتعدد تكون كالاتي :

$$y_i = \sum_{j=1}^{p} \beta_j X_{ij} + e_i \quad , i = 1,2,\cdots,n \quad j = 1,2,\cdots,p \quad \cdots (1)$$

حيث ان :

y : موجة من درجة $(n \times 1)$ لمشاهدات المتغير المعتمد .

X :مصفوفة من الدرجة $(n \times p)$ لمشاهدات المتغير المستقل .

$\beta = (\beta_1, \cdots, \beta_p)$: موجه من الدرجة $(p \times 1)$ للمعالم المطلوبة تقديرها .

e: موجه من الدرجة $(n \times 1)$ للخطأ العشوائي بمتوسط صفر وتباين σ^2 .

يعتبر الأنموذج المبين في الصيغة (1) هو أنموذج معلمي يعتمد على عدة افتراضات في تحليل العلاقة بين المتغيرات ومن بين تلك الافتراضات خطية العلاقة بين المتغير المعتمد و المتغيرات المستقلة [8] .

وفي حالة عدم تحقق هذا الافتراض سيؤدي هذا الى مقدرات متحيزة و غير كفوءة لذا كان من المناسب ايجاد اسلوب جديد لدراسة العلاقة بين المتغيرات حيث اعتمد الباحثون في الاونة الاخيرة على النمذجة اللامعلمية التي لا تضع اي قيود او شروط عند صياغة الأنموذج ، ويمكن التعبير عن الأنموذج اللامعلمي بالصيغة الاتية :

$$y_i = m(x_i) + e_i \quad i = 1,2, \dots, n \qquad \dots (2)$$

حيث ان :

Y: هو متغير الاستجابة .

m(x) : الدالة المجهولة المراد تقديراها التي تتميز انها لا تحتوي على معالم .

X : قيم المشاهدات للمتغير المستقل .

e : الخطا العشوائي يتوزع توزيع طبيعي بمتوسط صفر وتباين σ^2 .

ولقد لوحظ أنَّ هذا الاسلوب ايضا لا يخلو من العديد من المشاكل ومن ابرزها مشكلة الابعاد العالية (curse of dimensionality)، تحدث هذه المشكلة اذا كان عدد المتغيرات المستقلة اكبر او يساوي حجم العينة [9] **.**

وكذلك ظهور مشكلة التباين العالي للمقدرات، ولتجنب المشاكل السابقة جميعها تم استعمال أنموذج المعاملات المتغيرة (varying Coefficients model)الذي يعد من الأنموذجات شبه معلمية (semiparametric models) وهي الأنموذجات التي تعد في الوقت الحاضر من اكثر الاساليب استخداما حيث تغلب على اسلوب الأنموذجات المعلمية (parametric model) والأنموذجات اللامعلمية (nonparametric models) وتكون صيغة أنموذج المعاملات المتغيرة (varying Coefficients model) كالاتي [2]:

$$Y_i = \sum_{j=1}^{p} \beta_j(U_i) X_{ij} + \varepsilon_i \qquad , i = 1,2, \cdots n \qquad \cdots (3)$$

حيث ان :

Y: هو موجه من الدرجة $(n \times 1)$لمشاهدات المتغير المعتمد .

X: مصفوفة من الدرجة $(n \times p)$ لمشاهدات المتغير المستقل .

$\beta(U) = \left(\beta_1(U), \cdots, \beta_P(U)\right)$: متجه من الدوال اللامعلمية المجهولة المراد تقديرها . ويمكن ملاحظة ان $(U_1, U_2, \cdots U_n)$ التي تكون قيمتها U=Xβ تتغير مع المعاملات $\left(X_{i1}, X_{i2}, \cdots, X_{ip}\right)$ من خلال الدوال $\beta(U) = \left(\beta_1(U), \cdots, \beta_P(U)\right)$ وان اعتماد $\beta_j(U), j = 1,2,\cdots,p$ على U هذا يشير الى حالة خاصة من تفاعل بين X و U وهذا التفاعل يجعل دوال $\beta_j(U)$ ليست ثابتة و متغيرة (Varying) حيث تكون الفكرة الاساسية لهذه الأنموذجات هي السماح لمعاملات الانحدار بالتغير بسهولة (تفاعل) مع متغير اخر في حين أن الأنموذج المبين في الصيغة (1) تكون فيه معاملات المتغيرات التوضيحية ثابتة [13] [2] .

(2-2) نموذج المعاملات المتغيرة زمنيا Time- varying Coefficients Model

غالبا يكون $(U_i, i = 1,2,\cdots,n)$ هو متغير كأن يكون الزمن . وعلى فرض ان البيانات تتكون على شكل قياسات مكررة للمتغيرات $\{Y, X_1, \cdots, X_P\}$ عبر نقاط زمنية $(t_1, t_2, \cdots, t_n)$ في هذه الحالة سوف يكون أنموذج المعاملات المتغيرة (Time- varying Coefficients Model) بالصيغة الاتية :

$$y_i = \sum_{j=1}^{p} \beta_j(t) X_{ij} + \varepsilon_i, i = 1,2,\cdots n \qquad \cdots (4)$$

يسمى الأنموذج اعلاه أنموذج الخطي الديناميكي المعمم (dynamic generalized linear model) إذ يعد هذا الأنموذج حالة خاصة من أنموذج المعاملات المتغيرة (Varying Coefficients Models). و تستخدم بشكل خاص في الدراسات الطولية من خلال السماح باستكشاف التأثير الذي يتغير بمرور الزمن للمتغيرات التوضيحية على الاستجابة [5] .

(3-2) الشريحة التكعيبية cubic spline

وفي حاله تقدير المعلمات المتغيرة زمنيا (time-varying parameters) في نموذج المعلمات المتغيرة زمنيا (Time- Varying Coefficient Model) سيتم احتساب مجموع مربعات البواقي مضاف اليها حد يطلق علية حد الجزاء (roughness penalty) وكما موضح في الصيغة الآتية:

$$\sum_{i=1}^{n}\left[y_i - \sum_{j=1}^{p} X_{ij}\,\beta_j(t_i)\right]^2 + \lambda \int_a^b \left\{\beta_j{}^{(2)}(t_i)\right\}^2 dt \quad i = 1,2,\dots,n \;, j = 1,2,\dots,p \qquad \cdots (5)$$

λ: معلمة التمهيد

$\beta_j(t)$: الدالة المجهولة المراد تقديراها .

حيث نقوم بتصغير المعادلة (5) للحصول على قيمة المقدر $\hat{\beta}_j(t_i)$ إذ يمثل الجزاء الاول من المعادلة مجموع مربعات البواقي (RSS).اما الجزاء الثاني من المعادلة اعلاه يمثل حد الجزاء (roughness penalty) الذي يحتوي على معلمة التمهيد (λ) ويكون لها دور مهم في تحديد قيمة حد الجزاء غير الممهد اي عندما ($\lambda \to 0$) هذا يعني عدم ظهور حد الجزاء في المعادلة وعندما ($\lambda \to \infty$) هذا يعني أنَّ معلمة التمهيد كبيرة جدا وهذا ينتج منحنى ثابت الانحدار الخطي [4] .

وكما لهذه المعلمة دور مهم في التوازن بين التحيز والتباين .عند الفترة $[a, b]$ لنفرض ان نقاط زمن التصميم التي تستخدم كعقد هي ($T_1, T_2, \dots, T_P$) الدالة β هي شريحة تكعيبية في حالة وجود الشروط الاتية [10] :

- في كل فترة جزائية $[T_j, T_{j+1}]$ تكون الدالة متعددة الحدود .
- تكون المشتقة الاولى والمشتقة الثانية والدالة (β) هي دوال مستمرة .

نفرض ان β هو عبارة عن متجه المعلمات $\beta=[\beta_1, \beta_2, \dots, \beta_p]'$ من الدرجة (pX1) ومن خواص الشريحة التكعيبية $\beta^{(2)}(a)=0$ و $\beta^{(2)}(b)=0$ علما بان

$\beta^{(2)}(a)$: المشتقة الثانية عند الحد a .

$\beta^{(2)}(b)$: المشتقة الثانية عند الحد b .

عندئذ يكون المتجه S هو متجه المشتقة الثانية للموجه الدوال β حيث ان

$$S=[S_2, \dots, S_{p-1}]^T$$

من الدرجة ((p-2)X1) نفرض ان $\beta_j = \beta_j(t_i)$ و $S_j = \beta_j^{(2)}(t_i)$ i=1,2,…,n j=1,2,…,p

حيث ان شكل المنحنى β يعتمد على المتجهين اعلاه (S_j, β_j) بواسطة اثنين من المصفوفات نستطيع تعريف

المتجهات ولتكن هذه المصفوفات (D,E)

ليكن a متجه الفرق بين T_j و T_{j-1} وكالاتي

$$a_{j=T_{j+1}-T_j} \qquad j=1,2,\dots,p$$

نفرض أنَّ E مصفوفة بدرجة(px(p-2)) عناصرها e_{ij} و تحسب كالاتي :

$$e_{(j-1,j)} = a^{-1}(j-1)$$

$e_{(j,j)} = a^{-1}(j-1) - a_j^{-1}$

$e_{(j+1,j)} = a_j^{-1}$

$e_{(i,j)} = 0 \ \forall \ |i-j| \geq 2 \quad i=1,2,\ldots,n \quad j=2,\ldots,p-1$

وكذلك نفرض أنَّ (D) مصفوفة متماثلة (symmetric) من الدرجة (p-2) × (p-2) تحسب عناصرها كالاتي :

$d_{jj} = (a_{j-1} + a_j)/3 \quad , \quad j = 2,3 \ldots, p-1$

$d_{(j,j+1)} = d_{j+1,j} = \frac{a_j}{6} \quad , \quad j = 2,3,\ldots,p-2$

$d_{ij} = 0 \ \forall \ |i-j| \geq 2 \ , \quad i=1,2,\ldots,n \quad j=2,\ldots,p-1$

وبما أنَّ (D^{-1}) موجود نستطيع تعريف المصفوفة L كالاتي :

$$L = E \ D^{-1} E^T \quad \cdots (2-15)$$

حيث وحسب رأي الباحثين (Green and silverman)[10] اذا تحقق الشرط التالي $E^T \beta = DS$ يكون المتجهان S, B شرائح تكعيبية (cubic smoothing spline) وعند توفر الشرط اعلاه يمكن حساب مربع المشتقة الثانية للدالة β وكما يلي :

$$\int_a^b (\beta^{(2)}(t))^2 dt = S' D \ S = B' L \ B$$

ويمكن اعادة كتابة الصيغة (5) كالاتي :

$$[y - x\beta]'[y - x\beta] + \lambda\beta' L\beta \quad \cdots (6)$$

حيث يمثل مقدر الشريحة التكعيبية $\hat{\beta}$ وكالاتي[12] :

$$\beta^{\wedge}_{\lambda} = (x^T x + \lambda L)^{-1} X'Y \quad \cdots (7)$$

$$\beta^{\wedge}_{\lambda} = R_{\lambda} x' y \quad \cdots (8)$$

$\beta^{\wedge}_{\lambda}$: مقدر الشريحة التكعيبية الذي يمثل بل متجه الاتي :

$$\beta^{\wedge} = \left[\beta^{\wedge}{}_{\lambda}(t_1), \dots, \beta^{\wedge}{}_{\lambda}(t_n)\right]'$$

R_{λ} : مصفوفة التمهيد من الدرجة (p×p) .

وبالاعتماد على معلمة التمهيد λ وعلى نقاط العقد $(T_1, \dots, T_P)$ يتم حساب مصفوفة (R_{λ}) .

(3-2-1) اختيار معلمة التمهيد *smoothing parameter selection*

تلعب معلمة التمهيد (λ) Smothing Spline دوراً مهماً في تجانس الشريحة بشكل اكثر سلاسة حيث أن (λ) لها تأثير قوي في التحيز و التباين حيث إنَّ نقصان معلمة التمهيد يؤدي الى قلة التحيز وزيادة التباين والعكس صحيح .وللحصول على قيمة جيدة للمعلمة (λ) من خلال تطبيق طرائق اختيار معلمة التمهيد الآتية :

1- معيار العبور الشرعي العام (Generalized Cross-Validation)

تم اقتراح معيار العبور الشرعي العام (GCV) بواسطة Wahba (1 977) and Craven and Wahba (1979)

حيث نستطيع أن نحصل على (GCV)من صيغة معيار العبور الشرعي (CV) ويمكن تمثيل معيار العبور الشرعي بالصيغة الاتية :

$$CV_{\lambda} = n^{-1} \sum_{i=1}^{n} \left[\frac{y_t - y^{\wedge}{}_t}{1 - r_{tt}}\right]^2 \dots (9)$$

r_{tt}: تمثل العنصر t من عناصر القطر الرئيسي للمصفوفة التمهيد R_{λ} .

$y_t - y^{\wedge}{}_t$: انحرفات المشاهدات .

من الصيغة (9) يمكننا الحصول على معيار العبور الشرعي العام (GCV)وذلك عن طريق استبدال عناصر قطر الرئيسي r_{tt} لمصفوفة الممهد (R_{λ}) بواسطة

$$n^{-1} \sum r_{tt} = n^{-1} tr(R_{\lambda}) = n^{-1}\, df$$

وتكون صيغة معيار (GCV) كالاتي :

$$GCV_{\lambda} = \frac{n^{-1} \sum_{i=1}^{n} \left[y_t - y^{\wedge}{}_t\right]^2}{\left[1 - tr\frac{(R_{\lambda})}{n}\right]^2} \quad \dots (10)$$

$$GCV_{\lambda} = \frac{n^{-1}\, SSE}{\left(1-\frac{df}{n}\right)} \quad \ldots (11)$$

المجموع المتبقي للمربعات حول $\hat{\beta}$ مقسوما على عامل التصحيح $n[1-tr(R_{\lambda})]^2$ وعلية تكون الفكرة الاساسية لتحديد قيمة معلمة التمهيد بطريقة (GCV) هو *تصغير* مجموع مربعات البواقي لا يجاد القيمة المثلى لمعلمة التمهيد [12] .

(2- 4) انحدار اساس الشريحة Regression Basis spline (B-spline)

قبل التطرق الى هذه الطريقة سيتم تعريف معنى الشرائح من نوع (Basis spline (B-spline)

(4-2-1) الشرائح من نوع B-spline :

تم التعامل مع الشرائح B-spline من قبل Nikolai Lobachersk في مطلع القرن التاسع عشر وهي عبارة عن دالة متعددة الحدود من الدرجة p في المتغير x حيث يشير الحرف B الى اختصار Basis ان هذا النوع من الشرائح متعددة حدود تكون مرتبطة فيما بينها عن طريق نقاط ربط تسمى عقد knots ويجب أن تكون هذه العقد اصغر من حجم العينة وايضا يجب أن يكون عددها صحيحا .إن مفهوم الشريحة spline هو عبارة عن منحنى متعدد الحدود مستمر يستخدم لتقريب حل مشكلة رياضية و يعتمد هذا المنحنى على العلاقة بين دالة الاساس Basis function ونقاط التحكم control point .ولقد استخدمت هذه الشرائح في العديد من المجالات مثل الرياضيات والهندسة وعلوم الكمبيوتر ولقد اصبح استخدام هذا النوع من الشرائح شائعا بصورة كبيرة في السنوات الاخيرة ومن مميزات الشريحة (Spline) أنَّها لا تحتاج حساب المشتقة ولا تشترط افتراضات لا سيما في حل المعادلات لهذا تختصر الجهد والوقت عند الحساب والتعامل معها [3].

وللتقدير بطريقة انحدار اساس الشريحة (Regression Basis spline) يتم فيها تقسيم مجال قيم المتغير التوضيحي المتمثل بالفترة [a , b] تسمى انحدار الشريحة B حيث إنَّ نقاط التقسيم تسمى بالعقد والتي يمكن التعبير عنها كما يأتي:

$$a = T_1, T_2, \ldots\ T_m, T_{m+1} = b$$

حيث تقسم هذا العقد مجال المتغير التوضيحي والمتمثل بالفترة[a , b] الى m من الجور الموضعي (الفترات الفرعية) وهذا يعني يتم استخدام متعدد الحدود من رتبة معينة يمكن تحديدها مسبقا بين اي عقدتين متجاورتين وعندئذ يتم بناء ما يسمى بانحدار الشريحة باستعمال دوال الاساس [11] [6]

$$N_{j,0}(t) = \begin{cases} 1 & x \le t < x_{j+1} \\ 0 & o.w \end{cases}$$

$$N_{j,p+1}(t) = \frac{t-x_j}{x_{j+p}-x_j}\, N_{j,p}(t) + \frac{x_{j+p+1}-t}{x_{j+p+1}-x_{j+1}}\, N_{j+1,p}(t) \quad \cdots (12)$$

تم استخدام مصطلح (B-Spline) لاول مرة في عام 1946 من قبل Schoenberg's .حيث اجرى تحويلاً على البيانات متساوية الابعاد وذلك بتقريبها بواسطة الدوال التحليلية (دوال الاساس) Basis ويمكن التعبير عن اي شريحة من الدرجة (P) بسلسلة من العقد المعطاة يمكن التعبير عنها كتركيبة خطية وحيدة من B-spline من الدرجة نفسها .ويتم التعبير عن (B-Spline) كالاتي [7]:

$$B(t)=\sum_{j=0}^{k-1} \alpha_j \ N_{j,p}(t) \cdots (13)$$

$$\beta(t)= [\alpha_0 \quad \dots \quad \alpha_{k-1}] \begin{bmatrix} N_{0,P}(t) \\ \dots \\ N_{k-1,p}(t) \end{bmatrix} \quad \cdots (14)$$

$$\beta (t)=\alpha N^T \qquad \cdots \ (15)$$

$\beta(t)$: موجه الشريحة التي تعرف كمنحنى B-Spline .

$N_{j,p}(t)$: مصفوفة دوال الاساس (B-Spline) من الدرجة (k-1×p) .

α : موجه معاملات دوال الاساس الثابتة وايضا تسمى نقاط التحكم لأنها تتحكم في شكل المنحنى B-Spline من الدرجة (1×k-1) .

عدد نقاط التحكم لمنحنى (B-Spline) تساوي k من دوال الاساس لـ (B-Spline) وكذلك تعتمد على درجة المنحنى P وعدد العقد m+1 بواسطة العلاقة الأتية [7]:

$$K = m - p \qquad \cdots \ (16)$$

K: عدد دوال الاساس B-Spline .

m: عدد العقد .

P: درجة المنحنى .

وبالرجوع الى معادلة (12) في حالة وجود العقد المتكررة يكون قيمـــة المقام صفر، في هذا الحالة يفترض أن يكون (0/0) يكون صفرا . تكون B-Spline متعددة الحدود على سبيل المثال شرائح B من الدرجة الاولى هي شريحة خطية Linear spline ومن الدرجة الثانية quadratic spline ومن الدرجة الثالثة cubic spline مجموع (B-Spline) في اي نقطة داخل هذا المجال يساوي واحد [7].

$$\sum_{j=0}^{k-1} N_{j,p}(t) = 1 \qquad X_p \le t < X_{m-p}$$

يمكن التعبير عن المعادلة رقم (15) كالاتي :

$$\beta(t)= [\alpha_0 \quad \cdots \quad \alpha_{k-1}] \begin{bmatrix} N_{0,P}(t_1) & \cdots & N_{k-1,P}(t_1) \\ \vdots & \ddots & \vdots \\ N_{0,P}(t_n) & \cdots & N_{k-1,P}(t_n) \end{bmatrix} \quad \dots (17)$$

α : متجه معاملات دوال الاساس .(المعاملات الثابتة) من الدرجة (1× k-1)

N: مصفوفة التصميم من الدرجة k-1×p .

وبتعويض المعادلة (13) في معادلة (4) وباستخدام معيار المربعات الصغرى (Ordinary least square) نحصل على مقدر معاملات دوال الاساس الثابتة والذي يرمز له ($\widehat{\alpha}$) وكالاتي [1] :

$$\min_\alpha \sum_{i=1}^{n} \left(y_i - X_{ij}\, \alpha_j\, N_{j,p}(t_j)\right) \qquad \cdots (18)$$

حيث يمكن كتابة المعادلة اعلاه كالاتي

$$\min_\alpha \sum_{i=1}^{n} (y_i - \alpha V_i)^T\, (y_i - \alpha V_i) \qquad \cdots (19)$$

$$y_i = (y_1 \quad , \dots, \quad y_n)^T$$

$$V_{ij} = {X^T}_{ij}\, N_{j,p}(t)$$

حيث يكون V من الدرجة (n×k-1) .

$$V_{ij} = [V_{i0}, V_{i1}, \cdots V_{ik-1}]^T$$

عندئذ تكون الصيغة التقديرية لمتجه معاملات دوال الاساس الناتجة عن تصغير معيار المربعات الصغرى الاعتيادية والمبينة كالاتي :

$$\widehat{\alpha} = (V^T V)^{-1} V'\, y \qquad \cdots (20)$$

وبتعويض قيمة ($\widehat{\alpha}$) يكون مقدر B-Spline كالاتي :

$$\hat{\beta}(t)=\hat{\alpha} \quad N(t) \qquad \cdots (21)$$

وعند تعويض تقدير $\hat{\beta}(t)$ في معادلة (4) نحصل على متجه الاستجابة y المقدر وكما يلي :

$$\hat{y} = \sum_{j=1}^{p} \hat{\beta}(t)\, X_{ij} \quad \cdots (22)$$

3. الجانب التجريبي

(1- 3) وصف تجارب المحاكاة

لقد اعتمدنا في تجارب المحاكاة على ثلاثة حجوم عينات ($n_1 = 30$, $n_2 = 100$, $n_3 = 250$) وذلك في توليد البيانات التي تخص المتغيرات العشوائية التي تتضمن في أنموذج المعاملات المتغيرة زمنيا ويمكن تلخيص عملية المحاكاة كالاتي :

(2-1- 3) توليد المتغيرات العشوائية

1. توليد المتغير التوضيحي x بمتوسط صفر وتباين σ^2
2. توليد متغير الزمن t بمتوسط صفر وتباين σ^2
3.توليد الخطأ العشوائي e بمتوسط صفر وب ثلاث مستويات للانحراف المعياري (1, 0.5 , 0.01)
4.تعويض المتغير التوضيحي x ومتغير الزمن t والخطأ العشوائي e ولجميع المستويات في أنموذج المعاملات المتغيرة زمنيا للحصول على متغير الاستجابة y ولكل مستوى من مستويات الانحراف المعياري وكذلك لكل أنموذج من أنموذجات المتغيرة زمنيا .

(3-1-3) أنموذجات المستعملة في تجارب المحاكاة

تم الاعتماد على عدة دوال رياضية مختلفة لتمثيل المعاملات المتغيرة للمتغير التوضيحي وكما مبين ادناه:

1. $2\mathrm{Sin}\,(2\pi\, t)$
2. $\exp\,(2t - 1)^3$
3. $0.72+9t$

وبالاعتماد على هذه الدوال يتم صياغة نموذج المعلمات المتغيرة حيث تعوض هذه الدوال في نموذج المعلمات المتغيرة زمنيا و كما يأتي :

$$y = (2\mathrm{Sin}\,(2\pi\, t))x + e$$

$$y = (\exp\,(2t - 1)^3)x + e$$

$$y = (0.72 + 9t)x + e$$

(2- 3) نتائج تجارب المحاكاة

سنوضح في هذا الجانب نتائج تجارب المحاكاة التي اجريت لحجوم العينات الثلاثة و مستويات الانحراف المعياري المعتمدة وباستخدام سبعه دوال رياضية مختارة لتمثل الدوال المتغيرة ا أذ اجريت هذه تجارب لمقارنة طرائق التقدير اللامعلمية و كالاتي :

جدول (1) يبين قيم معيار (MAPE) لجميع حجوم العينات و مستويات الانحراف المعياري لمقارنة للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الاول

		Standard deviation levels for the boundary of error		
N	nonparametric methods	1	1.5	2
50	RBS	1.142	2.709478	4.360352
	CS	1.102208	2.2368212	4.2131534
80	RBS	1.113219	2.464955	4.357315
	CS	1.033021	2.023041	3.761702
150	RBS	1.063232	2.441591	4.282632
	CS	1.0012122	2.012939	3.213212

جدول (2) يبين قيم معيار (MAPE) لجميع حجوم العينات و مستويات الانحراف المعياري لمقارنة للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الثاني

		Standard deviation levels for the boundary of error		
N	nonparametric methods	1	1.5	2
50	RBS	1.141705103	2.709533087	4.361579
	CS	1.093264123	2.541208986	4.012391
80	RBS	1.113413985	2.464553338	4.360539
	CS	1.034751984	2.31096556	3.876511
150	RBS	1.063236856	2.44136103	4.281660
	CS	1.0213503321	1.858556176	3.632029

جدول (3) يبين قيم معيار (MAPE) لجميع حجوم العينات و مستويات الانحراف المعياري لمقارنة للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الثالث

		Standard deviation levels for the boundary of error		
N	nonparametric methods	1	1.5	2
50	RBS	1.141968992	2.709474818	4.360479
	CS	1.432187912	2.678943156	4.316578
80	RBS	1.113237857	2.464903317	4.357647
	CS	1.021589033	2.129072091	3.760654
150	RBS	1.063229714	2.441565238	4.282520
	CS	1.00654741	2.094272771	3.290191

الاستنتاجات

1. نلاحظ بزيادة حجم العينة يقل قيمة معيار (MAPE) ولجميع الطرق وهذا يلائم الطريقة الاحصائية
2. نلاحظ أنه كلما زاد الانحراف المعياري يؤدي ذلك الى زيادة قيمة (MAPE)
3. نلاحظ أنَّ طريقة الشريحة التكعيبية (CS) هي الطريقة الأفضل لوجود حد الجزاء .
4. نوصي باعتماد طريقة (cubic Basis spline) في تقدير أنموذج المعاملات المتغيرة زمنيا
5. نوصي بإجراء دراسات مستقبلية لطرائق لامعلمية حصينة لتقدير نموذج المعاملات المتغيرة زمنيا بوجود بيانات ملوثه

المصادر

1. Ahkim, M., Gijbels, I., & Verhasselt, A. (2017). Shape testing in varying Coefficients models. Test, 26(2), 429-450.
2. Fan, J., & Zhang, W. (2008). Statistical methods with varying Coefficients models. Statistics and its Interface, 1(1), 179.
3. Gálvez, A., & Iglesias, A. (2013). Firefly algorithm for explicit B-spline curve fitting to data points. Mathematical Problems in Engineering, 2013..

4. Green, P. J. and Silverman, B. W. (1994). “Nonparametric Regression and Generalized Linear Models: a Roughness Penalty Approach”. Chapman and Hall, London .
5. Hastie, T., & Tibshirani, R. (1993). Varying-Coefficients models. Journal of the Royal Statistical Society: Series B (Methodological), 55(4), 757-779.
6. Johnson, R.W., (2005). “A B-spline Collocation Method for solving the Incompressible Navier- Stokes Equations Using an ad hoc Method: the Boundary Residual Method,” Computers& Fluids 34: 121-149
7. Karagöz, R. (2020). Tensor Network B-splines for high-dimensional function approximation.
8. Kiebel, S., & Holmes, A. (2011). The general linear model. Statistical Parametric Mapping: The Analysis of Functional Brain Images, 101-125.
9. Rigollet, P., & Hütter, J. C. (2018). High dimensional statistics lecture notes. Accessed May, 2018.
10. Rodrıguez, G. (2001). Smoothing and non-parametric regression. Princeton University.
11. Wang, B. &Miao, Z.(2014)." Comparative Analysis for Robust Penalized Spline Smoothing Method" Volume 2014, 11 pages
12. Wu, H., & Zhang, J. T. (2006). Nonparametric regression methods for longitudinal data analysis: mixed-effects modeling approaches. John Wiley & Sons.
13. Xue, L., & Qu, A. (2012). Variable selection in high-dimensional varying-Coefficients models with global optimality