الطرائق اللامعلمية في تقدير أنموذج المعاملات المتغيرة دراسة مقارنة

أ.د. ظافر حسين رشيد* م.م. حسام عبد الرزاق رشيد**

المستخلص

ان المشكلة الرئيسيه لبحث ودراسة اي ظاهرة هي في كيفية اعتماد التحليل المناسب الذي يمكننا من الاعتماد على اساسه في اتخاذ العديد من القرارات التي ترتقي لحل تلك الظاهرة وان فهمنا للظاهرة يمثل المرحلة الاولى اذ سنتمكن من جمع كل المعلومات ذات الصله بها وتمثيلها بطريقة ذات معنى لكي يتسنى لنا رؤية العلاقات بين المعلومات المختلفه. من المعروف سابقاً ان قيم معلمات أنموذجات الانحدار الخطيه تكون ثابتة (Fixed) على مدى جميع المشاهدات علماً ان تلك المعلمات تكون غير معلومة وينصب الاهتمام على تقدير قيمها. حديثاً تم التوجه في بعض الحالات التطبيقيه على دراسة انموذجات خطية تكون معلماتها ذات قيم متغيرة بمعنى ان تلك المعلمات يمكن اعتبارها دالة تعتمد على احد المتغيرات المستقلة ان هذه الهيكلية في بناء الأنموذج بين المتغيرات المستقلة والمتغير المعتمد تعرف بأنموذجات المعاملات المتغيرة Varying (VCM) Coefficient Models)) وقد تم تقدير معاملات أنموذج المعاملات المتغيرة بأعتماد الطرائق اللامعلمية وهي طريقة النواة لمتعدد الحدود الموضعي (LPK) Local Polynomial Kernel)) وطريقة انحدار الشريحة (RS) Regression Spline)) اذ تم استخدام اسلوب المحاكاة لاجل المقارنه ووجد ان طريقة LPK هي الافضل .

مصطلحات البحث: أنموذج المعاملات المتغيرة، طريقة النواة لمتعدد الحدود الموضعي، طريقة انحدار الشريحه، معيار العبور الشرعي العام .

Abstract

The main problem to search and study any phenomenon is how the adoption of appropriate analysis that we can employ to take many decisions that elevate resolve this phenomenon, and our understanding of the phenomenon represents the first phase as it will be able to collect all the relevant information out and represented in a meaningful way between different information. Previously known that parameters of linear regression be fixed over all the observations and those parameters are unknown and the focus to estimate it. Recently was addressed in some cases to study linear models which the parameters are variable in the sense that those parameters can be considered as a function dependent on a independent variable. This structural building model between the independent variables and the dependent variable known as Varying Coefficient Models (VCM) . We estimate the varying coefficients by using nonparametric methods which is the Local Polynomial Kernel (LPK) and Regression Spline (RS), and by

* جامعة بغداد / كلية الادارة والاقتصاد .

** الجامعة المستنصرية / كلية الادارة والاقتصاد .

مقبول للنشر بتأريخ 2013/12/9

مستل من أطروحة دكتوراة

using simulation technique for comparison we found that the LPK method is the best.

Keywords: Varying Coefficient Model, Local Polynomial Kernel, Regression Spline, Generalized Cross Validation.

1. المقدمة

أنموذجات المعاملات المتغيرة (Varying Coefficient Models (VCM)) تعد بديل للأنموذجات التجمعية (Additive Models) كما انها تعتبر توسيع لامعلمي لأنموذجات الانحدار الخطية، على عكس الأنموذجات المعلمية التي تهمل الطبيعة الديناميكية للعلاقات بين المتغيرات فأن أنموذجات المعاملات المتغيرة تسمح لمعاملات الأنموذج بالتغير اعتماداً على متغيرات مستقلة معينة وبالتالي تمكنها من استكشاف الانماط الديناميكية التي تتصف بها البيانات وتمثيل العلاقات بشكل جيد وكفوء. لهذه الأنموذجات تطبيقات واسعة في العديد من المجالات كالاقتصادية والمالية والصحية والبيئية وغيرها. بالاضافة لأنموذجات المعاملات المتغيرة هناك الأنموذجات شبه المعلمية التي تبنى على اساس ان تأثير احد المتغيرات التوضيحية يكون متغيراً في حين بقية المتغيرات التوضيحية تكون تأثيراتها ثابتة[2][9][3] . ان المشكلة التي تعاني منها الطرائق اللامعلمية هي البعدية (الابعاد) (Curse of Dimensionality) تلك المشكلة تجعل الطرائق اللامعلمية ضعيفة عملياً وذلك عندما يكون هنالك اكثر من متغير توضيحي، ولتفادي هذه المشكلة سنعتمد الأنموذجات التجمعية (Additive) أو أنموذجات المعاملات المتغيرة (VCM) والتي تعد من الأنموذجات الشبه معلمية (Semiparametric) من ناحية الهيكلية . اقترح كل من Hastie وTibshirani في عام (1993) استخدام أنموذجات المعاملات المتغيرة العامة (General Varying Coefficient Models) في تمثيل عدة تطبيقات وأستخدم كل من Fan و Zhang في عامي (١٩٩٩ و ٢٠٠٠) طريقة المرحلتين (Two-Step) في تقدير معاملات أنموذج المعاملات المتغيرة ايضاً أستخدم كل من Cai و Fan و Yao في عام (٢٠٠٠) أنموذج المعاملات المتغيرة في بيانات السلاسل الزمنية غير الخطية وفي عام (٢٠٠٢) أقترح Huang وآخرون أستخدام تمهيد الشريحة في توفيق أنموذج المعاملات المتغيرة [10]. لتقدير معاملات أنموذج المعاملات المتغيرة يمكن اعتماد الطرائق اللامعلمية اذ يهدف البحث الى تناول طريقة النواة لمتعدد الحدود الموضعي (Local Polynomial Kernel (LPK)) وبتوظيف دالة نواة مقترحة ومقارنتها مع طريقة انحدار الشريحة (Regression Spline (RS)). تمت المقارنة بأستخدام المحاكاة فيما يخص تقدير الدالة التي بموجبها يعتمد معامل المتغير التوضيحي على متغير مستقل آخر.

2. الجانب النظري

1-2 أنموذجات المعاملات المتغيرة (Varying Coefficient Models (VCM))

ان صيغة الأنموذجات التجمعية المعممة (Generalized Additive Models ((GAM) والمبينة بالصيغة التالية التي تفترض ان الاستجابة عبارة تركيب خطي لمجموعة الدوال غير المعلومة [10].

$$\mathrm{E}(y) = a_1(x_1) + a_2(x_2) + \cdots + a_L(x_L) \qquad \cdots (1)$$

اذ ان $\mathrm{E}(y)$ يمثل توقع متغير الاستجابة y وان x_k ، $k = 1,2,\cdots,L$ هي المتغيرات التوضيحية وان a_k تمثل دوال غير محددة بدلالة المتغيرات التوضيحية (التي نريد تقديرها) غير ان هنالك حالات تكون فيها معاملات المتغيرات التوضيحية غير ثابتة وانما هي دالة تعتمد على متغير آخر مما يجعلها ذات قيما متغيرة والأنموذج الرياضي الانسب لوصف مثل هذه العلاقات يعرف بأنموذج المعاملات المتغيرة (Varying Coefficient Model (VCM)) وصيغته الرياضية تكتب كالاتي [4] [5] :

$$y_t = \sum_{k=1}^{L} a_k(u_t)\, x_{tk} + \varepsilon_t \quad , t = 1,2,\cdots,n \qquad \cdots (2)$$

اذ ان y_t ، $t = 1,2,\cdots,n$ يمثل متغير قيم الاستجابة وان x_{tk} ، $k = 1,2,\cdots,L$ هي المتغيرات التوضيحية المراد تقدير تأثيرها على متغير الاستجابة y اما $a_k(u_t)$ فتمثل معاملات المتغيرات التوضيحية في معادلة الانحدار والتي هي عبارة عن دالة في المتغير u وليست ثابتة و ان ε_t تشير الى الخطأ عند الزمن t وهو متغير عشوائي بمتوسط صفر وتباين ثابت σ^2. من الأنموذج (2) نلاحظ ان المتغيرات u تؤثر في المعاملات $a_k(u_t)$ الامر الذي يؤدي الى جعلها ذات قيما متغيرة وليست ثابتة ومن خلال اعتمادية تلك المعاملات على المتغيرات u نجد انها تؤدي الى حالة خاصة من التفاعل بين المتغير u والمتغير التوضيحي x.

2-2 طريقة النواة لمتعدد الحدود الموضعي ((LPK) Local Polynomial Kernel)) [5] [8]

تعتمد طريقة النواة لمتعدد الحدود الموضعي على فكرة تقريب معاملات المتغيرات التوضيحية لأنموذج المعاملات المتغيرة، والمبين في الصيغة (2) اذ ان تلك المعاملات $a_k(u)$ هي عبارة عن دوال في المتغير u. هذا التقريب يعتمد بالدرجة الاساس على متعدد الحدود من الدرجة p وكذلك النقطة الثابتة u_0، بأفتراض ان الدالة $a_k(u)$ التي نود تقديرها تمتلك مشتقات مستمرة لغاية المشتقة $p+1$ وبالاعتماد على تقريب تايلر (Taylor Expansion) يمكن كتابة الدالة $a_k(u)$ وكالاتي:

$$a_k(u) = \sum_{j=0}^{p} \frac{1}{j!}\, a_k^{(j)}(u_0)(u_i - u_0)^j$$

$$= \sum_{j=0}^{p} \beta_{jk}\,(u_i - u_0)^j \qquad \cdots (3)$$

اذ ان $\beta_{jk} = \frac{1}{j!}\, a_k^{(j)}(u_0)$ تمثل المعاملات التي سنقوم بتقديرها وان $a_k^{(j)}(u_0)$ تمثل مشتقة الدالة $a_k(u_0)$ للرتبة j عند النقطة u_0. ان متعدد الحدود (3) يمكننا من التقدير الموضعي لمعاملات متعدد الحدود β_{jk}، $j = 0,1,\cdots,p$ وذلك بتصغير معيار المربعات الصغرى الموزون (Weighted Least Squares (WLS)) والذي يكتب بهيئة المصفوفات وكالاتي:

$$\min_{\beta} (Y - X\beta)^T W (Y - X\beta) \qquad \cdots (4)$$

اذ ان $Y = [y_1, y_2, \cdots, y_n]^T$ يمثل متجه لقيم الاستجابة ذو درجة $n \times 1$ وان X تمثل مصفوفة التصميم ذات درجة $n \times (p+1)L$ والتي تكتب كالاتي:

$$X = \begin{bmatrix} x_1^T & (u_1 - u_0)x_1^T & \cdots & (u_1 - u_0)^p x_1^T \\ x_2^T & (u_2 - u_0)x_2^T & \cdots & (u_2 - u_0)^p x_2^T \\ \vdots & \vdots & \cdots & \vdots \\ x_n^T & (u_n - u_0)x_n^T & \cdots & (u_n - u_0)^p x_n^T \end{bmatrix}$$

علما ان x_t^T، $t = 1,2,\cdots,n$ تحدد تبعاً لعدد المتغيرات التوضيحية وتكتب كالاتي:

$$x_t = \left[x_{t1}\ ,\ x_{t2}\ ,\ \cdots\ ,\ x_{tL}\right]^T \qquad , \qquad t = 1,2,\cdots,n$$

اما المصفوفة W فتمثل مصفوفة قطرية للاوزان من الدرجة $n \times n$ وتكتب كالاتي:

$$W = diag\ \left\{K_h(u_t - u_0)\right\} \qquad , \qquad t = 1,2,\cdots,n$$

اذ ان K_h يتم تحديدها بالاعتماد على دالة النواة K اذ تم اعتماد دالة النواة المتمثلة بالصيغة الرياضية التالية والتي تم اقتراحها من قبل الباحثين:

$$K(u) = \frac{3315}{4096}(1 - u^4)^4\ ,\ I_{\{|u| \le 1\}}$$

وباعتماد الصيغة التالية K_h اذ ان عرض الحزمة (Bandwidth) الذي يرمز له بـ h

$$K_h(u_t - u_0) = \frac{1}{h} K\left(\frac{1}{h}(u_t - u_0)\right) \qquad \cdots (5)$$

ان دالة $K_h(\cdot)$ نحصل عليها بأعادة احتساب دالة النواة $K(\cdot)$ وتحديد عرض الحزمة المناسب الذي يكون ذو قيماً موجبة $(h > 0)$ ويستخدم لتحديد البيانات ضمن مدى الجوار الموضعي (Local Neighborhood) ، والمعرف بالشكل $I_h(u_0) = [u_0 - h, u_0 + h]$. ان دالة النواة تمكننا من تحديد المشاهدات ضمن المدى $I_h(u_0)$ التي تساهم في توفيق منحنى دالة الانحدار عند نقطة (u_0). ان متجه

المعاملات المراد تقديره هو $\beta=\left[\beta_0^T,\beta_1^T,\cdots,\beta_p^T\right]$ وتكون ذات سعة $(p+1)L\times 1$، اذ ان عناصر هذا المتجه هي β_j^T، $j=0,1,\cdots,p$ وتحدد بشكل عام تبعاً لعدد المتغيرات التوضيحية وتمثل بالمتجه $\beta_j=\left[\beta_{j1},\beta_{j2},\cdots,\beta_{jL}\right]^T$، $j=0,1,\cdots,p$ والصيغة التقديرية لمتجه معاملات متعدد الحدود الناتج من تصغير معيار المربعات الصغرى الموزون والمبين بالصيغة **(4)** تكتب كالآتي:

$$\hat{\beta}=(X^TWX)^{-1}X^TWY \qquad \cdots(6)$$

بايجاد تقدير معاملات دوال متعدد الحدود نتمكن من ايجاد مقدر الدالة $a_k^{(j)}(u_o)$ والذي يكتب وفق الصيغة الأتية:

$$\hat{a}_k^{(j)}(u_o)=j!\;e_{(j,k)}^T\;(X^TWX)^{-1}X^TWY \qquad \cdots(7)$$

اذ ان $e_{(j,k)}$ تمثل متجه احادي **(Unit Vector)** ذو درجة $((p+1)L\times 1)$ والذي يحتوي على واحد عند الموقع (j,k) اما المواقع المتبقية فتكون اصفار. ومن الجدير بالملاحظة ان الاهتمام ينصب على ايجاد مقدر الدالة $a_k^{(j)}(u_o)$ عندما $(j=0)$ اي عند المشتقة الصفرية لتلك الدالة والممثلة بالاتي:

$$\hat{a}_k(u_0)=e_k^T(X^TWX)^{-1}X^TWY \qquad \cdots(8)$$

ان تعويض تقدير المعاملات المتغيرة **(8)** في معادلة الانحدار **(2)** يمكننا من ايجاد القيم التنبؤيه لمتغير الاستجابة y_t , $t=1,2,\cdots,n$

$$\hat{y}_t=\sum_{k=1}^{L}\hat{a}_k(u_t)\,x_{tk} \quad ,t=1,2,\cdots,n \qquad \cdots(9)$$

3-2 طريقة انحدار الشريحة (Regression Spline (RS))

في حالة اعتماد طريقة انحدار الشريحة في تقدير معاملات أنموذج المعاملات المتغيرة فان التقريب في هذه الطريقة يعتمد على دوال اساس قوة القطع التي تكتب كالاتي[7] [13] [8] :

$$1\ ,\ u,....,u^p,(u-\kappa_1)_+^p,....,(u-\kappa_K)_+^p \qquad \cdots(10)$$

وبالاعتماد على دوال اساس قوة القطع التي تتضمن $p+K+1$ من الدوال باعتماد متعدد الحدود من الدرجة p وعدد K من العقد فان معاملات الأنموذج المتغيرة يمكن التعبير عنها وكالاتي:

$$a_k(u)=\sum_{j=0}^{p}\beta_{jk}u^j+\sum_{r=1}^{K}\beta_{(p+r)k}(u-\kappa_r)_+^p \qquad \cdots(11)$$

اذ ان $(\beta_0,\beta_1,\cdots,\beta_p,\beta_{p+1},\cdots,\beta_{p+K})$ تمثل معاملات دوال اساس قوة القطع **(10)** اذ ان تسمية انحدار الشريحة يرتبط بدرجة متعدد الحدود p المعتمدة فأذا كانت $p+1$ يطلق عليها انحدار الشريحة الخطية **(Linear)** اما اذا كانت $p=2$ فيطلق عليها انحدار الشريحة التربيعية **(Quadratic)** واذا كانت $p=3$ فتسمى انحدار الشريحة التكعيبية **(Cubic)**، ويمكن اعادة كتابة دالة معاملات الأنموذج المتغيرة بهيئة المصفوفات وكالاتي:

$$a_k(u)=Z\beta \qquad \cdots(12)$$

اذ ان **Z** هي مصفوفة التصميم ذات سعة $n\times(p+K+1)L$ وتكتب كالاتي:

$$Z=\begin{bmatrix} x_1^T & u_1x_1^T & \cdots & u_1^px_1^T & (u_1-\kappa_1)_+^px_1^T & \cdots & (u_1-\kappa_k)_+^px_1^T \\ x_2^T & u_2x_2^T & \cdots & u_2^px_2^T & (u_2-\kappa_1)_+^px_2^p & \cdots & (u_2-\kappa_k)_+^px_2^T \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^T & u_nx_n^T & \cdots & u_n^px_n^p & (u_n-\kappa_1)_+^px_n^T & \cdots & (u_n-\kappa_k)_+^px_n^T \end{bmatrix}$$

وان x_t ، $t=1,\cdots,n$ تكتب تبعاً لعدد المتغيرات التوضيحيه كالاتي:

$$x_t=\left[x_{t1},x_{t2},\cdots,x_{tL}\right]^T$$

وان β هو متجه معاملات دوال اساس القطع ذو سعة $(p+K+1)L\times 1$ ويكتب كالاتي:

$$\beta=\left[\beta_o^T,\cdots,\beta_p^T,\beta_{p+1}^T,\cdots,\beta_{p+K}^T\right]^T$$

وان β_q ، $q=0,1,\cdots,m$ ، $m=p+K+1$ تكتب تبعاً لعدد المتغيرات التوضيحيه وكالاتي:

$$\beta_q=\left[\beta_{q1},\beta_{q2},\cdots,\beta_{qL}\right]^T$$

وبالاعتماد على معيار المربعات الصغرى الاعتيادية **((OLS) Ordinary Least Square)** والمبين كالاتي:

$$\min_{\beta}(Y-Z\beta)'(Y-Z\beta) \qquad \cdots(13)$$

يمكن الحصول على تقدير المربعات الصغرى لمتجه معاملات دوال اساس قوة القطع وفق الصيغة الاتية:

$$\hat{\beta}=(Z^TZ)^{-1}Z^TY \qquad \cdots(14)$$

اذ ان $Y=\left[y_1,y_2,\cdots,y_n\right]^T$ يمثل متجه لقيم الاستجابة ذو درجة **(n×1)**، وان **Z** كما ذكرت في اعلاه.

ومن تقدير معاملات دوال اساس قوة القطع نحصل على تقدير معاملات الأنموذج المتغيرة وكالاتي:

$$\hat{a}_k(u)=Z\hat{\beta}=Z(Z^TZ)^{-1}Z^TY \qquad \cdots(15)$$

وبذلك يمكننا ايجاد متجه الاستجابة المقدر تبعاً الى نقاط الزمن $t=1,2,\cdots,n$ وكالاتي:

$$\hat{y}_t=\sum_{k=1}^{L}\hat{a}_k(u_t)\,x_t \qquad \cdots(16)$$

4-2 الممهدات الخطية (Linear Smoothers) [14]

الممهد الخطي يعرف بانه التعبير عن متجه قيم متغير الاستجابة المقدر $\hat{y}$ ولكل قيمة من قيم المتغير التوضيحي بهيئة تركيبة خطية في متجه قيم متغير الاستجابة الحقيقية y ويعتبر الممهد خطياً اذا استطعنا ان نكتب متجه قيم متغير الاستجابة المقدر $\hat{y}$ بدلالة متجه قيم متغير الاستجابة الحقيقية y وفق الصيغة المبسطة الاتية:

$$\hat{y}=A_\rho y \qquad \cdots(17)$$

اذ ان $\hat{y}=\left[\hat{y}_1,\cdots,\hat{y}_n\right]^T$ يمثل متجة قيم متغير الاستجابة المقدرة وان A_ρ هي مصفوفة من الدرجة $n\times n$ تعرف بمصفوفة التمهيد **(Smoothing Matrix)** والتي تأخذ صيغا مختلفة تبعاً لطريقة التقدير المعتمدة والتي تعتمد بدورها على معلمة التمهيد **(Smoothing Parameter)** ρ .

1-4-2 درجة حرية الأنموذج (df) degree of freedom))

درجة حرية التمهيد تستخدم كمقياس لتحديد عدد المعلمات ذات التأثير والمتضمنة في الأنموذج فمثلا في انحدار الشريحة درجة حرية التمهيد تساوي $m=p+K+1$ والتي تمثل عدد دوال الاساس متعددة الحدود من الدرجة p مضافاً اليها دوال قوة القطع K. ان درجة حرية التمهيد يعبر عنها بأثر **(Trace)** مصفوفة التمهيد A_ρ اي ان :

$$df=tr(A_\rho)=\sum_{t=1}^{n}a_{tt} \qquad \cdots(18)$$

اذ ان a_{tt} هي العناصر القطرية للمصفوفة A_ρ [14].

2-4-2 معيار العبور الشرعي (Cross Validation (CV))

معيار العبور الشرعي CV_ρ يستعمل بشكل واسع لاختيار انسب قيمة لمعلمة التمهيد ρ اذ ان صيغته تكتب بالشكل الاتي:

$$CV_\rho = n^{-1}\sum_{t=1}^{n}\left\{\frac{y_t-\hat{y}_t}{1-a_{tt}}\right\}^2 \qquad \cdots(19)$$

اذ ان a_{tt} تمثل العنصر t من عناصر القطر في مصفوفة التمهيد A_ρ وبذلك سيكون بامكاننا حساب قيمة العبور الشرعي لكل قيمة من قيم معلمة التمهيد ρ . نلاحظ من الصيغة **(19)** ان هذا المعيار يعتمد على الانحرافات المشاهدة $y_t-\hat{y}_t$ وعناصر القطر الرئيسي لمصفوفة التمهيد فقط مما يسهل عملية الحساب [14].

3-4-2 معيار العبور الشرعي العام (Generalized Cross Validation (GCV))

يعد معيار العبور الشرعي العام **GCV** تقريب لمعيار العبور الشرعي **CV** اذ يستعمل لاختيار انسب قيمة لمعلمة التمهيد ρ ويمكن الحصول على معيار العبور الشرعي العام **GCV** من صيغة معيار العبور الشرعي **CV** المبين بالصيغه **(19)** وذلك باستبدال عناصر القطر الرئيسي a_{tt} لمصفوفة التمهيد A_ρ بمعدلها اي ان $n^{-1}\sum_{t=1}^{n}a_{tt}=n^{-1}tr(A_\rho)=df/n$ وبهذا فأن صيغة معيار العبور الشرعي العام **GCV** تكون بالشكل الاتي [14]:

$$GCV_\rho=\frac{n^{-1}\sum_{i=1}^{n}[y_i-\hat{y}_i]^2}{\{1-tr(A_\rho)/n\}^2}=n^{-1}SSE_\rho/(1-df/n)^2 \qquad \cdots(20)$$

3. الجانب التجريبي

لغرض تطبيق طرائق التقدير اللامعلمية المعتمدة التي تم ذكرها في اعلاه تم اعتماد اسلوب المحاكاة **(Simulation)** والذي يحاكي عدداً كبيراً من الحالات الافتراضية المماثلة لنظيراتها في الواقع العملي بما يضمن شمولية اكبر للنتائج المتوقعة وذلك لأعتمادها في عملية التقدير والمقارنة بين تلك الطرائق المعتمدة في تقدير أنموذج المعاملات المتغيرة. تم تنفيذ تجارب المحاكاة بأعتماد ثلاث حجوم للعينات هي **(n_3=150, n_2=100, n_1=50)** وذلك لتوليد البيانات الخاصة بالمتغيرات العشوائية المتضمنة في أنموذج المعاملات المتغيرة **(VCM)** اذ تم تكرار كل تجربة **1000** مرة وذلك لاجل الحصول على نتائج متسقة. لقد تم الاعتماد على عدة دوال رياضية مختلفة لتمثيل المعاملات المتغيرة للمتغير التوضيحي حيث تم مراعاة الصيغة الخطية واللاخطية في تلك الدوال وندرج ادناه الدوال المعتمدة :

الدالة الخطية وصيغتها[15] $a(u)=2u$

الدالة الخطية بوجود حد ثابت[1] [12] $a(u)=2u+1$

الدالة من الدرجة الثانية[15] $a(u)=8(u-0.5)^2$

الدالة من الدرجة الرابعة[1] [11] $a(u)=1-48u+218\,u^2-315\,u^3+145\,u^4$

الدالة الغير خطية[1] [6] $a(u)=u+2\exp(-16u^2)$

وبالاعتماد على دوال المعاملات المتغيرة المذكورة اعلاه يمكن كتابة صيغة أنموذجالمعاملات المتغيرة والتي ستجري محاكاتهما وكالاتي:

الأنموذج الخطي $y=(2u)x+\in$

الأنموذج الخطي $y=(2u+1)x+\in$

الأنموذج من الدرجة الثانية $y=8(u-0.5)^2x+\in$

الأنموذج من الدرجة الرابعة $y=(1-48u+218\,u^2-315\,u^3+145\,u^4)x+\in$

الأنموذج الغير خطي $y=(u+2\exp(-16u^2))x+\in$

تضمن تنفيذ تجارب المحاكاة حول أنموذجات المعاملات المتغيرة كتابة عدة دوال بلغة **MATLAB** الاصدار **(2011)** لكل أنموذج من الأنموذجات الخمسة ذات المعاملات المتغيرة اذ تم تنفيذ مايأتي:

1- توليد المتغير التوضيحي x_1 بحيث يتوزع توزيعا طبيعياً بمتوسط يساوي صفر وتباين يساوي **0.02**.

2- توليد المتغير **u** الذي تعتمد عليه المعاملات المتغيرة بحيث يتوزع طبيعياً بمتوسط يساوي صفر وتباين يساوي **0.03**.

3- توليد المتغير العشوائي $\in$ الذي يمثل حد الخطأ العشوائي في أنموذج المعاملات المتغيرة قيد البحث بحيث يتبع التوزيع الطبيعي بمتوسط يساوي صفر وثلاث مستويات للانحراف المعياري هي $(\sigma_1 = \frac{1}{2}\ ,\ \sigma_2 = \frac{1}{4}\ ,\ \sigma_3 = \frac{1}{8})$.

4- تعويض المتغيرات التي تم توليدها في أنموذجات المعاملات المتغيرة لأجل الحصول على متغير الاستجابة y ولكل مستوى من مستويات الانحراف المعياري في (3) وكذلك لكل أنموذج من أنموذجات المعاملات المتغيرة الخمسة قيد البحث.

وفيما يلي الطرائق اللامعلمية التي تمت المقارنه بينها:

1. طريقة النواة المتعدد الحدود الموضعي (LPK).
2. طريقة انحدار الشريحة (RS).

بعد الحصول على نتائج تجارب المحاكاة التي تم تكرارها 1000 مرة وبأعتماد خمسة دوال تمثل معامل المتغير التوضيحي والتي اختيرت لمحاكاة أنموذج المعاملات المتغيرة والمبينة بالجدوال من (1) الى (5) اذ حصلت طريقة النواة لمتعدد الحدود الموضعي (LPK) على اصغر قيمة لمعدل معيار المفاضلة (MSE) ولجميع الأنموذجات الخمسة وجميع حجوم العينات ومستويات الانحراف المعياري لحد خطأ الأنموذج وقد لوحظ ان قيم الانحراف المعياري لمتوسط مربعات الخطأ (MSE) تشير الى تجانس قيم متوسط مربعات الخطأ (MSE) ولجميع التكرارات.

جدول (1)

يوضح معدل قيم معيار (MSE) ولجميع حجوم العينات ومستويات الانحراف المعياري للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الاول

		مستويات الانحراف المعياري لحد الخطأ		
حجوم العينات	الطرائق اللامعلميه	σ=1/2	σ=1/4	σ=1/8
n_1=50	LPK	0.059426387108747 (0.011485497828554)	0.003540283147170 (0.000720458974552)	0.000098914540261 (0.000021083960484)
	RS	0.061921947771345 (0.011777671448559)	0.003649676593534 (0.000736823039692)	0.000103496800571 (0.000021985577464)
n_2=100	LPK	0.062103312813904 (0.009111342621385)	0.003520111722584 (0.000493242050924)	0.000098395054915 (0.000016115197025)
	RS	0.063013346151030 (0.009103533373188)	0.003578507595288 (0.000487410778206)	0.000101729982176 (0.000016243810016)
n_3=150	LPK	0.063117685396637 (0.007011833628771)	0.003547376481152 (0.000354393273961)	0.000098505444721 (0.000011959660920)
	RS	0.063880113689738 (0.007131320763386)	0.003584413987733 (0.000355848715817)	0.000101011045558 (0.000012395588011)

جدول (2)

يوضح معدل قيم معيار (MSE) ولجميع حجوم العينات ومستويات الانحراف المعياري للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الثاني

		مستويات الانحراف المعياري لحد الخطأ		
حجوم العينات	الطرائق اللامعلميه	σ=1/2	σ=1/4	σ=1/8
n_1=50	LPK	0.060173596028792 (0.012797842563552)	0.003531070990053 (0.000751890214986)	0.000097760005020 (0.000020011694892)
	RS	0.062797339014056 (0.012955984368568)	0.004064759432146 (0.000810447675413)	0.000495273905435 (0.000101874148937)
n_2=100	LPK	0.060752328848187 (0.009020842424664)	0.003472382724420 (0.000521521099063)	0.000097100132212 (0.000013316786534)
	RS	0.062355586500900 (0.009051604180285)	0.003947972974728 (0.000614315985310)	0.000504466249582 (0.000067774210163)
n_3=150	LPK	0.063137862452745 (0.007707729024013)	0.003570544685812 (0.000430160580780)	0.000097253839503 (0.000010960992400)
	RS	0.064353270371794 (0.007907040734634)	0.004027061843998 (0.000460819703867)	0.000495282165481 (0.000059382020113)

جدول (3)

يوضح معدل قيم معيار (MSE) ولجميع حجوم العينات ومستويات الانحراف المعياري للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الثالث

		مستويات الانحراف المعياري لحد الخطأ		
حجوم العينات	الطرائق اللامعلميه	σ=1/2	σ=1/4	σ=1/8
n_1=50	LPK	0.000099160379671 (0.000017827402944)	0.003407244317579 (0.000631520784691)	0.000099716648962 (0.000019781534729)
	RS	0.001769486632493 (0.000401255063652)	0.005150973476085 (0.000959251048433)	0.001725174771571 (0.000312328011849)
n_2=100	LPK	0.000101150691607 (0.000014355486515)	0.003512686532970 (0.000601706264224)	0.000102936081512 (0.000014996785179)
	RS	0.001725662089331 (0.000243953634764)	0.005278500915887 (0.000730698313195)	0.001730488781618 (0.000257525608189)
n_3=150	LPK	0.000101125411601 (0.000011488215779)	0.003578629814083 (0.000455903249400)	0.000101984947886 (0.000011150497774)
	RS	0.001723398344383 (0.000200081119826)	0.005300230597138 (0.000545039512412)	0.001745597529413 (0.000197443587612)

جدول (4)

يوضح معدل قيم معيار (MSE) ولجميع حجوم العينات ومستويات الانحراف المعياري للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الرابع

		مستويات الانحراف المعياري لحد الخطأ		
حجوم العينات	الطرائق اللامعلميه	σ=1/2	σ=1/4	σ=1/8
n_1=50	LPK	0.061334996125405 (0.012840559155659)	0.003456598463195 (0.000712757650223)	0.000231708027587 (0.000075621891892)
	RS	0.064983144931886 (0.013310830223094)	0.004692698213326 (0.001101642197213)	0.001573314249699 (0.000661809043103)
n_2=100	LPK	0.060879625970763 (0.008146681734796)	0.003609688033503 (0.000535546820757)	0.000251262794899 (0.000072742546539)
	RS	0.063489007735646 (0.008792131687689)	0.005001544033703 (0.000871793520073)	0.001551937163715 (0.000496502258710)
n_3=150	LPK	0.061542042562676 (0.007857340633162)	0.003703242040865 (0.000496162046133)	0.000246362645900 (0.000050365665839)
	RS	0.063554103445571 (0.007891323027349)	0.005011394698181 (0.000783132412986)	0.001473597100692 (0.000377129515026)

جدول (5)

يوضح معدل قيم معيار (MSE) ولجميع حجوم العينات ومستويات الانحراف المعياري للطرائق اللامعلمية المستخدمة في تقدير أنموذج المعاملات المتغيرة الخامس

		مستويات الانحراف المعياري لحد الخطأ		
حجوم العينات	الطرائق اللامعلميه	σ=1/2	σ=1/4	σ=1/8
n_1=50	LPK	0.058866914405853 (0.012700183512511)	0.003517485859272 (0.000784395723644)	0.000098429346371 (0.000016179081966)
	RS	0.062186098988939 (0.013344743546606)	0.005198006846445 (0.001045286955305)	0.001669245937114 (0.000308529391262)
n_2=100	LPK	0.060454289406132 (0.008496145301877)	0.003625519855759 (0.000511697947306)	0.000097763872490 (0.000013750173817)
	RS	0.063152085806882 (0.008951219373768)	0.005223548724160 (0.000674849576477)	0.001643953122107 (0.000229086514178)
n_3=150	LPK	0.061492130799033 (0.007276389872775)	0.003548350550404 (0.000417107774795)	0.000100487286035 (0.000011764707673)
	RS	0.063664025833598 (0.007470133680575)	0.005171915562507 (0.000668690867718)	0.001642445340791 (0.000186335359739)

4. الاستنتاجات والتوصيات

1. من خلال مقارنة طرائق التقدير اللامعلمية المعتمدة في تقدير أنموذج المعاملات المتغيرة وجد ان افضل طريقة كانت النواة لمتعدد الحدود الموضعي (LPK) اذ حققت أقل قيمة لمعيار متوسط مربعات الخطأ (MSE) ولجميع حجوم العينات الثلاثة ومستويات الانحراف المعياري وكذلك لجميع الأنموذجات الخمسة المعتمدة في تجارب المحاكاة.
2. لوحظ ان قيم الانحراف المعياري لمتوسط مربعات الخطأ لكافة تكرارات تجارب المحاكاة لكل حالة من حجوم العينات الثلاثة ومستويات الانحراف المعياري الثلاثة وكذلك للأنموذجات الخمسة المعتمدة في المحاكاة، كانت صغيرة ومتقاربة مما يشير الى تجانس قيم متوسط مربعات الخطأ لكل حالة محاكاة.
3. وجد ان كلما زاد حجم العينة ادى الى زياده قيم معدل متوسط مربعات الخطأ وتقليل قيم الانحراف المعياري لمتوسط مربعات الخطأ مما يشير الى تحسين نتائج تقدير أنموذج المعاملات المتغيرة.
4. تبين انهُ كلما قلت قيمة الانحراف المعياري المستخدم في توليد قيم حدود الخطأ العشوائي للأنموذج المعاملات المتغيرة أدى الى تقليل قيمة متوسط مربعات الخطأ لتلك الأنموذجات.

5. المصادر

1. خمو ، خلود يوسف (2004) " مقارنة أساليب بيز مع طرائق أخرى لتقدير منحنى الانحدار اللامعلمي " أطروحة دكتوراه في الإحصاء ، كلية الإدارة والاقتصاد ، جامعة بغداد.
2. Breitner, S. (2007), "Time-varying coefficient models and measurement error"
3. Cai, Z., Fan, J., Li, R., (1999), "Generalized varying-coefficient models".
4. Chan, S. C., Zhang, Z. G.(2009) "Local polynomial modeling and bandwidth selection for time-varying linear models", IEEE
5. Chan, S. C., Zhang, Z. G.(2011) "Local polynomial modeling and variable bandwidth selection for time-varying linear system", IEEE Vol 60, No. 3
6. Denison, D.G.T., Mallick, B.K. & Smith, A.F.M., (1998),"Automatic Bayesian Curve Fitting", J.R. Statist. Soc. B, Vol. 60, pp.333-350.
7. Eubank, R.L. (1988). "Spline Smoothing and Nonparametric Regression". Marcel Dekker. New York.
8. Eubank, R.L. (1999). "Nonparametric Regression and Spline Smoothing". Marcel Dekker, New York.
9. Kuruwita C. (2011), "Nonparametric methods in varying coefficient models and quantile regression models".
10. McCullagh,P.and Nelder,J.A.;(2001);"Generalized linear models"; Chapman and Hall, London.
11. Ruppert, D., Sheather, S.J. & Wand, M.P., (1995), "An effective Bandwidth Selector for Local Least Squares Regression", J.Amer. Statist. Assoc.,Vol. 90, pp.1257-1270.
12. Smith, M. & Kohn, R., (1996), "Nonparametric Regression using Bayesian variable Selection", Journal of Econometrics, Vol. 75, pp.317-343.
13. Wahba, G. (1990). Spline Models for Observational Data. SIAM, Philadelphia. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59.
14. Wu, H. and Zhang, J., (2006), "Nonparametric regression methods for longitudinal data analysis: Mixed-Effects modeling approaches", John Wiley & Sons, New Jersey.
15. Zhang, Chongshan, Yin, Xiangrong, (2012) "Trending time-varying coefficient market models" Quantitative Finance Volume 12, Issue 10.

.....

.....

.....