مجلة القادسية لعلوم الحاسوب والرياضيات المجلد(٧) العدد(١) السنة(٢٠١٥)

صفحة ٤٠-٥٤ محمد حبيب/ زينب حسن

تمهيد نموذج الانحدار اللامعلمي بأستخدام أسلوب بيز

أ.د.محمد حبيب الشاروط
كلية الادارة والاقتصاد
جامعة القادسية
malsharood@gmail.com

زينب حسن راضي
كلية علوم الحاسوب والرياضيات
جامعة القادسية
zainbhassan70@yahoo.com

استلام البحث ١٦ /١١/٢٠١٤ إرسال التعديلات ١٨ / ١٢/ ٢٠١٤ قبول النشر ١١/١/٢٠١٥

الخلاصة Abstract

يهدف هذا البحث الى تمهيد نموذج الانحدار اللامعلمي بأستخدام أسلوب بيز في التقدير لما لهذا الاسلوب من نتائج دقيقة لكونه يعتمد على فكرة أفتراض معلمات النموذج كمتغيرات عشوائية ويتم تقديرها من خلال جعل توقع دالة الخسارة اقل مايمكن وتكمن عملية تمهيد نموذج الانحدار اللامعلمي من خلال اعتماد مؤشر ثنائي يتحكم في تنسيق اختيار المتغيرات ، وقد تم اعتماد شفرة كراي في تحديث المؤشر الثنائي واعتماد المعاينة المركزة في خوارزمية تمهيد دالة الانحدار اللامعلمي. وقد توصل الباحث الى ان شفرة كراي كفوءه في هذا المجال ولكنها تتطلب مسحا كاملا لفضاء تغير المؤشر الثنائي بينما أدت المعاينة المركزة دورا اساسيا في التركيز على المتغيرات النشطة من خلال تجزئة المتغيرات الى مجاميع سهلت بموجبها الحصول على تمهيد دالة الانحدارالتي تمتعت بنتائج جيدة لتقدير انحدار الدالة اللامعلمية وفقا لنتائج معايير المفاضلة المحتسبة.

Abstract

This research is aimed to smooth the nonparametric model by applying Bayesian approach in estimation where this approach has accurate results through minimizing estimation of loss function. This approach is achieved by considering a binary variable whose value controls the process of choosing of variables. The Gray code is dependent in updating the stream of binary variable while the focused sampling is dependent in algorithm of smoothing of nonparametric regression function. We deduced: (1) Gray code is efficient in this application but it needs complete scan to space of variables, (2) the focused sampling takes essential role in focusing on the active variables in each iteration and in partitioning the variables in subsets to simplify procedure of smoothing. We get very good results of plotting of regression also good values for criterions of errors.

١- المقدمة

ان تمهيد الانحدار اللامعلمي Smoothing nonparametric regression يعد من الاساليب المهمة المستخدمة في نمذجة البيانات عند عدم توفرالمعلومات عن الدالة التي تصف انحدار تلك البيانات او معالمها ، وقد زاد الاهتمام في السنوات الاخيرة في دراسة تمهيد الانحدار اللامعلمي لمرونته في التعامل مع البيانات وعدم وضعه لشروط مسبقة على الباحث الاحصائي اضافة الى ان التطور الكبير في اجهزة الحواسيب وانجاز معالجة البيانات بواسطتها قد ذلل الصعوبات المرافقة لهذا التحليل بسبب تطلبه لاجراء عمليات احصائية مطولة ومعقدة.

يوصف نمودج الانحدار بالعلاقة الاتية [1] [2]:

$$y_i = g(x_i) + \varepsilon_i \qquad i = 1, 2, ...n \qquad (1)$$

حيث y_i يمثل المتغير المعتمد و x_i يمثل المتغير المستقل ε_i يمثل الاخطاء العشوائية التي تتوزع بمتوسط صفر وتباين σ^2 وأن g تمثل دالة الانحدار المجهولة المطلوب تقديرها . ان تحديد اسلوب مناسب للنموذج (1) يتوقف على الافتراضات الممكنة للدالة g ففي الانحدار المعلمي يفترض ان الدالة معلومة عدا بعض معالمها تكون مجهولة ، اما اذا كانت g غير معلومة فيتم اللجوء الى الطرق اللامعلمية لتقديرها. ومن الاساليب اللامعلمية في تمهيد الدالة g هو اسلوب بيز في تقدير منحني الانحدار، ويمكن ذكر بعض ممن أستخدموا هذا الاسلوب في تمهيد الانحدار اللامعلمي مثل (2004خلود يوسف خمو) [1] و (Frederick Wong1997 , Mark Hansen ،Michael Smith) [2] و (Edward I.George1997 و Robert E. McCulloch) [3]

١-١ هدف البحث

يهدف هذ البحث الى دراسة تمهيد الانحدار اللامعلمي بأستخدام اسلوب بيز واختيار المتغيرات المناسبة من اجل الحصول على التقارب مع المنحني الحقيقي بأستخدام شفرة كراي وخوارزمية المعاينة المركزة بالاستعانة بنماذج المحاكاة اضافة الى مقارنة بين نتائج تمهيد بيز لبعض من نماذج أنحدار الشريحة Spline المستخدمة في البحث للحصول على التقارب مع المنحني الحقيقي .

2- مقدر بيز اللامعلمي Nonparametric Bayesian Estimator [1] [3]

يعرف مقدر بيز لاي معلمة غير معرفة θ بأنه القيمة $\theta^{\wedge}$ التي تؤدي الى تصغير القيمة المتوقعة لدالة الخسارة بالنسبة للتوزيع البعدي للمعلمة θ ، أي :

$$Bayesian\,Estimator = \min\left(E\left[L(\theta,\theta^{\wedge})\right]\right) \qquad \text{.........} \quad (2)$$

حيث يرمز لدالة الخسارة بالرمز L وهي دالة حقيقية تلعب دورا كبيرا في تحديد مقدر بيز كونها تمثل الخسارة نتيجة أتخاذ قرار التخمين ، وهناك انواع عديدة لدوال الخسارة أكثرها أستخداما هي دالة الخسارة التربيعية والتي تعرف على انها دالة متماثلة تحقق الشروط التالية:

$$L(\theta^{\Lambda},\theta) = (\theta - \theta^{\Lambda})^2$$
$$L(\theta^{\Lambda},\theta) = (\theta - \theta^{\Lambda})^2 = 0 \qquad if \;\; \theta^{\Lambda} = \theta$$

ان اسلوب بيز يفترض ان المعلمة المراد تقديرها هي متغير عشوائي لها توزيع اولي Prior Distribution يحتوي معلومات اولية عن معلمات العينة ثم نحصل على التوزيع اللاحق للمعلمات من خلال حاصل ضرب التوزيع الاولي ودالة الامكان الاعظم لمشاهدات العينة بمعنى ان التوزيع اللاحق يمثل تركيب بين المعلومات الاولية وبيانات العينة الحالية بعد المشاهدة. وبهذا فأن مقدر بيز يعتمد على دالتين الاولى هي دالة الكثافة الاحتمالية اللاحقة والثانية هي دالة الخسارة حيث يتم الحصول على توقع دالة الخسارة بالاعتماد على التوزيع اللاحق للمعلمة من خلال الصيغة الاتية :

$$E[\,L(\,\theta,\theta^{\wedge})] = E[\,L(\,\theta - \theta\,)^2/\,x] \qquad \text{.........}(3\,)$$
$$= \int (\,\theta - \theta^{\wedge})^2 \pi(\,\theta/\,x\,)d\theta$$

وحسب هذا التعريف فأن مقدر بيز اللامعلمي للمعلمة المجهولة θ سوف يمثل قيمة الوسط الحسابي $\theta^{\wedge}$ للتوزيع اللاحق لهذه المعلمة .

1-2 أنحدار بيز اللامعلمي non-parametric regression] [3] [4] [5]

بفضل التطور الكبير في برامجيات الحاسوب الذي ساعد في تطبيق طرائق بيز اللامعلمية وتقليل التعقيد في تنفيذها ، اصبح من الممكن استخدام طرائق بيز اللامعلمية التي ترتكز على عملية اختيار المتغيرات من خلال المناورة في حذف بعض المتغيرات والابقاء على البعض الاخر منها وصولا الى تحقق المجموعة المثلى من المتغيرات. ان عملية الحذف للمتغيرات تتم من خلال اعتماد مؤشر ثنائي γ عناصره صفراو واحد ، بحيث

يتم حذف المتغير اذا كانت قيمة المؤشر الثنائي تساوي صفر او الابقاء على المتغير اذا كانت قيمة المؤشر تساوي واحد.

توجد العديد من الطرق لتحديث قيمة المؤشر الثنائي في كل تكرار (والذي سوف ينعكس على تحديث لمجموعة المتغيرات التي سوف يتم اعتمادها في كل تكرار) منها اسلوب الاختيار الامامي (Forward Selection) واسلوب اضافة الخطوات المتسلسلة (Stepwise Addition) وشفرة كراي ولكل من هذه الطرق سلبياتها وايجابياتها .

ان تطبيق اسلوب بيز في اختيار المتغيرات على النموذج (١) سوف يؤدي الى النموذج التالي :

$$Y = X_\gamma \beta_\gamma + \varepsilon \qquad \text{..........} \quad (4)$$

حيث ان Y : متجه (nx1) للمتغير المعتمد

X_γ: المصفوفة المتبقية بعد حذف صف من صفوفها او عمود من مدورها

β_γ : متجه المعاملات المقابل لقيم γ_i

ε : متجه الاخطاء العشوائية بمتوسط صفر وتباين σ^2

γ : متجه المؤشر الثنائي

2-2 تخمين المتوسط اللاحق لعمود المعاملات β [2]

لتخمين المتوسط اللاحق يجب اولا ان يكون التوزيع الاولي معلوم لـ β الذي سوف يقودنا الى تصغير النموذج المختار من β من خلال اختيار المجموعة الجزئية المثلى من متغيراته التي تضمن الحصول على اقل قيمة لمعيار الخطأ عند المقارنة بين دالة الاختبار وتخمين دالة الانحدار البيزية التي استخدمنا فيها عمود β الامثل بما يضمن افضل تقارب بين المنحنيين. وعلى النقيض من ذلك ، اي اننا اذا اخترنا توزيع اولي غير مناسب β فأن ذلك قد يقود الى عدم الحصول على التقارب المطلوب او الحصول عليه بوقت تنفيذ اطول للبرامجيات.

ان المتوسط اللاحق لعمود المعاملات β هو:

$$E(\beta \mid y) = \sum_{\gamma} E(\beta \mid y, \gamma)\, p(\gamma \mid y) \qquad \text{..........} (5)$$

وهذا المتوسط اللاحق β يمثل المعدل الموزون للمتوسطات اللاحقة لتوقع $E(\beta \mid y, \gamma)$ مع اخذ الاوزان طبقا الى الاحتماليات اللاحقة $p(\gamma \mid y)$ وبحل المعادلة (5) يعطي تخمين قيمة عمود المتغيرات وهو التخمين الذي نحصل عليه كاحد نواتج خوارزمية اختيار المتغيرات :

$$\beta^{\wedge} = \frac{1}{M} \sum_{i=1}^{M} E(\beta \mid y, \gamma_i) \qquad \text{..........} (6)$$

حيث M هي عدد التكرارات و γ_i يمثل قيمة المتغير الثنائي عند التكرار رقم i

هناك طريقتان لحساب المتوسط في التوزيع اللاحق:

١) بشكل مباشرمن خلال التعامل مع 2^r من الحالات في فضاء المؤشر الثنائي حيث r يمثل طول متجه المعلمات β .

٢) من خلال استخدام احد سلاسل ماركوف مونت كارلو للنمذجة .

3-2 اسلوب تنفيذ المعاينة المركزة لتحديد التوزيع اللاحق [1] [2]

يركز اسلوب المعاينة المركزة في كل تكرار على المتغيرات التي كانت نشطة في التكرار السابق بدلا من التعامل مع جميع المتغيرات بشكل متساوي كما يتم في كل تكرار التعامل آنيا مع متغيرين او اكثر. ان هاتين الميزتين (والتي تختلف فيهما المعاينة المركزة عن معاين جبس) يجعلاها مخطط نمذجة كفوء حسابيا. عند تنفيذ خوارزمية المعاينة المركزة ، فأن الخوارزمية تقوم بتجزئة المجموعة الكلية للمتغيرات C الى مجاميع جزئية C_1, C_2....... ثم تقوم الخوارزمية بأيجاد المتغيرات المثلى في كل مجموعة جزئية التي تحقق افضل توزيع لاحق $p(\gamma \mid y)$ ثم يتم ايجاد المجموعة الكاملة من المتغيرات المثلى من خلال اتحاد المجاميع الجزئية ذات المتغيرات المثلى .... $C = C_1 \cup C_2$. ان تعامل خوارزمية المعاينة المركزة مع عدد اقل من المتغيرات ضمن كل مجموعة جزئية بدلا من التعامل مع كافة المتغيرات بنفس الوقت (كما هو الحال في معاين جبس) ينعكس على تحسن واضح في سرعة الوصول الى التقارب المطلوب.

4-2 ممهد بيز بأستخدام الشرائح التكعيبية [4]

ان الصيغة التالية تمثل صيغة الشريحة التكعيبية في تقدير دالة الانحدار:

$$g(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \sum_{k=1}^{m} \beta_k (x - x_k)_+^3 \qquad \text{..........} (7)$$

حيث ان $(x - x_k)_+$ يشير الى الشريحة kith للدالة التكعيبية وحسب التالي :

$$\begin{cases} (x - x_k) : & if x - x_k > 0 \\ 0 & : if x - xk \le 0 \end{cases}$$

حيث أن $x_1, x_2, \ldots\ldots x_m$ تمثل العقد بعدد m والتي يتم وضعها على طول مجال تغير المتغير المستقل بحيث أن $\min(x_i) < x_1 < x_2 < \ldots\ldots\ldots\ldots < x_m < \max(x_i)$. ان اعادة كتابة معادلة الانحدار (1) بأستخدام الشريحة التكعيبية في المعادلة رقم (7) يعني ان نموذج الانحدار اللامعلمي يعاد صياغته كنموذج انحدار خطي مماثل للنموذج الخطي الوارد في (4) حيث ان β يمثل متجه المعلمات $\beta = (b_0, b_1, b_2, \mathrm{b}3, \beta_1 \ldots\ldots \beta_m)$ بطول $(r \times 1)$ حيث $(r = m + 4)$ وهي تمثل عدد العقد مضافا لها الاربعة معاملات للحدود السابقة لظهور العقد في الشريحة التكعيبية في المعادلة رقم (7) .

ان عمود المتغيرات X وفقا للشرائح التكعيبية هو:

$$X = (1, x, x^2, x^3, (X - x_1)_+^3 \quad , \ldots\ldots\ldots (X - x_m)_+^3 \qquad \text{..........} (8)$$

وبذلك يعاد كتابة النموذج الوارد في (4) بصيغة المصفوفات بالشكل الاتي:

$$y_i = \begin{bmatrix} 1 & x & x^2 & x^3 & (X - x_1)^3 & (X - x_2)^3 & \ldots\ldots & (X - x_m)^3 \\ 1 & x & x^2 & x^3 & (X - x_1)^3 & (X - x_2)^3 & \ldots\ldots & (X - x_m)^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x & x^2 & x^3 & (X - x_1)^3 & (X - x_2)^3 & \ldots\ldots & (X - x_m)^3 \end{bmatrix} \begin{bmatrix} b_o \\ b_1 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}$$

أن الامر المهم والمتعلق بشرائح الانحدار بصورة عامة هو أختيار عدد العقد $x_1,\ldots\ldots\ldots x_m$ ومواقعها لان عدم التوفيق بين هذين المتطلبين سيؤدي الى اما لفقد الكثير من تفاصيل منحني الانحدار المطلوب ايجاده او الى الحصول على منحني انحدار بتغاير موضعي عالي (تمهيد واطيء).

في مجال اختيار العقد أعتمدنا في بحثنا هذا أسلوب اختيار عقده بين (3-5) من مواقع المتغيرات المستقلة ولغاية 40 عقدة كحد اعلى ، ولقد آثرنا اختيار عقدة بعد كل ٤ مواقع للمتغيرات المستقلة لذا فأن قيمة m ضمن مجال المتغير المستقل والمتوزع توزيعا منتظما $[0,1]$ سوف يكون $m=24$ وقيمة $r=28$ وهذا سوف يؤدي الى وجود عدد 28 من المعاملات β حيث يقوم البرنامج باختيار المجموعة المثلى منها التي تعطي افضل منحني انحدار بالشرائح التكعيبية بأصغر قيمة لمعيار تكامل مربع الخطأ ISE.

لقد تم اعتماد الخوارزمية الاتية لتنفيذ اسلوب بيز في اختيار المتغيرات:

الخطوة رقم (٠): اختيار قيمة ابتدائية للمتغير الثنائي $\gamma=\gamma^{[0]}$

الخطوه رقم (١): ايجاد عمود المعاملات ومصفوفة التصميم المقابلة لقيمة المتغير الثنائي

الخطوة رقم (٢): ايجاد تخمين دالة الانحدار ومقارنة قيمتها مع قيمة دالة الاختبار وحساب معيار الخطأ

الخطوة رقم (٣): تحديث قيمة المتغير الثنائي وفقا لشفرة كراي

الخطوة رقم (٤): تحديث عمود المعاملات ومصفوفة التصميم وفقا لتحديث المتغير الثنائي

الخطوة رقم (٥): اعادة تنفيذ الخطوات (٢) و (٣) و (٤) لحين الحصول على اقل قيمة لمعيار المفاضلة.

الخطوة رقم (٦): تكرار التجربة لعدد M من المرات

الخطوة رقم (٧): حساب قيمة المتوسط اللاحق من العلاقة رقم (٦).

وبنفس الاسلوب يمكن تمهيد منحني الانحداراللامعلمي بتطبيق اسلوب بيز بالشرائح التربيعية مع أجراء بعض التغيير في المعادلة (7)

$$g(x)=b_0+b_1x+b_2x^2+\sum_{k=1}^{m}\beta_k(x-x_k)_+^2 \qquad \ldots\ldots\ldots (9)$$

٣. الجانب التجريبي

من اجل تطبيق طرق التقدير المختلفة للانحدار اللامعلمي، فقد اعتمدنا الاسلوب التجريبي من خلال تطبيق أسلوب المحاكاة باستخدام (MATLAB "R2008a") حيث تم :

١. توليد المتغيرات المستقلة : قمنا بتوليد المتغير العشوائي X بأستخدام الاوامر المتاحة في برنامج الماتلاب[6] لتوليد المتغيرات العشوائية وبتوزيع منتظم بحيث حصلنا على المتغير التوضيحي $x_i \sim u(0,1)$ $i=1,2,\ldots n$ حيث n يمثل حجم العينة

٢. توليد الاخطاء العشوائية : يتم توليد الاخطاء العشوائية بنفس أسلوب توليد المتغير التوضيحي وذلك بأستخدام الاوامر المتاحة في برنامج الماتلاب [39]، بحيث يتم الحصول على اخطاء عشوائية تتوزع توزيع طبيعي بتوقع صفر وتباين σ^2

٣. المتغير المعتمد : يتم توليد المتغير المعتمد من خلال جمع دوال المتغير التوضيحي $g(x_i)$ مع متجه الاخطاء العشوائيه ε_i

٣-٢ معاييـر الاخطاء المستخدمة

لقد استخدمنا معايير الاخطاء التالية للمقارنة بين النماذج المنفذة :

١. متوسط مربع الخطأ (Mean Square Error) [1] وصيغته :

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\{g(x_i) - \hat{g}(x_i)\}^2$$

٢. تكامل مربع الخطأ (Integrated Squared Error) [1] وصيغته :

$$ISE = \int \left(g(x) - \hat{g}(x)\right)^2 dx$$

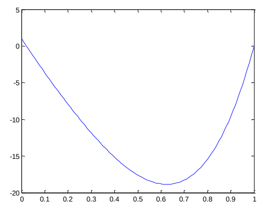
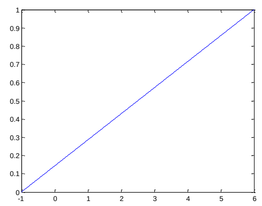
حيث ان $\hat{g}$ هي تقدير لدالة الانحدار g.

٣-٣ النماذج المستخدمة في تجارب المحاكاة [1]

- الدالة متعددة الحدود من الدرجة الخامسة بالصيغة

$$g(x) = 1 - 48x + 18x^2 - 15x^3 + 45x^4 - x^5$$

- الدالة الخطية : وصيغتها $g(x) = 7x - 1$

دالة الاختبار	الصيغة	شكل المنحني
متعددة حدود من الدرجه الخامسه	$g(x)=1-48x+18x^2-15x^3+45x^4-x^5$	
خطيه	$g(x)=7x-1$	

جدول رقم (1) : يمثل الدوال المستخدمة في المحاكاة

٤-٣ نتائج تجارب المحاكاة

(أ) الانحدار اللامعلمي لطريقة بيز بالشرائح التكعيبية والتربيعية مع دالة الاختبار متعددة الحدود

$$g(x)=1-48x+18x^2-15x^3+45x^4-x^5$$

كانت نتائج تمهيد الانحدار اللامعلمي بطريقة بيز بالشرائح التكعيبية لدالة الاختبار الاولى مبينة في الجدول رقم (2) ويتضح من النتائج كفاءة و بدرجة كبيرة لمقدر بيز بالشرائح التكعيبية Bayesian Cubic Spline مع عقد مكيفة ولم يظهر اي تذبذب بالمنحني للدالة بالاعتماد على تقدير المتوسط اللاحق ونلاحظ تفوقه على مقدربيز بالشرائح التربيعية Bayesian Quadratic Spline المبينة في الجدول رقم(3) حيث تمت المقارنة بالاعتماد على معيار المفاضلة ISE والاشكال المبين(1)(2) تبين الفرق بين التقديرين حيث يلاحظ تفوق الشرائح التكعيبية مقارنة بالشرائح التربيعية .

الطريقة	حجم العينة	σ = 0.25		σ = 0.5		σ = 1	
		AMSE	ISE	AMSE	ISE	AMSE	ISE
Bayesian-Cubic	20	0.489	0.3616	0.7858	0.7184	0.6491	0.8365
	50	0.2274	0.0927	0.2632	0.2026	0.1815	0.816
	100	0.3499	0.0796	0.2901	0.1246	0.3755	0.2367

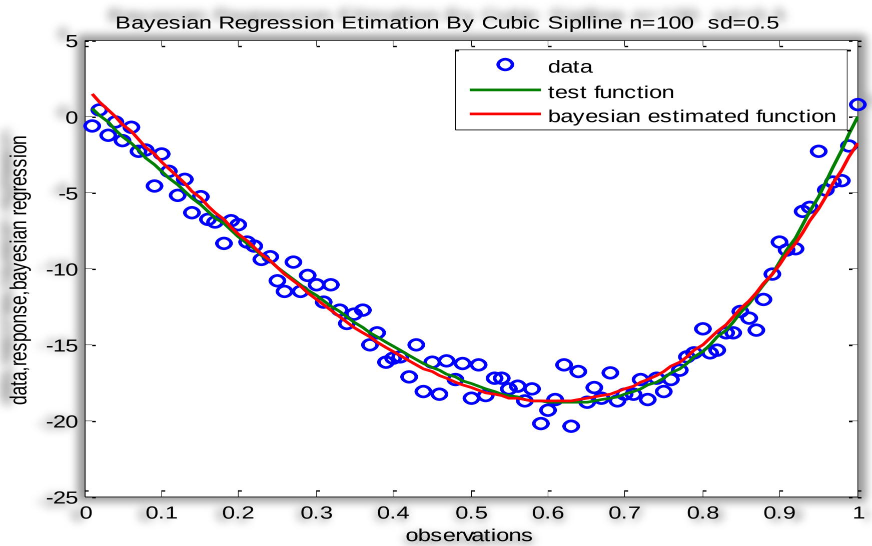
جدول رقم (2): يمثل نتائج تمهيد الانحدار اللامعلمي باستخدام طريقة بيز بالشرائح التكعيبية

لدالة الاختبار الاولى

الطريقة	حجم العينة	σ = 0.25		σ = 0.5		σ = 1	
		AMSE	ISE	AMSE	ISE	AMSE	ISE
Bayesian-quadratic	20	1.8119	5.1649	1.4799	2.1392	1.4719	2.9857
	50	1.2286	2.0041	1.2011	1.9899	1.1914	1.9551
	100	1.2586	2.0131	1.2237	1.9866	1.2019	1.9854

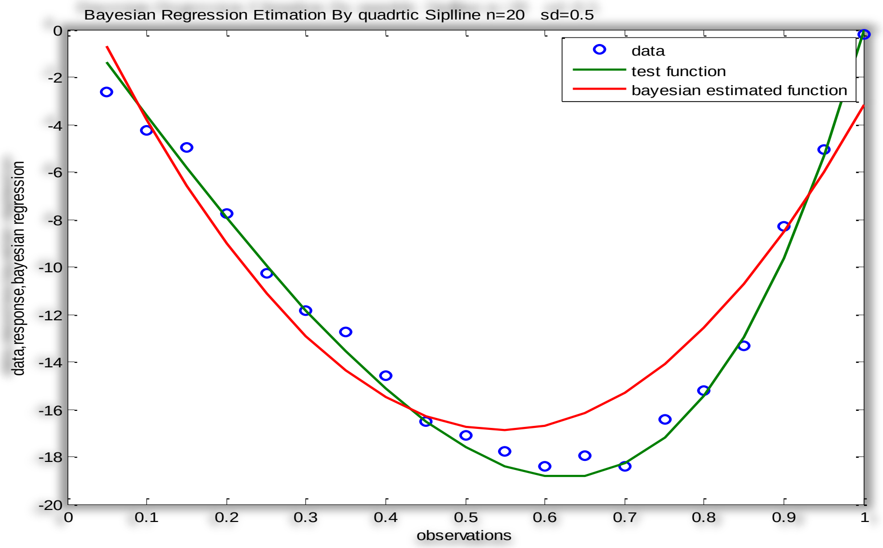
جدول رقم (3): يمثل نتائج تمهيد الانحدار اللامعلمي باستخدام طريقة بيز بالشرائح التربيعية

لدالة الاختبار الاولى



شكل رقم (1): يبين مقدر بيز لنموذج الانحدار اللامعلمي بأستخدام الشرائح التكعيبية لدالة الاختبار الاولى

عند σ = 0.5 وحجم عينة n=100



شكل رقم (2) :يبين مقدر بيز لنموذج الانحدار اللامعلمي بأستخدام الشرائح التربيعية لدالة الاختبار الاولى

عند $\sigma = 0.5$ وحجم عينة n=100

(ب) الانحدار اللامعلمي لبيز بالشرائح التكعيبية والتربيعية مع دالة الاختبار الخطية

$$g(x) = 7x - 1$$

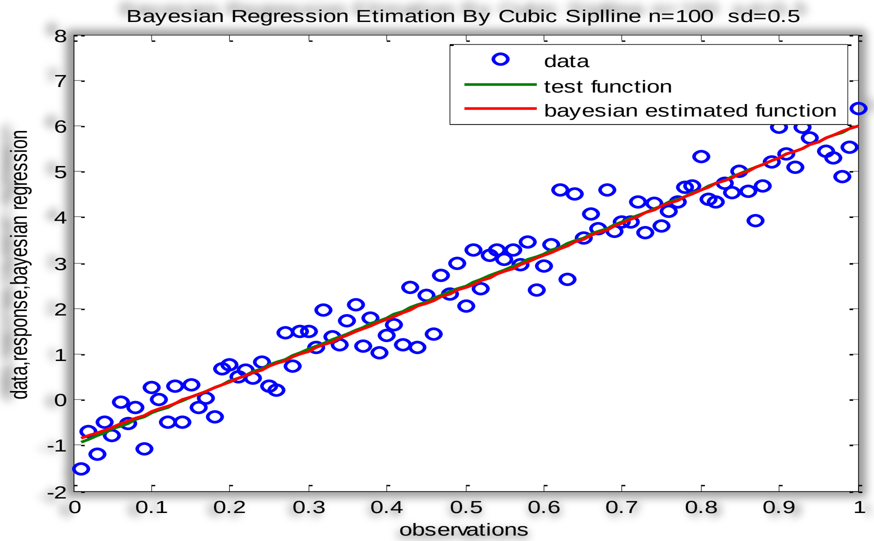
ان النموذج الخطي يعطي كفاءة عالية جدا لاستخدام الشرائح بنوعيها التربيعية والتكعيبية حيث ان اسلوب البحث العشوائي باستخدام شفرة كراي قد اعطى نتائج افضل في هذه الحالة مقارنة مع جميع الطرائق اللامعلمية ونلاحظ ذلك من خلاصة نتائج التقدير المبينة في الجدولين(5)(4) والرسوم البيانية بالاشكال (3)و(4) ، حيث ان تقدير المتوسط اللاحق لبيز لم يعاني من التذبذب وكما موضح بالشكل (3) و (4) هذا بالاضافة الى سهولة تتبع المخمن لاثر البيانات لعدم احتوائها على انحناءات في دالة الاختبار ، فقد تقاربت النتائج بالطريقتين لان الشريحة التربيعية أستطاعت تتبع نموذج الاختبار الخطي بشكل جيد مما يجعلها تنافس نتائج الشريحة التكعيبية

الطريقة	حجم العينة	σ = 0.25		σ = 0.5		σ = 1	
		AMSE	ISE	AMSE	ISE	AMSE	ISE
cubic	20	0.331	0.0051	0.074	0.0781	0.0681	0.0012
	50	0.0521	0.0066	0.0447	0.0035	0.0208	0.0066
	100	0.0514	0.0040	0.0269	0.0003	0.0422	0.0029

جدول رقم (4) :يمثل نتائج تمهيد الانحدار اللامعلمي بأستخدام بيز باستخدام الشرائح التكعيبية لدالة الاختبار الثانية

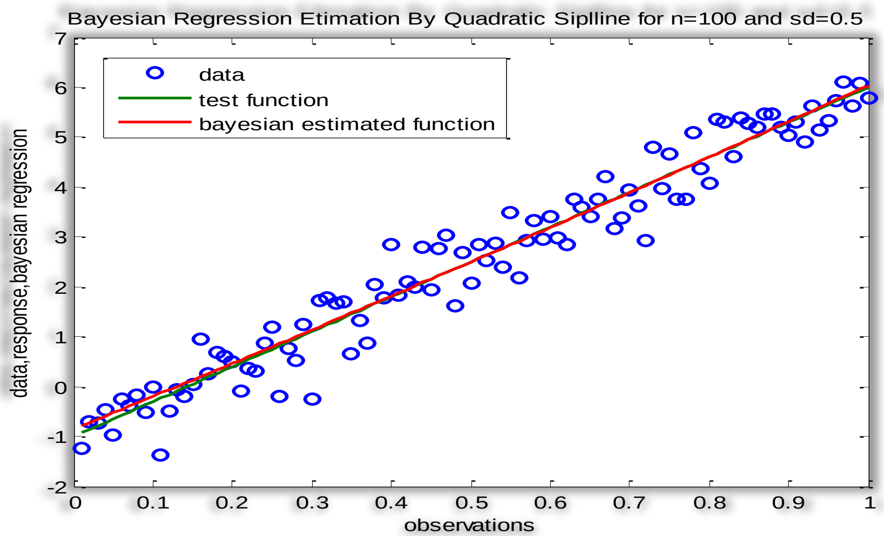
الطريقة	حجم العينة	σ = 0.25		σ = 0.5		σ = 1	
		AMSE	ISE	AMSE	ISE	AMSE	ISE
quadratic	20	0.1205	0.1542	0.0767	0.0163	0.2699	0.2699
	50	0.0110	0.00082	0.0046	0.00025	0.1094	0.0083
	100	0.0315	0.00018	0.0335	0.0030	0.0358	0.0027

جدول رقم (5) :يمثل نتائج تمهيد الانحدار اللامعلمي بأستخدام بيز باستخدام الشرائح التربيعية لدالة الاختبار الثانية



شكل رقم (3) :يبين مقدر بيز لنموذج الانحدار اللامعلمي بأستخدام الشرائح التكعيبية لدالة الاختبار الثانية

عند $\sigma = 0.5$ وحجم عينة n=100



شكل رقم (4) :يبين مقدر بيز لنموذج الانحدار اللامعلمي بأستخدام الشرائح التربيعية لدالة الاختبار الثانية

عند $\sigma = 0.5$ وحجم عينة n=100

٤- الاستنتاجات

بالنظر الى ماتوصلنا اليه من نتائج في الجانب التجريبي فقد استنتجنا الاتي:

١. أن اسلوب بيز اللامعلمي يوفر حلا ناجحا عند عدم توفر المعلومات عن معلمات دالة الانحدار المطلوب تخمينها ولكنها تنطوي على صعوبة بالغة جدا في تحقيق البرامجيات وحاجتها الى مستويات برمجة عالية بكفاءة حاسوب متطورة وعالية جدا.
٢. أن الاسلوب الذي اتبعناه في تقدير بيز اللامعلمي اعتمد على شفرة كراي والتي اخترنا السلوك البحثي فيها بسبب القلة القليلة جدا ممن استخدموا هذا الاسلوب وقد حققت هذه الشفرة نتائج جيدة الا انها تستغرق وقتا اطول في الوصول الى النموذج النهائي.
٣. اظهرت النتائج التي حصلنا عليها توافق مع الحقيقة القائلة بأن مؤشرات الخطأ في التقدير تقل عند زيادة حجم العينة بالرغم من ان بعض النتائج كانت غير ذلك كما في الجدول رقم ٢ ويعزى ذلك بسبب وجود القيم المتطرفة التي اثرت بشكل ملحوظ على بعض النتائج.
٤. أن نتائج تقدير بيز بالشرائح التكعيبية والتربيعية للدالة الخطيةأفضل من نتائج تقدير بيز بالشرائح التكعيبية والتربيعية لمتعددة الحدود .
٥.

٥- التوصيات:

١- اعتماد الشرائح التكعيبية في الانحدار اللامعلمي عند استخدام التمهيد باسلوب بيز لكفائتها في الوصول الدقيق الى تخمين نموذج الانحدار اللامعلمي.

٢. استخدام المعاينة المركزة كبديل عن اساليب المعاينة الاخرى خاصة في حالة كون عدد المتغيرات كبير كونها كفوءة في سرعة الوصول الى التقارب المطلوب بالمقارنة مع باقي اساليب المعاينة.
٣. ضرورة أستخدام حواسيب ذات سعات خزنية عالية وامكانيات جيدة وأصدارات حديثة للبرامجيات التطبيقية (مثل برنامج الماتلاب) من اجل اختزال زمن التنفيذوذلك لان اسلوب بيز يتطلب معالجة بيانية وبرامجيات عالية المستوى .
٤. ينبغي أستخدام شفرة كراي في أجراء التحديث المطلوب لتجاوزها الصعوبات المرافقة للمعاينات المرتكزة على سلسلة مونت كارلو

6- المصادر

١) يوسف ،خلود خمو (٢٠٠٤) "مقارنة أساليب بيز مع طرائق أخرى لتقدير منحني الانحدار اللامعلمي " أطروحة دكتوراه في الاحصاء ،جامعة بغداد ،كلية الادارة والاقتصاد

2) Frederick Wong, Mark Hansen, Robert Kohn and Michael Smith(1997) "Focused sampling and its application to nonparametric and robust regression "University of New South Wales

3) Edward I. George and Robert E. McCulloch (1997)"Approaches for Bayesian Variable Selection "University of Chicago

4) Persil Diaconis and Suan Holmes (1994)"Gray Codes for randomization Procedure " University Harvard

5) Xibin Zhang, Maxwell L. King (2013) "Bayesian Bandwidth Selection for a Nonparametric Regression Model With Mixed Types of Regresses "University of Southampton

6) MATLAB, Learn mat lab, version 6(release 12).