**تقنية ROBUST – LOESS في تحليل الانحدار**

**زينب حسن راضي**

**جامعة القادسية**

**علوم الحاسوب وتكنولوجيا المعلومات**

**الخلاصة :**

**تناولنا في هذا البحث دراسة أستخدام تقنية robust-loess في تقدير دالة الانحدار اللامعلمي ، حيث أن هذه الطريقة تستخدم في حالة وجود القيم المتطرفة خاصة وان طريقة loess هي طريقة غير حصينة كونها تعتمد استخدام المربعات الصغرى في التمهيد التي تتاثر بوجود الشواذ ويتم تنفيذrobust-loess باستخدام تقنية bisquar والتي تضيف المتانة الى المربعات الصغرى الموزونة في loess .**

**Abstract:**

**In this paper the robust-loess method is used to estimate the nonparametric regression function. The Loess is non-robust method and used in case of outliers where it bases on the less squares in regression which affects by presence of outliers. In this paper, the robust-loess has been implemented through applying bisquare technique which enhanced robustness to the weighted least squares in loess.**

**المقدمة :**

**تحليل الانحدار هو تخمين العلاقة بين المتغيرات ففي حال كوننا لانستطيع تحديد هذه العلاقة السلوكية وبدلا من اشتراط ان البيانات تاخذ قالبا او شكل دالي يوصف الظاهرة نلجأ الى استخدام اسلوب مرن للغاية هو الاسلوب اللامعلمي وقد استخدمنا في هذا الاسلوب تمهيد مخطط التشتت لانه يساعدنا في رؤية تلك العلاقة فهو يلخص متوسط البيانات باستخدام دالة تمهيد للنقاط اضافة الى انه يوفر لنا تقدير او تنبؤ لكل قيمة معطاة من  يطلق على هذه الطريقة loess وقد تم استخدامها من قبل باحثين منهم [1] (2004) خلود يوسف خمو و [2] (2015) زينب حسن راضي و William(1979)[3] و [4]  William and Devlin(1988)**

**صيغة نموذج الانحدار اللامعلمي : **

** يمثل الخطأ العشوائي،  دالة الانحدار المجهولة ، المتغير المعتمد .**

**هدف البحث :**

**يهدف البحث الى اضافة استخدام تقنية Robust الى طريقة loess لتمهيد الدالة اللامعلمية لاضفاء متانة ودقة في التقدير للوصول الى التقارب من المنحني الحقيقي.**

**البواقي :**

**من طبيعة تحليل الانحدار سواء كان معلمي او لامعلمي هو وجود البواقي والتي هي الفروق بين القيمة التي نحسبها من نموذج الانحدار والقيمة الحقيقية وتعرف بالصيغة التالية :**

****

**حيث ان  قيم المتغير المعتمد نسبة الى متغير المستقل ،  قيمة متغير التنبؤ ونعمل على عرض هذه الفروق او الاختلافات باستخدام احدى الطرق مثلا المدرج التكراري او box plot او استخدام مخطط التشتت البواقي .**

**:Loess**

**هي طريقة لامعلمية (بمعنى ليس لها تحديدات اولية لوصف شكل العلاقة بين المتغيرات ) وهي اختصار لمصطلح الانحدار الموضعي الموزون Cleveland(1979) والمصطلحان Loess , Lowess مشتقان من مصطلح locally weight scatter plot smoothing الطريقتين مختلفتين بالموديل المستخدم للانحدار، Lowess تستخدم دالة انحدار متعددة حدود خطية وLosee يستخدم دالة متعددة حدود تربيعية ، عملية التمهيد هنا تعتبر موضوعية لان كل قيمة ممهدة تحدد بواسطة جوار الاقرب للبيانات الواقعة ضمن الفترة و موزونة لان دالة انحدار الوزن تعرف لكل نقاط البيانات الواقعة ضمن الفترة وبأستطاعتنا ان نستخدم دالة الوزن الحصينة لجعلها مقاومة للقيم الشاذة او المتطرفة كما في طريقة Robust-loessوالخطوات ادناه توضح وصف خطوات loess :**

1. **نحسب انحدار الموزون لكل نقطة من نقاط البيانات ضمن المجال ، الوزن يعطى بأستخدام دالة الوزن الثلاثية :**

****

**تمثل قيمة التنبؤ المرتبطة بالتمهيد وقيم نقاط الجوار الى المعرفة في الفترة وطول المسافة على المحور السيني لمعظم قيم المؤشر في الفترة .**

**الوزن له الخصائص التالية :**

* **نقاط البيانات الممهدة تمتلك وزن اكبر ولها التاثير الاكبر على التمهيد**
* **نقاط البيانات الشاذة في المجال لها وزن صفر وليس لها تاثير على التمهيد**

1. **نستخدم انحدار المربعات الصغرى الموزونة ، اذا كان انحدار lowess نستخدم متعددة حدود من الدرجة الاولى ،اذا كان انحدار loess تستخدم متعددة حدود تربيعية**
2. **قيمة التمهيد تعطى بواسطة انحدار الموزون لقيمة التنبؤ ، اما اذا كان حساب التمهيد يتضمن نفس العدد للجوار الاقرب لنقاط البيانات على جانبي نقاط البيانات الممهدة فأن دالة الوزن متماثلة واكثر من ذلك اذا كان عدد نقاط بيانات الجوارالتغير غير متماثل حول نقاط البيانات الممهدة فأن دالة الوزن غير متماثلة .**

**Robust loess:**

**أن تقنية losee هي تقنية غير حصينة لانها تعتمد الية المربعات الصغرى التي تتاثر بوجود القيم المتطرفة التي تؤثر على قيمة التمهيد وتجعلها لاتعكس سلوك الجزء الاكبر من قيم البيانات لذا نلجأ الى استخدام اجراء robust لمعالجة هذا التطرف في البيانات حيث تضيف متانة الى اوزان المربعات الصغرى في خطوات loess والى جانب استخدامها تقنية bisquare هي اعادة اوزان النقاط الاساسية الى البواقي فاذا كانت البواقي كبيرة (انحراف كبير في النموذج ) فان اوزان هذه النقاط منخفض وبالعكس ويمكن وصف هذه الطريقة بعدة خطوات :**

* **حساب البواقي من الصيغة التالية :**

****

* **حساب الاوزان الحصينة لكل نقطة من نقاط الفترة ، والاوزان تعطى بواسطة دالة bisquar :**

****

**حيث ان تمثل البواقي لكل نقاط الفترة ،MAD هو المتوسط المطلق لانحرافات البواقي الذي هو مقياس لكمية انتشار البواقي**

****

**اذا كانت صغيرة مقارنة مع MAD فأن الاوزان الحصينة قريبة الى 1 اما اذا كانت كبيرة فأن الاوزان الحصينة تساوي صفر ونقاط البيانات المرتبطة تستبعد من حساب التمهيد**

* **نعيد تمهيد البيانات بأستخدام الاوزان الحصينة حيث ان قيمة التمهيد النهائية تستخدم كلا من الانحدار الموضعي الموزون والاوزان الحصينة**
* **نكرر الخطوات السابقة للحصول على افضل النتائج .**

**التطبيق :**

**لتطبيق ماقد تم ذكره اعلاه استخدمنا برنامج MATLAB 2013 ،قمنا بتوليد المتغير العشوائي و الاخطاء العشوائية بأستخدام الاوامر المتاحة في البرنامج ، أضافة الى توليد المتغير المعتمد من خلال جمع دوال المتغير التوضيحي مع متجه الاخطاء العشوائية ،واعتمدنا دالة اختبار متعددة الحدود لتوضيح تمهيد البيانات .**

**في هذا البحث استخدمنا قيم مجموع مربعات البواقي الحاصل بسبب التمهيد SSE (التي تعرف بأنها الاختلافات بين قيمة المشاهدة وقيمة التنبؤ) واحصائية معامل الارتباط R2 في الطريقتين اعلاه ، ان استخدام SSE اداة تشخيص مفيدة من اجل تحديد ماذا كان المنحني ممهد بشكل يشمل اغلبية البيانات حيث ان القيمة الصغيرة له تشير الى ان النموذج يمتلك خطأ عشوائي صغير وبالعكس، اما R-Square فهي مقياس لمدى نجاحنا في وصف او بيان العلاقة بين المتغيرات او هي مربع الارتباطات بين متغير الاستجابة ومتغير التنبؤ وهذه الاحصائية تاخذ قيمة بين 0 و1 .**

****

****

**سجلنا نتائج المؤشرين الاحصائيين لطريقة loess ، SSE=43.4و=0.9415 R2 وطريقة robust-loessوR2=0.9573 SSE=31.73 .**



**شكل (1) يوضح طريقة loess في التمهيد**



**شكل (2) يوضح طريقة robust-loess في التمهيد**

**الاستنتاج:**

1. **ان استخدام loess يمثل نهجا مرنا للغاية لمشكلة تمثيل البيانات بالاضافة الى اننا لانحتاج الى تحديد اولي ببعلاقة بين المتغيرات المستقلة وامعتمدة اضافة الى ان تركيبها واضح ومفهوم**
2. **تعتمد طريقة loess استخدام امربعات الصغرى في التمهيد لذا نستطيع اعتبارها طريقة عامة لكل من الاسلوبين المعلمي واللامعلمي ولكن بنفس الوقت يؤاخذ عليها تأثرها بالقيم المتطرفة لان المربعات الصغرى تتاثر بوجود الشواذ**
3. **يكون استخدام robust-loessحل لمشكة اقيم المتطرفة .**

**التوصيات :**

**نوصي باستخدام تقنية robust-loess في التمهيد لما تضيفه من متانة الى طريقة المربعات الصغرى حيث انها تعيد اوزان النقاط الى البواقي وبأستطاعتنا اضافة robust الى الشرائح spline او عندما يكون توزيع الخطأ العشوائي غير طبيعي (ملوث).**

**المصادر :**

1. **خلود يوسف خمو "مقارنة اساليب بيز مع طرائق اخرى لتقدير منحني الانحدار" أطروحة دكتوراه في الاحصاء ،جامعة بغداد ، كلية الادارة واقتصاد (2004).**
2. **زينب حسن راضي "تمهيد دالة الانحدار اللامعلمي بطرائق تمهيد متنوعة" رسالة ماجستير في الاحصاء الرياضي ،جامعة القادسية ،كلية علوم الحاسوب والرياضيات (2015).**
3. **William S.Cleveland"Robust locally weighted regression and smoothing" Journal of American statistical association" vol.74،issue368 (1979)**
4. **William S.Cleveland and sudanJ.Devlin"Locally weighted regression: Approach analysis by local" Journal of American statistical association" vol.83 ، NO.403، pp596-610، (1988) .**
5. **Mat lab "Learn mat lab" version 6 release 12 .**