KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

استخدام طريقة اختيار المتغيرات لاسو التكيفية البيزية – دراسة مقارنة

أ.د. احمد نعيم فليح

الباحث احمد جاسم حسن

جامعة القادسية / كلية الادارة والاقتصاد

DOI: https://doi.org/10.36322/jksc.176(B).19588

المستخلص

في هذا البحثِ نسلطُ الضوءُ على اكتشافِ الجانبِ النظريِ لمشكلةِ اختيارِ المتغيراتِ منْ خلالِ طريقةِ التنظيمِ لأسو التكيفيةَ وتقديرِ معالمِ نموذجِ الانحدارِ الخطيِ المتعددِ وفقَ أسلوبِ بيزْ , حيثُ تعتبرُ طريقةً لأسوِ التكيفيةَ أداةَ أوْ طريقةِ تقليصٍ تعملُ بنفسِ الوقتِ على اختيارِ المتغيراتِ التفسيريةِ , التي لها تأثيرٌ على متغيرِ الاستجابةِ وتقديرِ معالمِ نموذجِ الانحدارِ ذو القدرةِ التفسيريةِ العاليةِ معَ دقةِ تنبؤٍ عاليةٍ , حيثُ تمَ اقتراحُ نموذجٍ هرميٍ للتوزيعاتِ المسبقةِ , والجديدَ هنا أنَ في هذا البحثِ قدْ طورنا النموذجُ الهرميُ للتوزيعاتِ المسبقةِ منْ خلالِ اقتراحِ أنَ تباينَ البياناتِ σ^2 يمتلكَ توزيعٌ مسبقٌ يتبعُ توزيع معكوسٍ مربعٍ كأيَ . وبناءٌ على ذلكَ تمَ اشتقاقُ التوزيعاتِ اللاحقةِ . إضافةٌ إلى ذلكَ تمَ توظيفُ التوزيعاتِ اللاحقةِ في عملٍ خوارزميةٍ Gibbs للمعاينةِ لتوليدِ عيناتٍ لمعالمِ النموذجِ , وتمَ إجراءَ تجربةِ محاكاةٍ واحدةٍ وتحتَ قيمٍ مختلفةٍ لحجومِ العيناتِ , حيثُ تمَ استخدامُ طريقةٍ لأسوِ التكيفيةَ التقليديةِ وطريقةُ لأسوِ التكيفيةَ البيزية التي تمَ اقتراحها منْ قبلِ الباحثينَ Yi و Mallick عامُ 2014 ومقارنةَ النتائجِ معَ الطريقةِ المقترحةِ . وأظهرتْ النتائجُ وبناءً " على معياريْ MMAD و SD أنَ الطريقةَ المقترحةَ هيَ طريقةُ تنافسِ الطرائقِ الموجودةِ منْ ناحيةِ أداءِ التقديرِ , تمَ إجراءَ تطبيقٍ عمليٍ للبياناتِ التي تمثلُ دراسةَ نموذجِ الانحدارِ لمجموعةٍ منْ المتغيراتِ التفسيريةِ على متغيرٍ نقصِ الحديدِ في الدمِ لمجموعةٍ منْ المرضى , وتبينَ منْ

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.

مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

خلالِ توظيفِ الطريقةِ المقترحةِ أنها طريقةٌ قدْ وفرتْ أسلوبَ اختيارِ المتغيراتِ , بناءً " على القيمِ المقدرةِ ومقارنةُ النتائجِ معَ الطرائقِ المشارِ إليها في الجانبِ التجريبي . والجديرَ بالذكرِ أنَ النتائجَ التي تمَ الحصولُ عليها واعدةً ومشجعةً وتعطي نتائجُ أفضلُ مقارنةٍ بالطرقِ القائمةِ .

الكلمات المفتاحية: طريقة لاسو التكيفية , طريقة لاسو التكيفية البيزية , تحديد المتغيرات , المحاكاة.

Using the Bayesian Adaptive Lasso Variable Selection Method Comparative Study

Prof. Dr Ahmed Naim Flaih
Researcher Ahmed Jassim Hassan
Al-Qadisiyah University / College of Administration and Economics

Abstract

In this research, we highlight the discovery of the theoretical side of the problem of selecting the variables through a method Assessment of adaptive Lasso regularization and parameter estimation of the multiple linear regression model according to the method of Bayes. D. Lasso's adaptive method is a tool or method of contraction that works at the same time on the choice of language. Explanatory phrases, which have an effect on a variable The response and parameter estimation regression model has high explanatory power with high prediction accuracy A, where a hierarchical model for prior distributions was proposed, and what is new

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

here is that in this research we have developed the model A subtraction of the prior distributions by suggesting that the data variance σ^2 has a prior distribution that follows Inverted square sigmoid How much? Accordingly, the subsequent distributions were derived. In addition, the post-distributions were employed in the work of the Gibbs sampling algorithm to generate samples for dependents. of the model, and one simulation experiment was conducted under different values of the sample sizes, where the method was used. The method of the traditional adaptive lasso and the Bayesian lasso method which were proposed by the Researchers Yi and Mallick in 2014 And compare the results with the proposed method. And the results showed, and based on the MMAD and SD standards, that the proposed method is a method that competes with the other methods. exist in terms of estimating performance, a practical application was made of the data that represent the study of the donor model A study of a group of explanatory variables on the variant of iron deficiency in the blood of a group of patients , and it was revealed through the employment of The proposed method is a method that provides a method for selecting variables, based on the values of Estimate and compare the results with the methods referred to in the experimental aspect. It is

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

worth mentioning that the results obtained are promising and encouraging, and give better results compared to Existing roads.

Keywords: Adaptive Lasso Method, Bayesian Adaptive Lasso Method, selection Variables, Simulation.

1-المقدمة (Introduction)

يعتبر تحليل الانحدار وما ينتج عنه من نماذج التنبؤ اداة حيوية لتأسيس العلاقة بين متغير الاستجابة ومجموعة من المتغيرات التوضيحية , حيث تساعد هذه العلاقة في فهم سلوك الظاهرة المدروسة من خلال تحديد قيمة المتوسط لمتغير الاستجابة مما يساعد في التنبؤ بالقيم المستقبلية للظاهرة المدروسة . اضافة الى ذلك يمكن القول ان نماذج التنبؤ للظاهرة المدروسة يمكن ان توفر معلومات للمتغيرات التوضيحية التي تحدد قيمة متغير الاستجابة (outcome) ومدى قوة علاقة هذه المتغيرات لمتغير الاستجابة . ان لنماذج الانحدار تطبيقات غير محددة في مجالات مختلفة , من ضمنها المجالات الطبية حيث يستخدم نموذج الانحدار المقدر للمساعدة في الكشف عن المخاطر العالية لبعض الامراض مما يساعد في منع تطور مرض معين في وقت مبكر من خلال التدخلات الضرورية كذلك يمكن لهذه النماذج ان تتنبأ بالمرض في المستقبل مما يساعد في سهولة التواصل بين المريض والطبيب المختص اعتمادا" على معلومات اكثر واقعية وبالتالي يساعد النموذج المتنبأ به بتقديم المساعدة للأطباء والمرضى فيما يتعلق باختبار افضل , والمساعدة في الرعاية والخدمات الصحية وبالتالي ما ينعكس على اتخاذ القرار المناسب بشان التخطيط وادارة الجودة للمؤسسات الصحية .

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

ان اختيار الطريقة المناسبة في موضوع اختيار المتغيرات (Variable Selection) يعتبر من اهم الاساليب التي يحتاج لها الباحث في بناء نموذج الانحدار حيث يعتبر اختيار المتغيرات المناسبة (المهمة او ذات العلاقة بمتغير الاستجابة) الداخلة في نموذج الانحدار من اهم واصعب مراحل بناء نموذج الانحدار .

في عام 1960 قدم الباحث Efroymson بحثا" قدم فيه اسلوب لاختيار المتغيرات يدعى stepwise وهو تعديل لأسلوب اختيار المتغيرات المسمى (Forward selection) او اختصارا" (FS) حيث يعمل هذا الاسلوب على ادخال المتغير التوضيحي في النموذج او استبعاده من النموذج واحدا" تلو الاخر وبالتالي يتم اختيار مجموعة جزئية واحدة فقط من كل المجاميع الجزئية الممكنة للنماذج . اضافة الى كل ذلك يعتبر اسلوب (Backward eliminalion) احد هذه الاساليب في اختيار المتغيرات .

في عام 1973 قدم الباحث Akaike بحثا" قدم فيه احد معيار المعلومات Information criteria والتي تعتبر بمثابة معايير لاختبار النماذج وتم تسمية هذا المعيار Akaike Information criteria مختصرا"AIC ان هذا المعيار يعمل على اختيار افضل النماذج بحيث يتم الموازنة بين دقة الملائمة (Pit) وبساطة النموذج (اقل عدد من المتغيرات التوضيحية) .

في عام 1996 قدم الباحث Tibshirani بحثا" اقترح فيه طريقة مبتكرة في اختيار المتغيرات وتعتمد في تمثيلها الهندسي على جعل معالم بعض المتغيرات التوضيحية تساوي للصفر , اي انها طريقة تؤشر حلول صفرية sparse solution وتسمى هذه الطريقة بطريقة Lasso وهي احدى طرائق التنظيم Regularization التي تعالج مشكلة تضخم التباين بمقدرات طريقة OLS .

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

في عام 2006 قدم الباحث Zou بحثا" قدم فيه طريقة انكماش جديدة تدعى adaptive Lasso اي لاسو التكيفية وهي طريقة تنظيم مبنية على اساس انها طريقة لاسو ذات المرحلتين Two-stages Lasso والتي تعالج اعتماد طريقة Lasso على استخدام معلمة انكماش واحدة لكل من عملية اخيار المتغيرات وعملية التقدير , حيث يتم بهذه الطريقة المقترحة اكثر من معلمة انكماش . حيث تستخدم هذه الطريقة الجديدة عوامل لمعالم دالة الجزاء تتمثل بأوزان موجبة للسيطرة على التحيز في المقدر من خلال تقليص معالم نموذج الانحدار بكميات مختلفة (اوزان) وتعتبر ايضا" من طرائق اختيار المتغيرات لأنها تمثل معالم بعض المتغيرات مساوية للصفر , اضافة الى ذلك فان مقدراتها تتميز بخصائص oracle .

في عام 2008 قدم الباحثان Park and Casella اول بحث يتناول طريقة لاسو في تقدير معالم نموذج انحدار متعدد وفق اسلوب بيز . حيث اعتمد الباحثان الى اشارة Tibshirani في بحثه عام 1996 الى ان اسلوب بيز في تقدير معالم نموذج الانحدار يتطلب استخدام توزيع لابلاس كتوزيع مسبق لمعلمة الانحدار .في هذا البحث استخدم الباحثان تمثيل مختلط لتسهيل العمليات الحسابية في الخوارزمية وهذا التمثيل هو عبارة عن صيغة رياضية تتمثل بحاصل ضرب التوزيع الطبيعي مع التوزيع الاسي وبالتالي تطوير نموذج هرمي للتوزيعات المسبقة والتي من خلالها يتم احتساب التوزيعات اللاحقة في تقدير معالم نموذج الانحدار .

في عام 2014 قدم الباحث Leng واخرون اولا" بحث يتعلق بطريقة لاسو التكيفية وفق اسلوب بيز , حيث اقترح الباحثون استخدام معلمة الانكماش كمعامل وزن للمعالم المقدرة , اي ان كل معلمة مقدرة يوجد لها عامل مرافق يتمثل بالوزن . حيث تعتبر هذه الطريقة بمثابة طريقة لاختيار المتغيرات وطريقة لتقدير معالم نموذج الانحدار . اضافة الى ذلك وضح الباحثون الية اختيار النموذج الافضل باستخدام

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

الطريقة المقترحة اعتمادا" على قيم المنوال المقدرة من التوزيع اللاحق الشرطي . كذلك تم اختبار الطريقة المقترحة في الجانب التنبؤي للنموذج المفترض , حيث بينت الطريقة المقترحة افضلية الطريقة المقترحة في جانب دقة التنبؤ . ان تجارب المحاكاة وتحليل البيانات الحقيقية بينت قدرة الطريقة المقترحة وكفاءتها في كل من جانب اختيار المتغيرات وجانب التنبؤ .

في عام 2015 قدم الباحث Feng واخرون بحثا" عن طريقة لاسو التكيفية وتوظيفها في الانحدار الرتبي (Ordinal) بوجود متغيرات كامنة (Latent) حيث تدارس الباحثون في هذا البحث المتغيرات التفسيرية بافتراضها متغيرات كامنة على متغير الاستجابة بافتراضه متغير رتبي . كذلك تدارس الباحثون ان المتغيرات الكامنة تم التعبير عنها من خلال قيم الارتباطات بين المتغيرات التفسيرية مستخدمين نموذج التحليل العاملي التوكيدي . ان طريقة لاسو التكيفية وفق اسلوب بيز تم استخدامها لتقدير معالم المتغيرات واجراء اختبار المتغيرات , واظهرت نتائج تجارب المحاكاة وتحليل البيانات الواقعية كفاءة طريقة لاسو التكيفية في تحليل مثل هذا النوع من البيانات .

في عام 2020 قدم الباحث Flaih واخرون بحثا" قدموا فيه طريقة جزاء لاسو التكيفية وفق اسلوب بيز في نموذج انحدار Tobit والاسهام في هذا البحث كان من خلال تقديم تمثيل جديد لتوزيع لإبلاس من خلال صياغة توزيع مختلط للتوزيع الطبيعي مخلوطا" مع توزيع Rayleigh واعتمادا" على هذا التمثيل المختلط الجديد تم اقتراح نموذج هرمي جديد للتوزيعات المسبقة وتوظيفها في اشتقاق التوزيعات اللاحقة وباستخدام خوارزمية معاينة كبس تم تنفيذ العمليات الحسابية وفق لغة R من خلال استعراض سجل تجارب المحاكاة والبيانات الحقيقية في قدرة النموذج المقترح وكفاءته في اسلو اختيار المتغيرات والتفسير والتنبؤ وفق النموذج المقدر .

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

في عام 2022 قدم الباحثان Zhao , Duan بحثا" عن تقدير نماذج الانحدار الخطي باستخدام طريقة لاسو التكيفية وفق اسلوب بيز في حالة وجود قيم مفقودة في مشاهدات متغير الاستجابة , حيث تم استخدام الية لمعالجة القيم المفقودة من خلال استخدام نموذج الانحدار اللوجستي . وقد تم استخدام خوارزمية معاينة كبس و خوارزمية متروبولس – هاستنكز في تقدير معالم التوزيعات اللاحقة , ومعالم الانكماش اضافة الى ذلك تم حساب الاخطاء المعيارية للقيم المقدرة للمعالم . وقد تم تطبيق النموذج المقترح في عدة تجارب محاكاة وكذلك تطبيق النموذج على بيانات تلوث الهواء. واظهرت النتائج ان الأسلوب المقترح هو اسلوب فعال وممكن التطبيق على اساليب احصائية اخرى .

الهدف من هذا البحث هو مقارنة طريقة مقترحة لتقدير معالم نماذج الانحدار الخطي المتعدد مع طريقتي الجزاء Adaptive Lasso , Bayesian Adaptive Lasso , وتطبيقها في المجال الطبي لاسيما لتمثيل بيانات نقص الحديد في الدم كمتغير استجابة ومجموعة من المتغيرات التوضيحية . للحصول على النموذج الافضل من خلال اختيار المتغيرات المؤثرة بمتغير الاستجابة مما يسهل عملية تفسير سلوك البيانات المدروسة .

2-طرائق التقدير

2-1: طرائق الانكماش واختيار المتغيرات Variable Selection and shrinkage Method

اشار الباحث Tibshirani في عام 1996 الى ان طريقة المربعات الصغرى الاعتيادية لا يمكن الاعتماد على نتائجها عند تحليل البيانات وذلك لسببين . السبب الاول هو دقة التنبؤ , حيث عادة ما تكون مقدرات طريقة المربعات الصغرى تتصف بانها متحيزة قليلا" لكن لها تباين كبير , لذلك ومن اجل تحسين دقة التنبؤ وتحسين النتائج فان ذلك يتم من خلال تقليص بعض المتغيرات وجعلها مساوية للصفر , وبالتالي

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

فان هذا الاجراء سوف يؤدي الى زيادة تحيز المقدر ولكن بالنتيجة يقلل من تباين القيم المتنبأ بها الذي بدوره يؤدي الى تحسين دقة التنبؤ . السبب الثاني هو القدرة التفسيرية للنموذج , حيث عادة ما يتطلب اختيار بعض المتغيرات من مجموعة من متغيرات التي لها تأثير قوي على متغير الاستجابة مما يؤدي الى انتاج نموذج انحدار ذو قدرة تفسيرية عالية .

2-2: طريقة لاسو التكيفية Adaptive Lasso Method

اقترح الباحث Zou في عام 2006 طريقة جديدة من طرائق اختيار المتغيرات وهذه الطريقة التنظيمية تدعى adaptive Lasso اي طريقة لاسو التكيفية حيث استخدم الباحث Zou في هذه الطريقة لحساب مقدرات نموذج الانحدار الخطي المتعدد عوامل تتمثل بالأوزان الموجبة في دالة الجزاء المضافة الى مجموع مربعات الخطأ , والهدف من هذه الاوزان هو للسيطرة وبشكل مباشر على قيم التحيز في المقدرات من خلال المعالم بكميات مختلفة . ان مقدر طريقة لاسو التكيفية يحسب من خلال مشكلة الامثلية الاتية:-

$$\hat{\beta}_{\text{adaptive lasso}} = \arg\min_{\beta} \|Y - X\beta\|_2^2 + \lambda \sum_{j=1}^{p} W_j \left|\beta_j\right| \quad \dots(1)$$

حيث ان $\lambda > 0$ و $W_j > 0$ تمثلان معلمة الانكماش والاوزان على التوالي . وعرف الباحث Zou دالة للوزن وكما يأتي :-

$$\widehat{W}_j = 1/\left|\hat{\beta}_j(\text{ initial})\right|^{\alpha}$$

حيث ان $\alpha > 0$ وان $\hat{\beta}_j(\text{ initial})$ تمثل القيمة التقديرية الابتدائية للمعلمة β من OLS

3-2: طريقة لاسو التكيفية البيزية Bayesian Adaptive Lasso Method

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

اشار الباحث Tibshirani في عام 1996 الى امكانية افتراض ان معلمة نموذج الانحدار الخطي المتعدد β على انها متغير عشوائي عند محاولة تقديرها وفق اسلوب بيز وان هذا المتغير افترض انه يمتلك توزيعا" مسبقا" Prior كتوزيع لابلاس .

$$\pi(\beta_j) = \frac{1}{2\tau}\exp\left\{-\frac{|\beta_j|}{\tau}\right\} \quad \dots (2)$$

حيث ان $\tau = \frac{1}{\lambda}$ وبالتالي وعند استخدام توزيع للبيانات (Likelihood) وقاعدة بيز فانه يمكن استنتاج ان قيمة β المقدرة تمثل قيمة منوال التوزيع اللاحق Posterior وهذا الاسلوب بالتقدير تم تسميته بطريقة لاسو البيزية . ويعتبر عمل الباحثان Casella و Park في عام 2008 اول عمل في تقدير معالم نموذج الانحدار الخطي المتعدد وفق اسلوب لاسو البيزي مفترضا" ان التوزيع المسبق (2) يمكن تمثيلة من خلال توزيع مختلط للتوزيع الطبيعي مع التوزيع الاسي . تبع هذا العمل عدة بحوث انجزها كل من Griffin and Brown في عام 2010 و Hans في عام 2010 .

3- : النموذج الهرمي للتوزيعات المسبقةThe hierarchical model of prior distributions

الآن يمكننا كتابة نموذج التوزيعات الهرمية السابقة وكما يأتي :-

$$\tau = \frac{\sqrt{\sigma^2}}{\lambda}$$

$$\prod_{j=1}^{p} \pi\left(\beta_j|\,\sigma^2, \lambda_j\right) = \frac{\lambda_j}{\sqrt[2]{\sigma^2}} e^{-\frac{\lambda_j|\beta_j|}{2\sqrt{\sigma^2}}}$$

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

$$\beta \sim \prod_{j=1}^{p} \text{laplace}(\beta_j | 0, \frac{\lambda_j}{\sqrt{\sigma^2}})$$

وبالتالي يمكن اعادة كتابة التوزيع المسبق للمعلمة β_j من خلال التمثيل المختلط لمعلمة القياس Scale – mixture الاتي :

$$\prod_{j=1}^{p} \pi(\beta_j | \sigma^2, \lambda_j) \propto \prod_{j=1}^{p} \frac{\lambda_j}{2\sqrt{\sigma^2}} \exp\{-\frac{\lambda_j |\beta_j|}{2\sqrt{\sigma^2}}\}$$

$$\propto \prod_{j=1}^{p} \int_0^{\infty} N(\beta_j | 0, \sigma^2 \gamma_j^2) \, EXP\left(\frac{\gamma_j^2}{1}, \frac{\lambda_j^2}{2}\right) d\gamma_j^2$$

ومن خلال توظيف الصيغ اعلاه يمكن كتابة النموذج الهرمي التالي للتوزيعات المسبقة من خلال افتراض وجود المتغيرات الاضافية axuiliary Variable (Λ) :

$$y \sim N(y | X\beta, \sigma^2 I),$$

$$\beta \sim \prod_{j=1}^{p} N(\beta_j | 0, \sigma^2 \gamma_j^2),$$

$$\Lambda \sim \prod_{j=1}^{p} \exp\left(\gamma_j^2 | 1, \frac{\lambda_j^2}{2}\right),$$

$$\sigma^2 \sim \text{Inverse Chi} - \text{squar distribution } (\sigma^2 | a, b),$$

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

حيث ان Λ تمثل مصفوفة قطرية diagonal matrix عناصرها القطرية تمثل: $\gamma_1^2, \gamma_2^2, \gamma_3^2, \ldots, \gamma_p^2$

وبهذا يمكن كتابة دالة التوزيع اللاحق المشتركة حول كل التغيرات وهي :

$$\pi(\beta, \Lambda, y, X, \sigma^2) \propto N(y|X\beta, \sigma^2 I) \prod_{j=1}^{p} N(\beta_j|0, \sigma^2 \gamma_j^2) \left[\prod_{j=1}^{p} EXP\left(\gamma_j^2|1, \frac{\lambda_j^2}{2}\right)\right]$$

$$. Invrse - \chi^2(\sigma^2|a, b) \ldots (3)$$

حيث يستخدم النموذج الهرمي مرحلة العمليات الحسابية لتوليد العينات بتقنية MCMC ممكناً وبالتالي الاستدلال على خصائص التوزيعات اللاحقة . ان المتغيرات الاضافية يمكن استبعادها من خلال اجراء عملية التكامل لها اثناء عملية توليد العينات , وبالتالي يتم استخدام العينات المتحصل عليها في ايجاد توقع وتباين معالم نموذج الانحدار من خلال توزيعها اللاحق. وان التوزيع الشرطي للمتغير y هو

$$y|X, \beta, \sigma^2 \sim N(X\beta, \sigma^2 I),$$

اذن التوزيع اللاحق الشرطي الكامل لـ β يكون بالشكل الاتي :

$$= -\frac{1}{2\sigma^2}\{(\beta - A^{-1} X^T y)^T A(\beta - A^{-1} X^T y)\}$$

وباستذكار التوزيع الطبيعي نجد ان : $X \sim N(\mu, \sum)$

$$f(X; \mu, \sum) = \frac{e^{-\frac{1}{2}(X-\mu)^T \sum^{-1} (X-\mu)}}{(2\pi)^{\frac{n}{2}} |\sum|^{\frac{1}{2}}}$$

نستنتج ان توزيع β هو التوزيع الطبيعي المتعدد بمتوسط $A^{-1}X^T y$ وتباين $\sigma^2 A^{-1}$.

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

اما التوزيع اللاحق الشرطي الكامل لـ σ^2 موضح في ادناه حيث سيتم استخدام خوارزمية Gibbs لتوليد عينات من هذا التوزيع , ان المقادير الرياضية التي تضم σ^2 في التوزيع المشترك هي :

$$= (\sigma^2)^{-\left[\frac{n-1}{2}+\frac{p}{2}+\frac{a}{2}\right]-1} \exp\left[-\frac{1}{\sigma^2}\left\{\frac{(y-X\beta)^T(y-X\beta)}{2}+\frac{\beta^T W_\gamma^{-1}\beta}{2}+\frac{ab}{2}\right\}\right] \dots (4)$$

فانه يمكن الاستنتاج بان الصيغة (4) هي توزيع معكوس كاما بمعلمة شكل $\left[\frac{n-1}{2}+\frac{p}{2}+\frac{a}{2}\right]$ ومعلمة قياس $\frac{(y-X\beta)^T(y-X\beta)}{2}+\frac{\beta^T W_\gamma^{-1}\beta}{2}+\frac{ab}{2}$

اما التوزيع الشرطي اللاحق الكامل للمتغير γ_j^2 موضح في ادناه حيث سيتم اختيار المقادير الرياضية التي تضم هذا المتغير فقط من التوزيع المشترك :

ان الصيغة الاخيرة تم تقريبها بالاعتماد على توزيع GIG (Generalized Inverse Gaussian) لتصبح بالشكل الاتي :

$$= (\gamma_j^2)^{-\frac{3}{2}} \exp\left[-\frac{\beta_j^2\left((1/\gamma_j^2)-\sqrt{\lambda_j^2\sigma^2/\beta_j^2}\right)^2}{2\sigma^2(\frac{1}{\gamma_j^2})}\right] \dots (5)$$

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

من الصيغة (5) تبين لنا ان توزيع $1/\gamma_j^2$ هو توزيع GIG بمعلمة شكل λ_j^2 ومعلمة متوسط (mean parameter) هي $\sqrt{\lambda_j^2\sigma^2/\beta_j^2}$.

اما التوزيع الشرطي اللاحق الكامل للمتغير λ_j^2 كما موضح في ادناه بعد توظيف المقادير الرياضية التي تضم هذا المتغير من التوزيع اللاحق :

$$= (\lambda^2)^{p+k-1} \exp\left[-\lambda^2 \left(\frac{1}{2}\sum_{j=1}^{p}\gamma_j^2 + \frac{1}{\theta}\right)\right] \dots (6)$$

ان الصيغة الاخيرة ما هي الا توزيع كاما بمعلمة شكل $p + k$ ومعلمة قياس $\left[\frac{1}{2}\sum_{j=1}^{p}\gamma_j^2 + \frac{1}{\theta}\right]$

4-الجانب التجريبي / تجربة المحاكاة First simulation experience

في هذه التجربة تم الاعتماد على تقنية Monte Carlo للمحاكاة بهدف دراسة اداء الطريقة المقترحة ومقارنتها مع الطرائق الاتية :

1-طريقة لاسو التكيفية (A Lasso) Adaptive Lasso التقليدية .

2-طريقة لاسو التكيفية البيزية (BA Lasso) Bayesian Adaptive Lasso

حيث تم توليد بيانات دراسة المحاكاة باستخدام نموذج الانحدار الخطي المتعدد الاتي :

$$y_i = \acute{X}_i\beta + e_i \quad ; i = 1,2, \dots, n$$

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

حيث ان $e_i \sim N(0,\sigma^2)$ و $X_i \sim N(0,1)$ مفترضين ان تجانس البواقي هو مجانس Homogeneous. اضافة الى ذلك افترضنا ان المتغيرات التفسيرية ترتبط مع بعضها البعض بعلاقة الارتباط الاتية :

$$Corr\left(x_i, x_j\right) = \rho^{|i-j|}$$

حيث ان ρ يمثل قيمة معامل الارتباط .

ولكون ان المتغيرات التفسيرية قد تم تحويلها الى قيم معيارية اي انها تمتلك متوسط يساوي صفر وتباين يساوي واحد فان مصفوفة المشاهدات لهذه المتغيرات التفسيرية تتبع التوزيع الطبيعي بمتوسط صفر وتباين $\sum$, اي ان $X \sim N(0,\sum)$ حيث ان $\sum$ هي مصفوفة التباين والتباين المشترك , هنا سوف نفترض ان $\sum_{ij} = \rho^{|i-j|}$. ايضا اشتملت هذه التجربة على افتراض ان قيمة قيم المعالم الحقيقة للمتغيرات التفسيرية هو متجه dense وقيمه كالاتي :

$$\beta = \left(0.75, 0.75, 0.75, 0.75, 0.75, 0.75, 0.75, 075\right)^T$$

اي انه متجه غني بالمعلومات او البيانات , بمعنى اخر ان اغلب قيمه هي قيم غير صفرية .

ولنفترض ايضا ان العلاقة بين متغير الاستجابة (y_i) والمتغيرات التفسيرية يمكن التعبير على انها $E(y) = X\underline{\hat{\beta}}$ اي انه:

$$f(x) = \sum_{j=1}^{8} X_i \beta_j$$

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

$$= \sum_{j=1}^{8} 0.75\, X_j$$

هذا المتجه الافتراضي توجد له متجهات مشابه تم دراستها من قبل Tibshirani في عام 1996 وظهر ايضا في دراسات مثل Li و Fan في عام Zon , 2001 في عام Alhamzawi , 2006 واخرون عام 2012 . في هذه التجربة ايضا تم افتراض قيم تباين الاخطاء $\sigma^2 = 1$ و$\sigma^2 = 4$ لدراسة اثر التباين في البيانات وان قيمة الارتباط $\rho = 0.5$ لدراسة اثر الارتباط الخطي بين المتغيرات , وكذلك تم افتراض عدة قيم لحجوم العينات (50,100,150,200,250) . حيث تم توليد 10000 عينة (itcrations) تمثل قيم المنوال المتولدة من التوزيعات اللاحقة لمعالم نموذج الانحدار باستخدام خوارزمية Gibbs للمعاينة , حيث تم حرق (burn – in) او استبعاد اول 2000 عينة بهدف الحصول على تقديرات مستقرة من خوارزمية Gibbs ومن اجل تقييم دقة التقديرات للطريقة المقترحة والطرائق الاخرى تم تدوين قيم المعالم المقدرة للمعلمة β لكل عينة واحتساب قيم المعيار MMAD من خلال الصيغة الاتية

$$\mathrm{MMAD} = \mathrm{median}\left[\mathrm{mean}\left|\acute{X}_1\,\hat{\beta} - \acute{X}_1\,\beta^{\mathrm{true}}\right|\right]$$

علما انه تم توليد مشاهدات لكل متغير تفسيري بحجم 200 مشاهدة , اي ان MMAD يحسب كالاتي

$$\mathrm{MMAD} = \mathrm{median}\left[\frac{1}{200}\sum_{i=1}^{200}\left|\acute{X}_1\,\hat{\beta} - \acute{X}_1\,\beta^{\mathrm{true}}\right|\right]$$

هنا تم اخذ الوسيط median للقيم المتولدة التي تمثل متوسطات 8000 عملية توليد Simulations الجدول في ادناه نوضح قيم معيار وسيط متوسط الانحرافات المطلقة MMAD ومعيار الانحراف

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

المعياري (SD) والتي تمثل نتائج تجربة المحاكاة للطريقة المقترحة وطرائق المقارنة الاخرى , والهدف من هذا الجدول هو الوصول الى افضل طريقة جزاء تعمل كأسلوب لاختيار المتغيرات .

جدول (1) يمثل قيم معياري MMSD و SD لتجربة المحاكاة

Sample size	Methods	σ^2	MMAD	SD
n=50	Alasso	1	0.37058	0.09106
	BAlasso		0.31915	0.09139
	New BAlasso		0.31087	0.07025
	Alasso	4	0.44396	0.11075
	BAlasso		0.37513	0.10104
	New BAlasso		0.37088	0.09945
n=100	Alasso	1	0.24822	0.02270
	BAlasso		0.22882	0.02978
	New BAlasso		0.12930	0.01884
	Alasso	4	0.03264	00.0651
	BAlasso		0.03063	0.05417
	New BAlasso		0.01742	0.00453
n=150	Alasso	1	0.10453	0.03065
	BAlasso		0.15417	0.02105

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

	New BAlasso		0.10453	0.01845
	Alasso	4	0.04382	0.03422
	BAlasso		0.05710	0.03003
	New BAlasso		0.02316	0.01348
n=200	Alasso	1	0.22419	0.09184
	BAlasso		0.21043	0.09049
	New BAlasso		0.20803	0.08973
	Alasso	4	0.19213	0.06604
	BAlasso		0.19326	0.06729
	New BAlasso		0.23058	0.05131
n=250	Alasso	1	0.05723	0.06712
	BAlasso		0.05843	0.03225
	New BAlasso		0.05723	0.00230
	Alasso	4	0.05631	0.05734
	BAlasso		0.05723	0.04537
	New BAlasso		0.04732	0.015723

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

من الجدول (1) انفا يتبين الاتي :

1- عند حجم العينة الصغير n=50 وتحت قيم التباين $\sigma^2 = 1$ نلاحظ افضلية الطريقة المقترحة New BALasso حيث بلغت قيمة معيار الحكم على جودة التقدير ومدى تطابقها مع القيم الموجودة في المتجه الحقيقي (MMAD = 0.31087) وهي الاقل مقارنة مع طريقة A lasso و B A lasso وايضا نجد ان اقل SD هو للطريقة المقترحة حيث كان (SD = 0.07025) .

2- عند حجم العينة الصغير n=50 وعند زيادة قيم التباين ليكون $\sigma^2 = 4$, اي عند زيادة تشتت البيانات التي تم توليدها من النموذج , نجد ان هذا التشتت قد اثر على قيمة المعيار MMAD مما ادى الى زيادتها ليكون (MMAD = 0.37088) مقارنه مع سابقتها عندما كان $\sigma^2 = 1$, كذلك نجد ان اثر تغير التباين قد انعكس على قيم المعيار MMAD للطرائق الاخرى لكن بقيت الطريقة المقترحة هي الافضل وهذا ايضا واضح من قيمة المعيار (SD = 0.9945) وهو الاقل من بين الطرائق ككل .

3- عند حجم العينة الصغير n=100 و n=150 و n=200 و n=250 وعند قيم التباين $\sigma^2 = 1$ و$\sigma^2 = 4$ نجد ان الطريقة المقترحة New BALasso قد انتجت اقل قيم للمعايير MMAD و SD مما يعني افضلية نتائج هذه الطريقة مقارنة مع الطرائق الاخرى واعتبارها افضل اسلوب لاختيار المتغيرات . حيث ان زيادة عدد المشاهدات يعني حتما زيادة دالة الامكان (الوزن) المضروبة بدالة التوزيع المسبق مما يعني اعطاء اهمية اكبر للبيانات والتالي تأثير اكثر على دالة التوزيع المسبق مما ينعكس ايجابا على النتائج المتولدة من التوزيع اللاحق وبالتالي ينتج عنها اقل قيم للمعايير MMAD و SD .

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

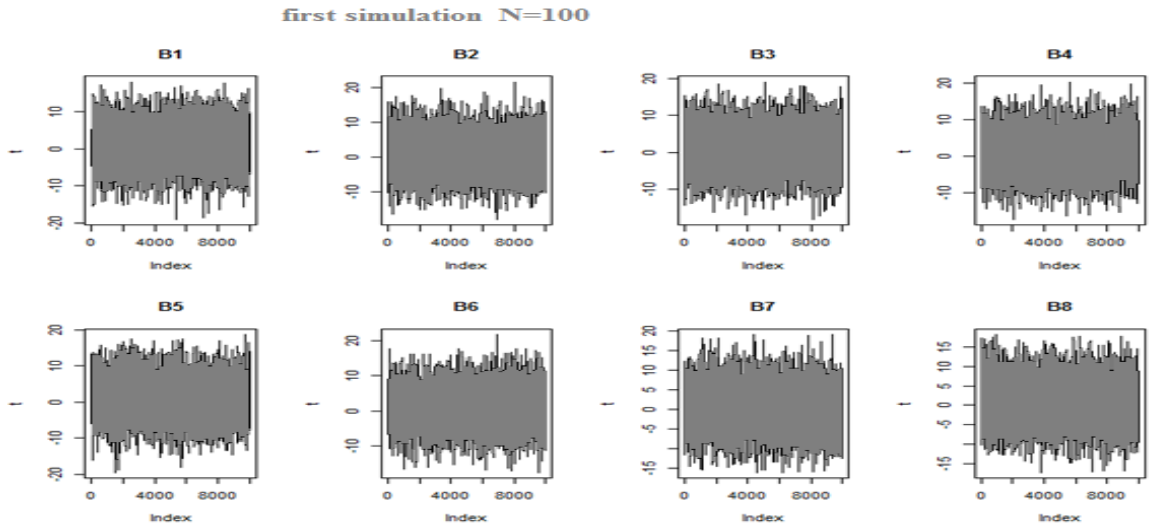
KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

4-نجد ان طريقة B A lasso هي الطريقة المنافسة بعد الطريقة المقترحة New B Lasso هي غير انها تعطي تأثر اقل قيم المعايير MMAD و SD مما يعني ان اسلوب بيز لتقدير معالم الطرائق المختلفة قد تتفوق في النتائج على طريقة بيز التقليدية والكلاسيكية كونه يأخذ بنظر الاعتبار التوزيع السابق لتمثيل الوزن المضروب بدالة الامكان للمشاهدات .في ادناه رسم trace plot لمعالم نموذج الانحدار التي تم احتسابها من خوارزمية معاينة Gibbs . يستخدم هذا الرسم للحكم على اداء الخوارزمية من ناحية تباطئ توليد سلسلة العينات للقيم المقدرة للمعالم , او فيما اذا كانت هناك flat bits , حيث عادة ما يتم تحسين (flat bits) كونها تمثل ان سلسلة العينات المتولدة لقيم المعالم تبقى على نفس حالها لفترة طويلة, كذلك نرغب في رسم trace plot ان نتجنب الخطوات المتتالية في اتجاه واحد (اي ان قيم السلسلة تكون مستقرة باتجاه واحد) . في تجربة المحاكاة هذه تم حرق اول 2000 تكرار ومن بعدها استقرت خوارزمية Gibbs للمعاينة في توليد المعالم المقدرة للنموذج المدروس مما يدل على ان هناك خلط جيد لسلسلة العينات المتولدة . والشكل التالي يوضح رسم trace plot لتجربة المحاكاة الاولى عند اختيارنا لحجم عينة n=200 .

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

شكل (1) رسم trace plot لتجربة المحاكاة للمعالم $\beta_1, \ldots, \beta_8$

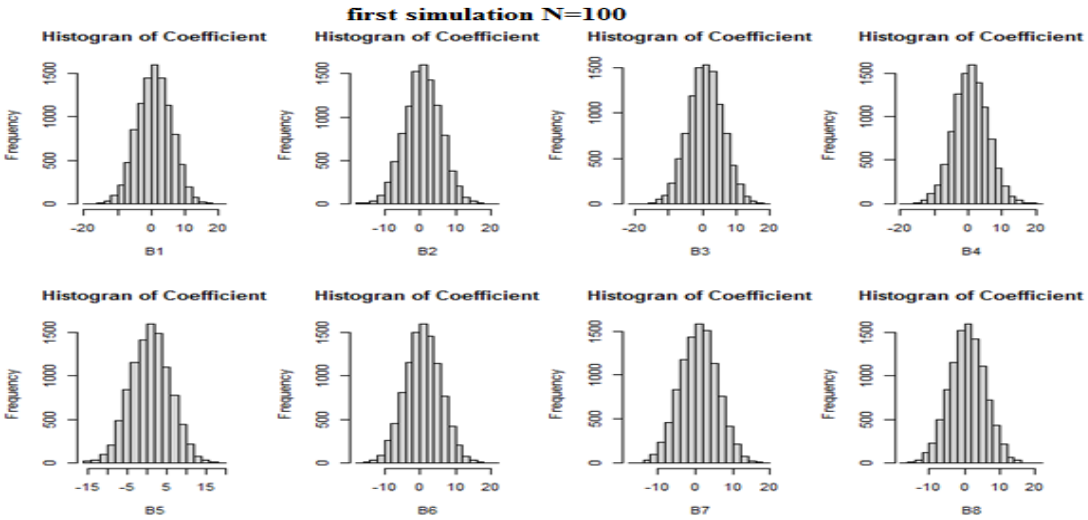


من خلال الشكل (1) نلاحظ ان رسم trace plot لعب دورا مهما باعتباره اداة لتشخيص التقارب في عمل خوارزمية Gibbs للمعاينة حيث يتبين ان الخوارزمية لا تعاني من مشكلة flat bits وبالتالي لا تعاني من التقارب البطيء , اي ان قيم المعالم المتولدة من التوزيع اللاحق تتقارب مع التوزيع الهدف target distribution .ومن اجل اثبات اداء الباحث حول ان التوزيع اللاحق لمعالم نموذج الانحدار β بانها تتبع التوزيع الطبيعي تم رسم المدرج التكراري Histogram للعينات المتولدة من خوارزمية Gibbs وكما موضح بالشكل ادناه :

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

شكل (2) المدرج التكراري لتجربة المحاكاة لمقدرات المعالم $\beta_1, \ldots, \beta_8$



حيث نلاحظ من الشكل (2) ان جميع المعالم $\beta_1, \ldots, \beta_8$ تتبع التوزيع الطبيعي , وهذا ينطبق مع ما جاء في الجانب النظري حيث افترضنا ان المعلمة β تتبع التوزيع الطبيعي . وهذا ينطبق مع جميع النتائج عند حجوم العينات الاخرى .

5-الجانب العملي / مرض نقص الحديد في الدم Iron deficiency disease/ practical side

يعتبر فقر الدم (الانيميا) من الأمراض الشائعة في العراق وخاصة في محافظة بابل ، وهو مرض خطير للغاية حيث يهدد الحياة ويؤدي إلى مضاعفات طبية خطيرة خاصة في حالات الأنيميا الشديدة ، لذلك من

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

الضروري معرفة سبب المرض أو المتغيرات التي تسبب المرض. هذا يساعد على الإصابة بالمرض ومعرفة أهم المتغيرات. مع تطور العالم وتحسن التكنولوجيا ، هناك العديد من الأساليب والتحليلات الإحصائية التي تساعد في دراسة المتغيرات التي لم يتم دراستها من قبل ، ومن أهم هذه الطرق استخدام انحدار لاسو التكيفي البيزي لفهم التأثير الأكبر على المتغير التابع من المتغيرات المستقلة . لذلك يعتبر فقر الدم من أخطر الأمراض التي تصيب الإنسان ، لأنه مرتبط بالدم المسؤول عن جميع الأنشطة الحيوية في الجسم. يحتاج الجسم إلى إمداد مستمر بالدم الغني بالأوكسجين (المحمل بالأوكسجين) ليعمل بشكل صحيح ، ويكمن الخطر في عدم معرفة ما الذي يؤثر على الاصابة بالمرض وفي التعرف على أسباب المرض.

6-متغيرات النموذج Sample variables

تم تحليل نموذج الانحدار الخطي المتعدد لظاهرة نقص الحديد في الدم على فرض وجود ان المتغير المعتمد (y) وهو HCT ويمثل حجم الكريات الحمراء المكدسة بالدم (نسبة الهيموجلوبين في الدم) اما المتغيرات المستقلة (xi) يمكن تفصيلها بان x_1 (WBC) ويمثل نسبة كريات الدم البيضاء و x_2 (Lym) ويمثل نسبة الخلايا اللمفاوية في الدم و x_3 (Gran) ويمثل عدد كريات الدم البيضاء في الدم و x_4 (RPC) ويمثل نسبة كريات الدم الحمراء في الدم و x_5 (HGB) ويمثل نسبة الهيموجلوبين الموجودة في كريات الدم و x_6 (MCV) ويمثل متوسط حجم كريات الدم الحمراء و x_7 (MCH) ويمثل متوسط حجم الهيموجلوبين الموجود في كرية الدم الواحدة و x_8 (MCHC) ويمثل متوسط تركيز الهيموجلوبين في كريات الدم الحمراء و x_9 (PLT) ويمثل نسبة الصفائح الدموية و x_{10} (MPV) ويمثل متوسط

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

حجم الصفائح الدموية و x_{11} (PDW) ويمثل عدد الصفائح الدموية و x_{12} (PCT) ويمثل هرمون البروكالستين الذي يحافظ على الكالسيوم في الجسم و x_{13} (cholesterol) ويمثل نسبة الكولسترول في الدم و x_{14} (R.B Sugar) ويمثل نسبة السكر في الدم و x_{15} (B.Urea) ويمثل نسبة اليوريا في الدم و x_{16} (S.Creatinine) ويمثل الكرياتين في الدم و x_{17} (LDL) ويمثل كمية البروتين الدهني منخفض الكثافة و x_{18} (HDL) ويمثل كمية البروتين الدهني عالي الكثافة و x_{19} (PCV) ويمثل كمية النسبة المئوية لخلايا الدم الحمراء في الدم كاملا" و x_{20} (E.S.R) ويمثل سرعة ترسب الدم و x_{21} (Weight) ويمثل وزن الانسان و x_{22} (Age) ويمثل عمر الانسان .

7- : تحليل البيانات Data analysis

في هذا الجانب تم توظيف طرائق التقدير المشار اليها في تجارب المحاكاة , حيث تم الحصول على مقدرات ومعالم نموذج الانحدار الخطي المتعدد الاتي :

$$y_i = \sum_{i=1}^{22} X_i \underline{\hat{\beta}}$$

والجدول ادناه يوضح تقديرات لمعالم النموذج اعلاه والبالغ عددها (22) معلمة

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

جدول (2) القيم التقديرية لمعالم نموذج الانحدار

Variables	Alasso	BAlaso	New BAlasso
X1	0.02307	0.03389	0.02366
X2	-0.01342	-0.01182	0
X3	0	0	0
X4	0.00139	0.0015	0.0015
X5	1.03856	1.04343	1.03877
X6	0.00465	0.08308	0.01495
X7	0.00584	0	0
X8	0	0	0
X9	0.00013	0.01029	-0.00023
X10	-0.34189	-0.00146	-0.00288
X11	-0.12232	0.00041	-0.00219
X12	-0.03432	-0.00901	-0.00041
X13	-0.00674	0.00253	-0.00037
X14	0.00363	0.001	0.00339
X15	0.00121	0.00234	0.00126

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

X16	-0.21832	-0.00386	0
X17	0.00391	0.00172	0.00359
X18	0	-0.01339	0
X19	0.00339	0.00438	0.00311
X20	0.00672	0.00678	0.00656
X21	0.000865	0.00275	0.0007
X22	0.00305	0.00522	0.0033

من الجدول (2) نجد ان النتائج تظهر تقديرات لمعالم نموذج الانحدار الخطي المتعدد وفق الطرائق المذكورة انفا", حيث نلاحظ ما يلي :

1-ان الطريقة المقترحة New BAlasso عملت كأسلوب لاختيار المتغيرات حيث استبعدت هذه الطريقة المتغيرات ($x_2, x_3, x_7, x_8, x_{16}, x_{18}$) في النموذج المقدر اي ان هذه المتغيرات ليس لها اي تأثير على متغير الاستجابة (نقص الحديد في الدم) . وهنا نستنتج ان المتغيرات المذكورة اعلاه هي (متغير نسبة الخلايا اللمفاوية في الدم و عدد كريات الدم البيضاء في الدم و متوسط حجم الهيموجلوبين الموجود في كرية الدم الواحدة و متوسط تركيز الهيموجلوبين في كريات الدم الحمراء و يمثل الكرياتين في الدم و كمية البروتين الدهني عالي الكثافة) لا يؤثر على مستوى الحديد في الدم ويجب استبعادها من النموذج للحصول على نموذج ذو قدرة تفسيرية عالية .

2-من نتائج الطريقة المقترحة نلاحظ ايضا ان بعض المتغيرات قد انتج عنها تقدير لتأثيرها وهو تأثير قليل جدا على التغير الحاصل في متغير الاستجابة وهذه المتغيرات هي (x_{13}, x_{12}, x_9) وهي نسبة

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

KUFA STUDIES CENTER JOURNAL
UNIVERSITY OF KUFA
مجلة مركز دراسات الكوفة - جامعة الكوفة
KSCJ OPEN ACCESS
ISSN (P) : 1993-7016
ISSN (E) : 2708-728X
www.journal.uokufa.edu.iq

الصفائح الدموية و هرمون البروكالستين الذي يحافظ على الكالسيوم في الجسم و نسبة الكولسترول في الدم على التوالي حيث كان تأثيرها سلبيا على متغير الاستجابة اما المتغير (x_{21}) والذي يمثل وزن الانسان كان تأثيره قليل لكنه تأثير موجب .

8-الاستنتاجات Inference

1-تعتبر مشكلة اختيار المتغيرات (Variable Selection) عند دراسة معادلة الانحدار من المواضيع المهمة سواء" في التحديد المسبق للمتغيرات التفسيرية للظاهرة المدروسة , وحسب النظرية العلمية لهذه الظاهرة او عند عدم التحديد المسبق للمتغيرات التفسيرية للظاهرة المدروسة , اي عند عدم وجود وضوح في تحديد تلك المتغيرات .

2-ان مشكلة اختيار المتغيرات (Variable Selection) مرتبطة مع التشخيص الصحيح لمعادلة الانحدار . حيث نجد ان اهم الاسئلة عند صياغة نموذج الانحدار هو اي المتغيرات التفسيرية . سيتم ادراجها في هذا النموذج , وبأية صيغة سيتم ادراجها . وهذا يعني يجب حل مشكلة اختيار المتغيرات وتشخيص او تحديد نموذج الانحدار بشكل متزامن في وقت واحد

3-ان مشكلة اختيار المتغيرات في موضوع تحليل الانحدار يمكن اعتبارها على انها مشكلة اختيار النموذج (Model Selection) .

4-تعتبر طريقة لاسو التكيفية adaptive lasso تستخدم معلمة تمثل قيمة الوزن الموجب positive weight للسيطرة بشكل مباشر على كمية التحيز من خلال تقليص قيم المعالم بواسطة هذه الكميات المختلفة في الاوزان .

5-تم اقتراح النموذج الهرمي للتوزيعات المسبقة بافتراض ان معلمة التباين σ^2 تتبع توزيع مربع كأي .

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

6-تم توظيف النموذج الهرمي المقترح في اشتقاق التوزيعات اللاحقة , لتقدير معلمة الانحدار الخطي المتعدد بافتراض ان هذه المعلمة تمتلك توزيع لابلاس من خلال التمثيل المختلط الذي اقترحه Park , Casella عام 2008 .

7-بسبب ان التوزيعات اللاحقة المقترحة تمتلك توزيعات معروفة الصيغة . تم توظيف خوارزمية Gibbs للمعاينة لتوليد العينات التي تمثل قيم متوسط معلمة الانحدار . حيث اظهرت النتائج تفوق الطريقة المقترحة على طريقة لاسو التكيفية التقليدية وطريقة لاسو التكيفية البيزية , التي اعتمدت على التوزيع المختلط المقترح من قبل الباحثان Park , Casella عام 2008 , وذلك بالاعتماد على معياري MMAD و SD حيث كانت الطريقة المقترحة تمتلك اقل قيم لهذه المعايير .

8-من خلال الجانب التطبيقي , اظهرت النتائج التي تم فيها استخدام بيانات واقعية لعينة من المرضى الذين يعانون من نقص نسبة الحديد في الدم , ان الطريقة المقترحة قد تفوقت بأدائها كطريقة او اسلوب لاختيار المتغيرات على الطرائق الاخرى المستخدمة في جانب المحاكاة .

9-المصادر References

[1] Efroymson, M. A. (1960). Multiple regression analysis. Mathematical methods for digital computers, 191-203.

[2] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), 267-288.

[3] Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American statistical association, 101(476), 1418-1429 .

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.
مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.

[4] Park, T., & Casella, G. (2008). The bayesian lasso. Journal of the American Statistical Association, 103(482), 681-686.

[5] Leng, C., Tran, M. N., & Nott, D. (2014). Bayesian adaptive lasso. Annals of the Institute of Statistical Mathematics, 66(2), 221-244.

[6] Feng, X. N., Wu, H. T., & Song, X. Y. (2017). Bayesian adaptive lasso for ordinal regression with latent variables. Sociological Methods & Research, 46(4), 926-953.

[7] Flaih , A.N.,Alshaybawee , T., Alhusseini, F .H. H. (2020). Sparsity via new Bayesian lasso. Periodicals of engineering and Natural Sciences.

[8] Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American statistical association, 101(476), 1418-1429.

This work is licensed under a https://creativecommons.org/licenses/by/4.0 International License.

مجلة مركز دراسات الكوفة - جامعة الكوفة مرخصة بموجب ترخيص المشاع الإبداعي 4.0الدولي.