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

دراسة نموذج انحدار لاسو المعمم البيزي باستخدام المحاكاة

أ.د. احمد نعيم فليح

الباحث فريد جخن كريم

جامعة القادسية / كلية الادارة والاقتصاد

DOI: https://doi.org/10.36322/jksc.176(D).19770

الملخص

في هذا البحث تم تناول احد مواضيع نظرية بيز في تقدير معالم نموذج الانحدار الخطي المتعدد حيث تعتبر عملية التقدير في النظرية الإحصائية من اهم المواضيع وأكثرها شيوعا في المجالات العلمية المختلفة. تناولت هذه البحث ما يسمى بطريقة بيز الامنة (Safe Baysian) في تقدير معالم نموذج الانحدار الخطي المتعدد حيث تعتبر هذه الطريقة أسلوبا لحل المشاكل التي ترافق التشخيص الخاطئ لعلاقات الانحدار بين المتغيرات . ان التشخيص الخاطئ لنموذج الانحدار للبيانات المفترضة يؤدي الى تقديرات غير متسقة لمعالم نموذج الانحدار قيد الدراسة مما يعني ان هذه التقديرات هي تقديرات غير مفيدة وبالتالي فان المتغيرات التفسيرية التي اشتمل عليها النموذج لا تفسر التغييرات الحاصلة بمتوسط متغير الاستجابة بصورة صحيحة. تم توظيف طريقة بيز في تقدير معالم نموذج الانحدار المتعدد التي اقترحها الباحثان Yi و Mallick في عام 2014 من خلال استخدام طريقة جزاء لاسو وافتراض ان التوزيع المسبق لمعلمة الانحدار هو توزيع لابلاس يمكن تمثيله من خلال خلط توزيعي $Gamma(2.\lambda)$ والتوزيع المنتظم . في هذا البحث قام الباحث بتوظيف نظرية بيز الامنة Safe Baysian على الأسلوب المقترح من قبل الباحثان Yi و Mallick وتم اجراء فحص تجريبي لسلوك طريقة جزاء لاسو وفق نظرية بيز الامنة من خلال تجربة محاكاة وتبين من خلال نتائج تجربة المحاكاة ان للطريقة المقترحة

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

افضلية في دقة التقدير مقارنة مع الطرائق الأخرى وهذا يعني ان هناك دور مهم لمعلمة التعلم في أسلوب بيز في زيادة احتمالية الحصول على التوزيع الصحيح (Correct) من خلال دور معلمة التعلم في دراسة سلوك توزيع الإمكان.

الكلمات المفتاحية: طريقة لاسو المعممة , الطريقة البيزية , اختيار المتغيرات , المحاكاة.

Studying the Bayesian Generalized Lasso via Simulation

Prof. Dr Ahmed Naim Flaih

Researcher Freed Jkheen Kareem

Al-Qadisiyah University / College of Administration and Economics

Abstract:

In this paper, one of the topics of Bayes' theory was dealt with in estimating the parameters of the multiple linear regression model, as the estimation process in statistical theory is one of the most important and common topics in various scientific fields. This paper dealt with the so-called Safe Bayesian method in estimating the parameters of the multiple linear regression model, as this method is considered a method for solving problems that accompany the misdiagnosis of regression relationships between variables. The misdiagnosis of the regression model for the assumed data leads to inconsistent estimates of the parameters of the regression model under study, which means that these estimates are unhelpful,

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

and therefore the explanatory variables included in the model do not correctly explain the changes in the average of the response variable. The Bayes method was employed in estimating the parameters of the multiple regression model proposed by the researchers Mallick and Yi in 2014 by using the lasso penalty method and assuming that the a priori distribution of the regression parameter is a Laplace distribution that can be represented by mixing the two distributions of Gamma (2,λ) and the Uniform distribution. In this paper, the researcher employed the Bayesian Safe theory on the method that proposed by Mallick and Yi, and an experimental examination was conducted for the behavior of the Lasso penalty method according to the Bayesian Safe theory through several simulation experiments, clearly the proposed method outperformed the other exits methods. This means that there is an important role for the learning parameter in the Bayes method in increasing the probability of obtaining the correct distribution through the role of the learning parameter in studying the behavior of the likelihood distribution.

Keywords: generalized lasso Method, Bayesian Method, Variable selection, Simulation.

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

المقدمة (Introduction)

كثيراً ما يتردد خصوصاً عند العمل في مجال تحليل نماذج الانحدار الخطي المتعدد السؤال الآتي : هل كثرة عدد المتغيرات التفسيرية هو ميزة او صفة للنموذج ام لا؟ الجواب هو لا لأنه اثبتت الكثير من الدراسات في مجال تحليل الانحدار أنه على الباحث أن يكون حذر في أختيار المُتغيرات التفسيرية وقد يكون الأمر خارج السيطرة عندما تُفرض نتائج التحليل واقع معين في الأختبار اي أن أختبار النماذج بعناية يؤدي الى تحسين دقة النموذج المقدر حيث ان اضافة عدد كبير من المتغيرات التفسيرية يمكن ان يؤدي الى حالة تُدعى (over fitting) وبالتالي تُدعى النماذج المُقدرة بنماذج (over fitting) اي أن النماذج التي تصف تأثير الحد العشوائي اي تأثير الأخطاء العشوائية على متغير الأستجابة أكثر مما تصف تأثير المُتغيرات التوضيحية (التفسيرية) على مُتغير الاستجابة وهذا يؤدي بالتالي الى أداء تنبؤي ضعيف للنموذج المقررعند أستخدامه على بيانات الأختبار (test data). بمعنى آخر ان النموذج الذي يتصف بحالة (over fitting) قد تمثل مجموعة معينة من البيانات أفضل تمثيل اي انه يلائم هذه البيانات أفضل ملائمة مما ينتج عنه خط انحدار او منحني يمر بالقرب من مشاهدات العينة ولكن عند تطبيق هذا النموذج على بيانات جديدة مسحوبة من نفس المجتمع فأننا نجد هذا النموذج لا يمثل او يلائم البيانات افضل تمثيل.أحياناً في مجال تحليل البيانات الطبيعية قد نواجه تجارب يكون فيها حجم العينة (n) قليل مقارنة مع عدد المتغيرات التوضيحية (p) وهذا واضح في تجارب تحليل (DNA) ففي مثل هذه التجارب تصبح عملية التنبؤ باختيار مجموعة (K) من (p) من المتغيرات أي اختبار مجموعة جزئية (subset) من المتغيرات التوضيحية يمكن ان يؤدي إلى تقدير نموذج انحدار أفضل من ناحية قدرة التنبؤ وبعدد (k) من المتغيرات التي ينتج عنها اقل مجموع مربعات خطأ.على مدار السنين السابقة تم تطوير الكثير من طرائق

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

وأساليب أختبار المتغيرات وأختبار النماذج حيث سيتم التطرق إلى هذه الطرائق في المباحث التالية من الفصل. أضافة الي ذلك فأن موضوع تحليل الأنحدار يعتبر من الادوات الاحصائية المهمة في تحليل البيانات الطبيعة التي تتطلب أيجاد نموذج يدرس العلاقة بين متغير الاستجابة ومجموعة من المتغيرات التوضيحية حيث يساعد تقدير معادلة الانحدار على فهم المخاطرة التي تنتج عن بعض الامراض والكشف عنها مبكراً مما يساعد على منع تطور معالجة هذا المرض مبكراً من خلال التدخلات الطبية اللازمة اضافة إلى ذلك فأن موضوع اختبار المتغيرات في الحالات الطبية يساعد سهولة التواصل بين الطبيب والمرضى.ومن هنا يمكن القول ان النماذج الاحصائية (نماذج الانحدار الخطي المتعدد) هي بمثابة اداة اساسية للباحثين في مجال الصحة العامة. حيث ان نجاح الباحث في أيجاد النموذج الافضل تعتمد بالتأكيد على عدة عوامل منها اختيارمجموعة مناسبة من المرضى، اختبار النموذج الصحيح، وتفسير النتائج بشكل دقيق. في هذه البحث سيتم التطرق إلى طريقة (لاسو المعممة) (Generalized lasso) وتقدير معالم نموذج الانحدار وفق اسلوب (بيز) اي ان هذه البحث سوف تحاول كشف تأثير أسلوب أختيار المتغيرات وبالتالي دراسة تأثير تقدير معالم النموذج وفهم العلاقة بين المتغيرات. يعتبر اسلوب اختيار المتغيرات (variable selection) من المواضيع الحيوية المهمة عند تحليل نماذج الانحدار لذلك فقد تطورت الاساليب الاحصائية المستخدمة في اسلوب أختيار المتغيرات على مدار السنين السابقة. هذا البحث ركز على حل مشكلة أختيار اهم المتغيرات التي لها تأثير على الاصابة بمرض الفشل الكلوي اي ان هذه البحث تساعد العاملين في المجال الطبي خصوصاً العاملين في مجال امراض الفشل الكلوي والمرضى المصابون بهذا المرض على فهم اهم العوامل التي تؤثر على المصابين وبالتالي سبل معالجة المرضى وايجاد افضل اساليب التواصل بين المختصين والمرضى. ومن الناحية الاحصائية تكمن مشكلة البحث في حل مشكلة

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

التشخيص الخاطئ لنموذج الانحدار الذي يمثل بيانات التجربة حيث من المعلوم ان التشخيص الخاطئ للنموذج سوف يؤدي إلى أيجاد متغيرات غير متسقة وهذا ما يؤثر على دقة التقدير ودقة التنبؤ لذلك نلجأ إلى استخدام ما يسمى (اسلوب بيز الآمن). في عام ١٩٩٦ قدم الباحث (Tibshirani) أول بحث عن طرائق الجزاء التي تعمل كأسلوب لاختيار المتغيرات. وقد سميت هذه الطريقة او الاسلوب بطريقة (لاسو lasso) حيث يقوم مبدأ هذه الطريقة على ايجاد مقدرات معالم نموذج الانحدار الخطي المتعدد من خلال تصغير مجموع مربعات الخطأ التي تخضع لشرط معين يدعى بدالة الجزاء. أن دالة الجزاء او الشرط يعمل على أيجاد قيمة معلمة الجزاء او (الانكماش shrinkage parameter) أو (معلمة الضغط tuning parameter) التي تجعل مجموع مربعات الخطأ أقل ما يمكن وتحت شروط معينة. تم تنفيذ خوارزمية حسابية تمثل مقدرات النموذج لأن تكون مساوية للصفر مما يجعلها أسلوباً لاختبار المتغيرات. كذلك ذكر الباحث Tibshirani . في هذا البحث يمكن استخدام اسلوب التقدير وفق اسلوب بيز او نظرية بيز بأفتراض أن معلمة الانحدار β_j هي متغير عشوائي يتبع في سلوكه متغير يتوزع وفق توزيع لابلاس كتوزيع اولي للمعلمة وبالتالي يمكن أستخراج القيمة المقدرة للمعلمة β على انها قيمة المنوال للتوزيع اللاحق. عام ٢٠٠٨ قدم الباحثان (Casella وPark) اول بحث تناول فكرة ان معالم نموذج الانحدار الخطي يتبع توزيع لابلاس كتوزيع مُسبق وذلك بالأعتماد على الفكرة التي أشار اليها الباحث Tibshirani في عام ١٩٩٦. حيث تم وضع نموذج هرمي للتوزيعات المُسبقة تحت افتراض ان النموذج المُسبق لمعلمة الأنحدار يمكن تمثيلها من خلال توزيع مُختلط يمثل توزيع لابلاس. هذا التوزيع المُختلط مبني على خط التوزيع الطبيعي مضروباً بالتوزيع الأسي. وبناءاً على النموذج الهرمي تم اشتقاق التوزيع اللاحق للمعالم الخاصة بالنموذج بأفتراض ان تقدير معلمة الانحدار β تمثل متوسط التوزيع اللاحق لهذه المعلمة. أضافة

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

إلى ذلك وسع فكرة البحث على الانحدار المُسمى (أنحدار Bridge) حيث تم تنفيذ خوارزمية كبس لتقدير معالم النموذج المقترح على بيانات تجريبية وبيانات حقيقية وأظهرت النتائج تفوق طريقة بيز على الطريقة التقليدية OLS وفق أسلوب طريقة الجزاء Lasso في اختيار المتغيرات. في عام ٢٠١٢ قدم الباحث Grunwald بحثاً عن طريقة saf-Bayesian حيث اعتمد فكرة هذه الطريقة على ان الاستدلال وفق اسلوب بيز التقليدي لا يؤدي إلى حلول مثلى اذا كان نموذج الانحدار المقدر لا يلائم البيانات بشكل دقيق اي ان طريقة بيز الآمنة هي بمثابه تعديل لبيز التقليدية , اي انه تم تدريب النموذج الخاطئ بمعادلات تعلم حيث تم الاستعانة بما يسمى معلمة معدل التعلم وأستخدامها لتحسين دالة البيانات likelihood حيث تم استخدام قيم مختلفة لمعلمة معدل التعلم لتقليل دالة الخسارة التجميعية . اثبت الباحث كفاءة الاسلوب المقترح من خلال عدة امثلة نظرية لاثبات صحة ادعائه حيث اثبت ان الاسلوب المقترح نظريا هو اسلوب كفوء مقارنة مع اسلوب بيز التقليدي. في عام ٢٠١٤ قدم الباحث (Yi, Mallick) بحثاً تناول تقدير معالم نموذج الانحدار الخطي المُتعدد وفق أسلوب بيز وبأفتراض وجود دالة جزاء لاسو. لكن عملهم يختلف عن عمل (Casella, Park) في عام ٢٠٠٨ حيث ان التوزيع المُسبق لمعالم نموذج الأنحدار β تم تمثيلها من خلال توزيع مختلط بناءاً على معلمة القياس وهذا التوزيع المختلط يمثل حاصل ضرب التوزيع المُنتظم مضروباً مع توزيع گاما Gamma$(2,\lambda)$ حيث تم أفتراض نموذج هرمي للتوزيعات المشتقة وبالتالي أشتقاق التوزيعات اللاحقة للمعالم المراد تقديرها. أشتملت الدراسة عدة تجارب محاكاة لأختبار دقة التنبؤ للنموذج المُقترح حيث أظهرت النتائج تفوق الطريقة المُقترحة مقارنة مع مع طريقة (Lasso) التقليدية وطريقة بيز لاسو. اضافة الى ذلك تم تحليل بيانات حقيقية مُختلفة واظهرت النتائج أن النموذج المُقترح يمتلك تقارب لخوارزمية كبس يمكنها من توليد قيم حقيقية للمعالم المُقدرة

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

واختبار المتغيرات. في عام ٢٠١٦ قدمت الباحثة Heide رسالة ماجستير حول موضوع (safe – Bayes) وفق طريقة جزاء لاسو حيث طورت في هذا البحث النموذج الهرمي والتوزيعات اللاحقة بافتراض وجود معلمة معدل التعلم . تم تقدير معالم نموذج الانحدار الخطي المتعدد حسب طريقة (الجزاء لاسو) بافتراض وجود معلمة معدل التعلم في دالة التوزيع للبيانات (Likelihood) اظهرت النتائج ان الطريقة المُقترحة تتفوق في كل من الجانب التجريبي وجانب تحليل البيانات الحقيقية على طريقة بيز التقليدية. في عام ٢٠٢٠ قدمت الباحثة Heide وآخرون بحثاً في الانحدار الخطي المعمم وتقدير المعالم وفق اسلوب بيز الآمن safe – Bayesian . ويسمى هذا الاسلوب باسلوب بيز المعمم حيث تنطوي فكرة هذا الاسلوب على رفع دالة الامكان likelhood إلى قوة تتمثل بمعلمة تسمى معلمة معدل التعلم (learning – rate) أي انه سيتم تحسين التوزيع اللاحق للمعلمة المقدرة من خلال ضرب التوزيع السابق بدالة الامكان المعممة المرفوعة لمعلمة التعلم . ان مبدأ هذه الفكرة جاء من ان كل نماذج الانحدار التي تم تقديرها تفتقر إلى عدم اليقين او التاكد من انها تلائم البيانات بشكل واقعي لذلك تم تدارس رفع دالة الامكان لمعلمة التعلم في سبيل تدريب النموذج للحصول على افضل تقدير لمعالم هذا النموذج . تم اشتقاق خوارزمية كبس بناءاً على التوزيعات اللاحقة المقترحة وتوظيفها في الانحدار اللوجستي حيث أظهرت النتائج تفوق الاسلوب المقترح وكل في جانب المحاكاة والجانب التحليلي للبيانات الواقعية على اسلوب بيز التقليدي. في عام ٢٠٢٠ قدم الباحثان WuوMartin بحثاً عن طريقة لاسو المعممة اي بوجود معلمة معدل التعلم او بكلام آخر طريقة لاسو الآمنة حيث قدم هذا البحث عدة أساليب مقترحة لاختيار معلمة معدل التعلم. هذه الأساليب المقترحة تهدف إلى التغلب على مشكلة التشخيص الخاطئ للنموذج الملائم للبيانات من خلال اختيار افضل معلمة تعلم تساعد على اجراء التنبؤ بأستخدام بيز المعمم.

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

في هذا البحث تم أجراء دراسة لقياس معالم نموذج الانحدار اللوجستي بوجود معلمة التعلم ومن خلال عدة امثلة تطبيقية تبين ان طريقة لاسو المقترحة تحت معلمة التعلم وافتراض عدة طرائق لتقدير هذه المعلمة قد اثبتت كفائتها مع طرائق تقدير أخرى بدلالة التقارب الاحتمالي لفترات Credible لمعالم نموذج الأنحدار اللوجستي المقدرة.

ان مساهمتنا في هذا البحث سيكون من خلال توظيف معلمة معدل التعلم في دالة الامكان او البيانات عند اشتقاق التوزيع اللاحق لنموذج الانحدار الخطي بوجود النموذج الهرمي للتوزيعات المسبقة التي اقترحها الباحثان Mallick, Yi في عام ٢٠١٤ وبالتالي اشتقاق توزيعات لاحقة جديدة وتوظيفها لايجاد خوارزمية كبس جديدة لتوليد العينات من هذه التوزيعات اللاحقة حيث لم يسبق لأي دراسة ان عملت في توظيف هذا النوع من safe-Bayes في النموذج الذي اقترحه الباحثان Mallick وYi .

2- النموذج الهرمي للتوزيعات المسبقة وفق لاسو المعممة

Hierarchical of prior's model for Generalized lasso

في هذا المبحث تم افتراض ان نموذج الانحدار قيد الدراسة هو نموذج انحدار خطي متعدد ومعرف بالصيغة الأتية :

$$Y = X\beta + e \dots\dots\dots (1)$$

حيث ان :

Y: هو متجه متغير الاستجابة $(n \times 1)$ وقد تم تحويل قيم هذا المتجه الى قيم مركزية centering أي ان $(y_i - \bar{y})$. حيث ان$y \sim N(X\beta, \sigma^2)$, X هي مصفوفة القيم المعيارية للمشاهدات $(n \times p)$, β هو

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

متجه معامل النموذج المراد تقديرها $(n \times 1)$.وان حد الخطأ يتبع التوزيع الطبيعي. وسوف نفرض ان المعلمة β تتبع التوزيع التمثيل الاتي:

$$\pi(\beta|\sigma^2,\lambda) = \frac{\lambda}{2\sqrt{\sigma^2}}\exp\left[\frac{-\lambda|\beta|}{\sqrt{\sigma^2}}\right.$$

$$= \int_{-w\sqrt{\sigma^2}<\beta<w\sqrt{\sigma^2}} \frac{1}{2w\sqrt{\sigma^2}}.\frac{\lambda^2}{\sqrt{2}}.w^{2-1}e^{-\lambda w}dw \dots\dots\dots (2)$$

سوف نعتمد على نموذج الأنحدار (1) والصيغة (2) واضافة معلمة معدل التعلم learning- α حيث ان النموذج الهرمي للتوزيعات المسبقة وفق طريقة لاسو البيزية المعممة كالآتي :

$$y|\, X,\beta,\sigma^2 \sim [N(X\beta,\sigma^2 I_n)]^{\alpha}$$

$$\beta|\, w,\sigma^2 \sim \prod_{j=1}^{p} \mathrm{UNIF}\left(-\sqrt{\sigma^2}w_j, \sqrt{\sigma^2}w_j\right)$$

$$w|\lambda \sim \prod_{j=1}^{p} \frac{\lambda^2}{\sqrt{2}} w_j^{2-1} e^{-\lambda w_j} dw_j$$

$$\sigma^2 \sim \pi(\sigma^2) \dots\dots\dots (2-6)$$ **خوارزمية Gibbs للتوزيعات اللاحقة**

1-2: خوارزمية Gibbs للتوزيعات اللاحقة

وبهذا يمكن تلخيص عمل خوارزمية كبس للمعاينه من اجل توليد قيم العينات للتوزيعات اللاحقة كالاتي :

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

1. توليد العينات للمتغير y : حيث سيتم توليد المشاهدات للمتغير y من التوزيع الطبيعي الاتي: $\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{\alpha n}{2}} \exp\left\{-\frac{\alpha}{2\sigma^2}(\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}\beta)\right\}$

2. توليد العينات للمتغير (w) : حيث سيتم توليد المشاهدات للمتغير $(w)^*$ من التوزيع الاسي المعياري بالمعلمة (λ) وبالتالي جعل $W_j = W_j^* + \frac{|\beta_j|}{\sqrt{\sigma^2}}$: . تم استخدام الخوارزمية التي اقترحها (Li and ghosh,2013) لتنفيذ هذه الخطوة.

3. توليد العينات للمتغير (β) : حيث سيتم توليد المشاهدات للمتغير (β) من التوزيع الطبيعي المبتور : $\beta \sim$ truncated multivarate normal $(\alpha A^{-1}\mathbf{X}'\mathbf{y}, \sigma^2 A^{-1})$.

4. توليد العينات للمتغير (σ^2) : حيث سيتم توليد المشاهدات للمتغير σ^2 من توزيع كاما المعكوس

$$\sigma^2 \sim \text{Inverse gamma}\left(\frac{\alpha n + p}{2}, \quad \frac{\alpha}{2}(\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}\beta)\right)$$

5. توليد العينات للمتغير λ : حيث سيتم توليد المشاهدات ل λ من توزيع كاما بمعلمه شكل $(2p+1)$ ومعلمة قياس (k). تم الاستعانة بخوارزمية اقترحها (Leng et al, 2014) $(\theta + \sum_{j=1}^{p}|\beta_j|)$ لتوليد قيم المعلمة λ .

1- المحاكاة Simulation

تضمن هذا المبحث تجربتين للمحاكاة وفي كل تجربة تم افتراض وجود متجه حقيقي للمعالم المقترحة لنموذج الانحدار الخطي المتعدد قيد الدراسة إضافة الى افتراض انه تم توليد مشاهدات لمتغيرات

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

مصفوفة X وفق حجوم عينات مختلفة وتحت قيم لتباين الأخطاء المقترحة وفي ادناه شرح مفصل لتجربتي المحاكاة . حيث تم الاستعانة بحزمة glmnet من برنامج R , (Friedman et al.,2010).

1-3 : وصف التجربة .

في هذه التجربة كانت عملية توليد قيم مشاهدات العينات لمتغيرات نموذج الانحدار الخطي المتعدد الاتي:

$$\mathbf{Y} = \mathbf{X}\beta + \boldsymbol{\varepsilon}$$

حيث افترض الباحث ان قيم مشاهدات المصفوفة X تم توليدها من التوزيع الطبيعي بمتوسط صفر وتباين واحد أي ان قيم المصفوفة X قد تم توليدها على انها قيم معيارية $X\sim N(0,1)$

وكذلك تم افتراض ان حد الخطأ العشوائي (ε) يتبع التوزيع الطبيعي أيضا بمتوسط صفر وتباين σ^2 أي ان $\varepsilon\sim N(0,\sigma^2)$ إضافة الى ذلك فانه لكل قيمة من قيم X هناك توجد قيمة مشاهدة لمتغير الاستجابة Y يتم توليدها التوزيع الطبيعي بمتوسط صفر وتباين σ^2 أي ان $Y\sim N(0,\sigma^2)$ بافتراض قيم ثابتة للتباين σ^2.

هنا افترض الباحث ان المتغيرات التفسيرية في المصفوفة X مترابطة مع بعضها البعض بمعامل الارتباط الاتي :

$$Corr(X_i, X_j) = \rho^{|i-j|}$$

وبسبب ان قيم مشاهدات المصفوفة X هي قيم معيارية أي انه تباينا تها واحد صحيح فهذا يعني ان مشاهدات المتغيرات التفسيرية في المصفوفة X تنتج التوزيع الطبيعي متعدد للمتغيرات الاتي :

$$X\sim N_n(0,\sum)$$

حيث ان : $\sum$ هي مصفوفة الارتباطات والمحسوبة وفق الصيغة الاتية :

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

$$\sum_{ij} = \rho^{|i-j|}$$

افترض الباحث ان $\rho = 0.5$ أي ان قيمة معامل الارتباط بين أي متغيرين هو 0.5 لاختبار اثر الارتباط المتعدد كذلك افترضنا ان المتجه الحقيقي لمعالم نموذج الانحدار الخطي المتعدد يأخذ القيم الاتية :

$$\beta = (0.85, 0.85, 0.85, 0.85, 0.85, 0.85, 0.85, 0.85)$$

وبذلك فان العلاقة بين متغير الاستجابة Y والمتغيرات التفسيرية X تأخذ الصيغة الاتية :

$$E(Y) = X\beta$$

وهذا يعني :

$$f(X) = \sum_{j=1}^{8} X_j \beta_j$$

وأيضا نجد ان النموذج الحقيقي يأخذ الصيغة الاتية :

$$f(X) = 0 \cdot 85X_1 + 0 \cdot 85X_2 + \cdots + 0 \cdot 85X_8$$

في هذه التجربة افترضنا انه تم توليد مشاهدات لحجوم عينات مختلفة أي ان $n \in (50,100,150)$ كذلك تم افتراض ان $\sigma^2 = 3$ لاختبار اثر تشتت البواقي في البيانات . ان تجربة المحاكاة تمت من خلال تنفيذ 10000 تكرار (iterations) وتم حرق 2000 تكرار بهدف الحصول على استقرار لتوليد المشاهدات من خوارزمية Gibbs للمعاينة وقد استخدم معيار التحيز (Bias) للحكم على دقة التقدير والذي يحسب وفق الاحصاءة الاتية :

$$\text{Bias}\left(\hat{\beta}_j\right) = \bar{\beta}_j - \beta_j$$

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 $\bar{\beta}_j$ يحسب من خلال الاحصاءة التالية ويمثل متوسط التقدير :

$$\bar{\beta}_j = \sum_{j=1}^{n} \frac{\hat{\beta}_j}{n}$$

حيث ان n يمثل حجم العينة للمشاهدات التي تم توليدها كذلك تم احتساب المعيار MSE وحسب الاحصاءة الاتية :

$$MSE(\hat{\beta}_j) = Var(\hat{\beta}_j) + (\beta ias(\hat{\beta}_j))^2$$

حيث ان :

$Var(\hat{\beta}_j)$ يمثل تباين التقدير ويحسب وفق الاحصاءة الاتية :

$$Var(\hat{\beta}_j) = \sum_{j=1}^{n} (\hat{\beta}_j - \bar{\beta}_j)^2 \; / \; n - 1$$

أيضا تم احتساب المعيار MAE وفق الاحصاءة الاتية

$$MAE = \sum_{j=1}^{n} |\hat{\beta}_j - \bar{\beta}_j| \; / \; n$$

وكذلك تم حساب المعيار SD وفق الاحصاءة الاتية :

$$SD = \sqrt{var(\hat{\beta}_j)}$$

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

الذي يعمل بمثابة الخطا المعياري لكل مقدر إضافة الى ذلك تم استخدام معيار K-fold لحساب قيم معلمة التعلم α وحسب الاحصاءة الاتية :

$$CV(\alpha) = \frac{1}{K}\sum_{k=1}^{K} error_k(\alpha)$$

حيث ان$error_k(\alpha)$: يحسب وفق الاحصاءة الاتية :

$$error_k(\alpha) = \sum_{i \in kth} (y_i - \hat{y}_j^{(k)}(\alpha))^2$$

حيث ان :

$\hat{y}_j^{(k)}$ هي القيمة المقدرة للمتغير y_i وحسب طريقة التقدير المستخدمة سواء وفق طريقة Ols او Lasso او غيرها من طرائق التقدير وتجمع المشاهدات بدون القطاع k .

في ادناه نتائج وتحليل تجربة المحاكاة الأولى مستخدمين طريقة التقدير المقترحة ومقارنة النتائج مع كل من طريقة Lasso التقليدية Classo وطريقة بيز لاسو Blasso وتحت قيم مختلفة لقيم معلمة التعلم α .

جدول (1) قيم المعالم المقدرة تحت حجوم عينات مختلفة لتجربة المحاكاة الاولى

Samples	Methods	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
n=50	CLasso	0.74159	0.73081	0.74952	0.73756	0.72999	0.72387	0.70838	0.72493
	NBL(α=0.5)	0.82359	0.81134	0.83343	0.81934	0.81095	0.80390	0.78754	0.80535
	NBL(α=0.7)	0.86457	0.85217	0.87420	0.86023	0.85137	0.84394	0.82723	0.84595
	NBL(α=0.9)	0.86642	0.85181	0.87711	0.86035	0.85277	0.84501	0.82653	0.84526

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

	BLasso	0.75552	0.83955	0.70487	0.71851	0.79837	0.72645	0.80079	0.75715
n=100	Classo	0.73215	0.73813	0.72805	0.72689	0.73251	0.71142	0.74397	0.74492
	NBL(0.5)	0.81309	0.82008	0.80860	0.80709	0.81388	0.79023	0.82596	0.82735
	NBL(0.7)	0.85338	0.86117	0.84910	0.84798	0.85491	0.82972	0.86720	0.86871
	NBL(0.9)	0.85367	0.86069	0.84940	0.84790	0.85499	0.82955	0.86714	0.86877
	BLasso	0.75839	0.77281	0.76455	0.80445	0.68263	0.76929	0.80303	0.76354
n=150	Classo	0.72820	0.72240	0.72562	0.72343	0.72458	0.72502	0.73069	0.74079
	NBL(0.5)	0.80896	0.80233	0.80605	0.80372	0.80484	0.80529	0.81140	0.82250
	NBL(0.7)	0.84943	0.84233	0.84624	0.84370	0.84503	0.84538	0.85189	0.86377
	NBL(0.9)	0.84791	0.84251	0.84386	0.84374	0.84612	0.84645	0.85132	0.86516
	BLasso	0.80463	0.80001	0.76919	0.78445	0.79784	0.77535	0.77554	0.81019

من الجدول رقم (1) انفا يتبين ما يلي:

1. عند حجم العينة $n = 50$ وهو ما يعتبر حجم عينة صغير نجد ان طريقة لاسو التقليدية كانت قيم المعالم المقدرة بعيدة تقريبا عن القيم الحقيقية في المتجه الذي تم افتراضه ذو القيم (85 ٠ 0) لكن نجد ان الطريقة المقترحة قد تطابقت قيم معالمها المقدرة مع قيم المتجه الحقيقي حيث نلاحظ انه كلما زادت قيم معلمة التعلم نجد تطابق القيم المقدرة مع القيم الحقيقية وأخيرا نجد كذلك ان طريقة لاسو البيزية Blasso كانت قريبة لكنها لا تتطابق مع قيم المتجه الحقيقي .

2. عند زيادة حجم العينة ليكون $n = 100$ و $n = 150$ نلاحظ ان نتائج أداء الطريقة المقترحة تفوقت على أداء عمل كل من طريقتي لاسو التقليدية ولاسو البيزية وعلى الرغم من ان حجوم العينات (150,100) اكبر من حجم العينة $n = 50$ مما يعني انها توفر معلومات اكبر في توزيع دالة الإمكان مما يؤدي مع ازدياد قيم معلمة التعلم الى ازدياد في أهمية دالة الإمكان وتأثيرها في التوزيع اللاحق .

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

في الجدول ادناه (2) تم احتساب قيم الانحراف المعياري (SD) للمعالم المقدرة في الجدول (1) والتي تعطي مؤشرا على جودة ودقة التقدير.

جدول (2) قيم الانحراف المعياري للمعالم المقدرة لتجربة المحاكاة الاولى

Samples	Methods	SD1	SD2	SD3	SD4	SD5	SD6	SD7	SD8
n=50	Class0	0.07142	0.07659	0.09037	0.08374	0.07691	0.06484	0.07546	0.06465
	NBL(0.5)	0.07972	0.08543	0.10133	0.09338	0.08525	0.07197	0.08384	0.07184
	NBL(0.7)	0.08326	0.08916	0.10601	0.09781	0.08972	0.07595	0.08775	0.07510
	NBL(0.9)	0.08440	0.08907	0.10647	0.09796	0.09167	0.07588	0.08734	0.07466
	BLasso	0.08387	0.08925	0.10844	0.10071	0.09192	0.08295	0.08746	0.07896
n=100	CLasso	0.08857	0.08083	0.08848	0.08452	0.06941	0.06957	0.07071	0.06894
	NBL(0.5)	0.09812	0.08991	0.09838	0.09386	0.07640	0.07747	0.07886	0.07695
	NBL(0.7)	0.10300	0.09437	0.10315	0.09866	0.08068	0.08117	0.08307	0.08023
	NBL(0.9)	0.10536	0.09345	0.10282	0.09675	0.08032	0.08100	0.08259	0.08065
	BLasso	0.11077	0.09329	0.10341	0.09690	0.08986	0.08611	0.08583	0.08320
n=150	Classo	0.04039	0.03357	0.03423	0.03750	0.04139	0.03939	0.04273	0.04238
	NBL(0.5)	0.04489	0.03725	0.03822	0.04184	0.04610	0.04357	0.04757	0.04702
	NBL(0.7)	0.04715	0.03913	0.04002	0.04382	0.04841	0.04570	0.04982	0.04947
	NBL(0.9)	0.04719	0.03898	0.03950	0.04381	0.04849	0.04549	0.04979	0.04987
	BLasso	0.04828	0.04035	0.04427	0.04760	0.04799	0.04678	0.04944	0.05024

حيث نلاحظ من الجدول ان افضل قيم للانحراف المعياري هي القيم الأقل كذلك نلاحظ ان قيم الانحراف المعياري تتناقص تدريجيا كلما زاد حجم العينة وزيادة قيم معلمة التعلم كما يعني ان البيانات وتوزيع دالة الإمكان كان لها وزن واهمية في التأثير على التوزيع اللاحق مما انعكس على جودة القيم المقدرة لمعالم النموذج من خلال هذا التوزيع اللاحق .

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

في ادناه الجدول (3) والذي تم فيه احتساب قيم معيار التحيز (Bias) للمعالم المقدرة في الجدول رقم (1).

جدول (3) قيم معيار التحيز للمعالم المقدرة لتجربة المحاكاة الاولى

Samples	Methods	Bias 1	Bias 2	Bias 3	Bias 4	Bias 5	Bias 6	Bias 7	Bias 8
n=50	CLasso	0.10841	0.11919	0.10048	0.11244	0.12001	0.12613	0.14162	0.12507
	NBL(0.5)	0.02641	0.03866	0.01657	0.03066	0.03905	0.04610	0.06246	0.04465
	NBL(0.7)	0.01457	0.00217	0.02420	0.01023	0.00137	0.00606	0.02277	0.00405
	NBL(0.9	0.01642	0.00181	0.02711	0.01035	0.00277	0.00499	0.02347	0.00474
	BLasso	0.09448	0.01045	0.14513	0.13149	0.05163	0.12355	0.04921	0.09285
n=100	CLasso	0.11785	0.11187	0.12195	0.12311	0.11749	0.13858	0.10604	0.10508
	NBL(0.5)	0.03691	0.02992	0.04140	0.04291	0.03612	0.05977	0.02404	0.02265
	NBL(0.7)	0.00338	0.01117	0.00090	0.00202	0.00491	0.02028	0.01720	0.01871
	NBL(0.9)	0.00367	0.01069	0.00060	0.00210	0.00499	0.02045	0.01714	0.01877
	BLasso	0.09161	0.07719	0.08545	0.04555	0.16737	0.08071	0.04697	0.08646
n=150	CLasso	0.12180	0.12760	0.12438	0.12657	0.12542	0.12498	0.11931	0.10921
	NBL(0.5)	0.04104	0.04767	0.04395	0.04628	0.04516	0.04471	0.03860	0.02750
	NBL(0.7)	0.00057	0.00767	0.00376	0.00630	0.00497	0.00462	0.00189	0.01377
	NBL(0.9)	0.00209	0.00749	0.00614	0.00626	0.00388	0.00355	0.00132	0.01516
	BLasso	0.04537	0.04999	0.08081	0.06555	0.05216	0.07465	0.07446	0.03981

حيث نلاحظ من هذا الجدول مدى أداء وجودة المقدرات التي تم احتسابها من الطريقة المقترحة والطرائق الأخرى ومن الجدول أيضا نلاحظ ان اقل قيم لمعيار التحيز كان للطريقة المقترحة وخصوصا عند تزايد حجم العينة وتزايد قيم معلمة التحيز.

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

الجدول ادناه (4) يوضح قيم معيار MSE والمعيار MAE’ والتي من خلاله يتبين أداء عمل طرائق تقدير المعالم.

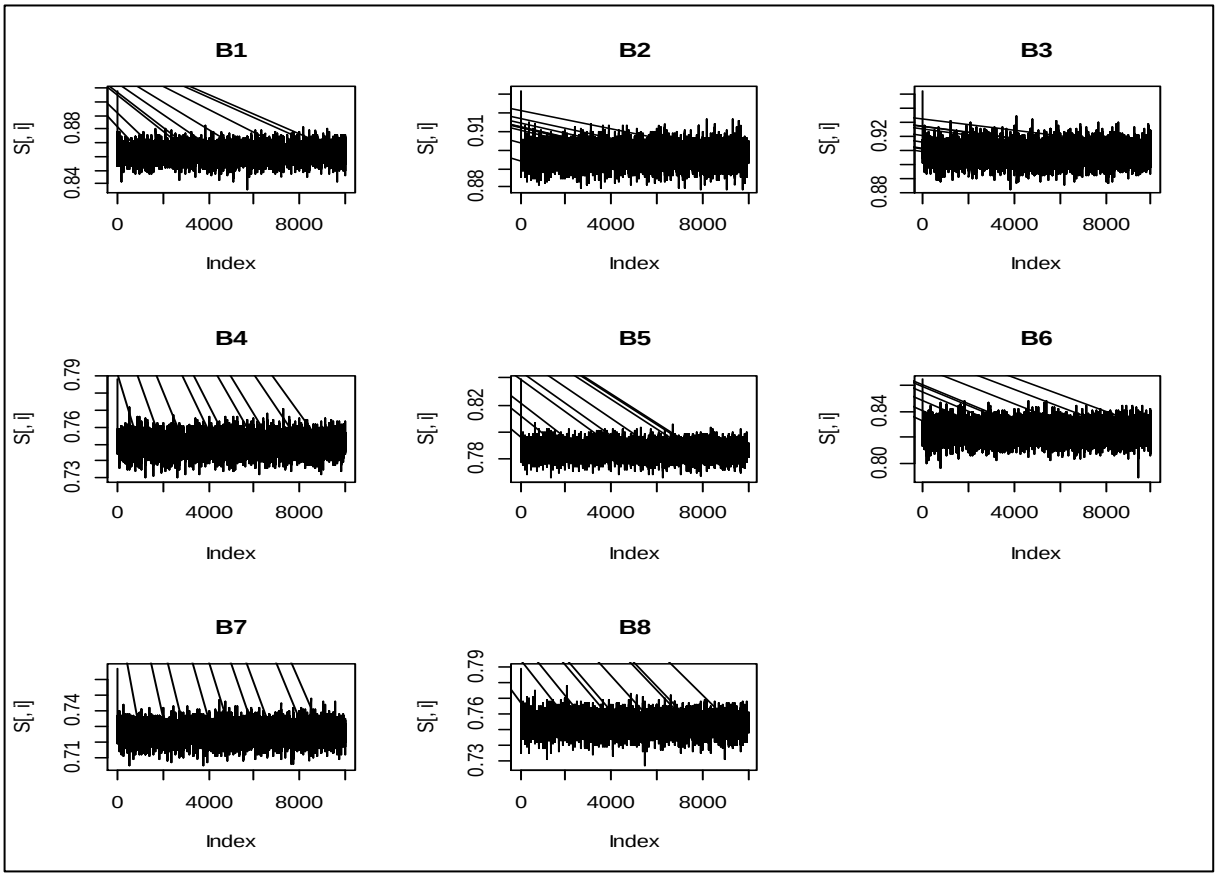
جدول (4) قيم معيار MSE و MAE لطرائق التقدير لتجربة المحاكاة الاولى

Samples	Methods	MSE	MAE
n=50	CLasso	1.53895	1.00542
	NBL(0.5)	1.21250	0.91832
	NBL(0.7)	1.17267	0.90493
	NBL(0.9)	1.17437	0.90535
	BLasso	1.29426	0.94406
n=100	CLasso	1.69070	1.02842
	NBL(0.5)	1.25177	0.91353
	NBL(0.7)	1.15284	0.88515
	NBL(0.9)	1.15288	0.88530
	BLasso	1.30800	0.89772
n=150	CLasso	1.38151	0.94452
	NBL(0.5)	1.00469	0.81215
	NBL(0.7)	0.94928	0.78672
	NBL(0.9)	0.94928	0.78696
	BLasso	1.05858	0.83168

من الجدول (4) نلاحظ ان معايير جودة أداء طرائق التقدير عند حجم العينة $(n = 50)$ كانت الأقل عدديا للطريقة المقترحة حيث تقل تدريجيا مع ازدياد قيم معلمة التعلم وكذلك الحال عند زيادة حجم العينة ليكون $n = 100$ و $n = 150$.

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

في ادناه تم رسم trace plot وهو بمثابة أداة لقياس تقارب سلسلة العينات التي تولدها خوارزمية Gibbs للمعاينة من قيم معالم نموذج الانحدار. أي ان هذه الأداة تمثل وسيلة لمعرفة مدى كفاءة خوارزمية Gibbs في توليد قيم المعالم.



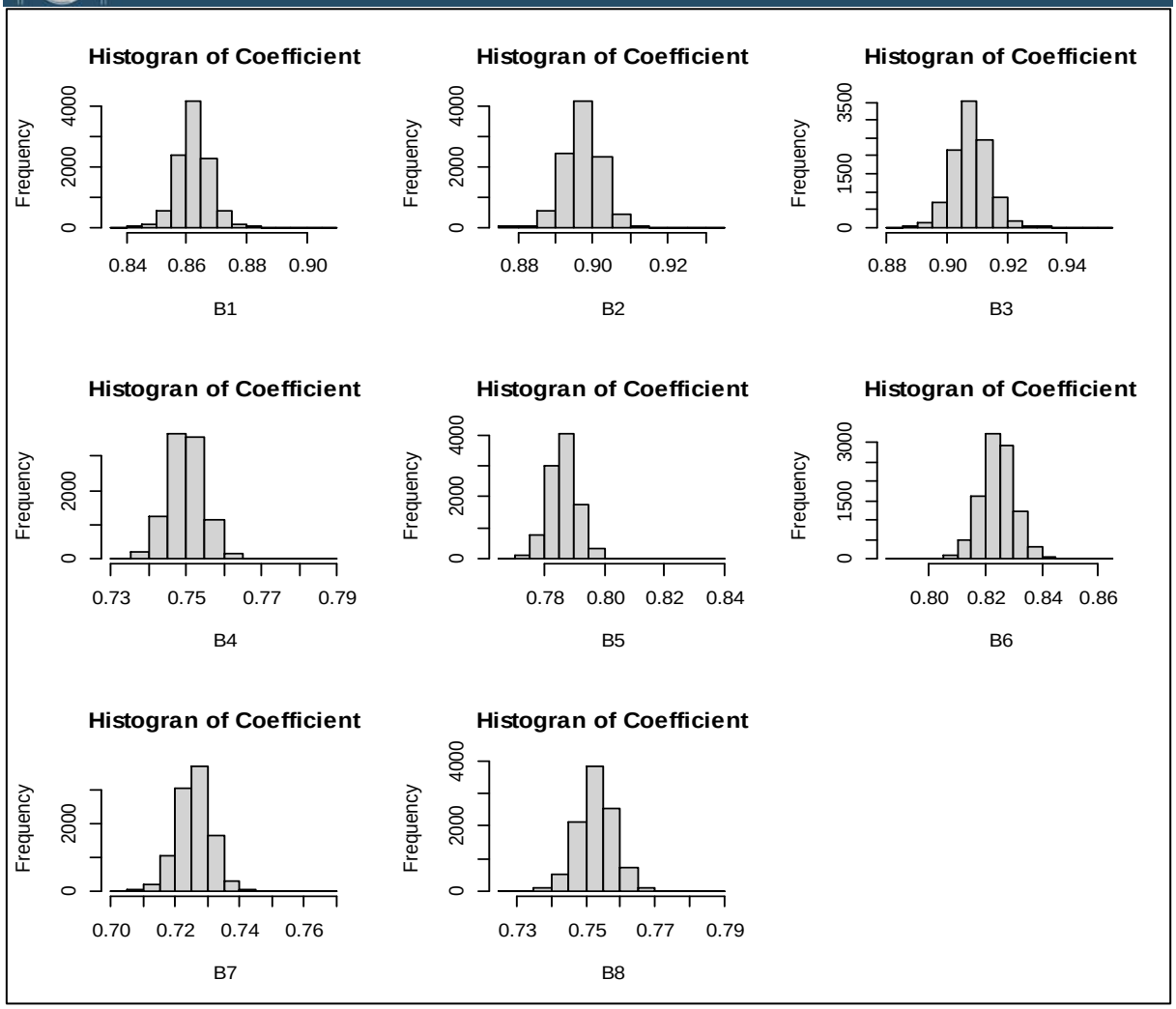
شكل (1) رسم trace plot للمعالم المقدرة وعند (α=0.5, n=50) لتجربة المحاكاة الاولى

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

ونلاحظ من الشكل أعلاه انه تم رسم (trace plot) لكل سلسلة متولدة من قيم معالم نموذج أي ان هناك (8) رسوم كل واحد يعود لمعلمة معينة حيث تبين من الرسومات ان أداء عمل الخوارزمية لا يعاني من تباطؤ في الاقتراب عند توليد قيم العينات في المعالم ولا تعاني من مشكلة flat bits (عدم وجود قيم متولدة بشكل رتيب و افقي) وان جميع مراكز السلاسل المتولدة قد كانت قريبة من قيم المتجه الحقيقي وبالنتيجة هذا يعطي انطباع على حسن اختيار كل من معلمة التعلم والتوزيع اللاحق .

كذلك قام الباحث برسم التوزيع التكراري Histogram لكل سلسلة متولدة من القيم المقدرة لمعالم نموذج الانحدار.

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) رسم المدرج التكراري لمعالم النموذج عند (α=0.5, n=50) لتجربة المحاكاة الاولى

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 نلاحظ من الشكل أعلاه ان جميع القيم المتولدة من التوزيع اللاحق لمعالم نموذج الانحدار الخطي المتعدد تتبع التوزيع الطبيعي وهذا ما يتطابق مع الجانب النظري الذي افترضنا فيه ان توزيع معلمة نموذج الانحدار β تتبع التوزيع الطبيعي.

1- الاستنتاجات

في هذا المبحث سيتم تناول اهم الاستنتاجات التي أدركها الباحث من خلال دراسته لهذه البحث بجانبها النظري وجانبها التجريبي وكانت اهم الاستنتاجات هي كما يأتي:

1. تم دراسة سلوك عمل أسلوب بيز الامن (Safe Bayesian) تحت طريقة اختبار المتغيرات لاسو Lasso حيث تم لأول مرة توظيف طريقة لاسو التي اقترحها الباحثان Yi و Mallick عام 2014 تحت أسلوب بيز الامن من خلال دراسة سلوك دالة الإمكان مع وجود معلمة التعلم (Learning rate) وبالتالي متوسط مقدرات التوزيع اللاحق لمعلمة نموذج الانحدار الخطي المتعدد .
2. تم اقتراح نموذج هرمي للتوزيعات المسبقة بافتراض وجود نموذج انحدار خطي متعدد و دالة امكان للبيانات مقيدة بمعلمة التعلم (Learning rate)
3. تم توظيف النموذج الهرمي المقترح للحصول على التوزيعات اللاحقة حيث تم استخدام خوارزمية Gibbs للمعاينة لتوليد العينات التي تمثل قيم متوسطات المعالم المقدرة من التوزيعات اللاحقة وفق طريقة لاسو ومن خلال أسلوب بيز الامن.
4. من خلال الجانب التجريبي تم اجراء تجربتي محاكاة من خلال افتراض وجود متجهات حقيقية لقيم معالم نموذج الانحدار الخطي المتعدد والمقترحة من قبل الباحث Tibshirani في عام 1996 وباعتبار هذه المتجهات على انها فضاء المعالم Parameter Space الحقيقي تم دراسة أداء الطريقة

المقترحة ودراسة سلوك حلولها لمعرفة مدى تطابق حلولها مع قيم المتجهات الحقيقية واظهرت النتائج افضلية الطريقة المقترحة اعتمادا على عدة معايير مثل (MAE ,MSE,Bias) وتحت حجوم عينات مختلفة .

5. أظهرت رسومات trace plot للعينات المتولدة من التوزيعات اللاحقة المقترحة . ان خوارزمية Gibbs للمعاينة تعمل بصورة جيدة ولا تعاني من أي مشاكل في إيجاد الحلول . إضافة الى ذلك تبين لنا من خلال رسوم Histograms ان هناك تطابق للنتائج المستحصل عليها من تجارب المحاكاة مع ما تم اقتراحه في الجانب النظري فيما يخص بالتوزيع الطبيعي لمعالم نموذج الانحدار الخطي المتعدد .

المصادر References

Friedman J, Hastie T, Tibshirani R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software. 33(1):1–22. [PubMed: 20808728].

Grunwald, P.D.(2012). The safe Bayesian: learning the learning rate via the mixability gap. In Proceedings 23rd International Conference on Algorithmic Learning Theory (ALT '12). Springer.

Leng, C., Tran, M., Nott, D. (2014). Bayesian adaptive lasso. Annals of the Institute of Mathematical Statistics. 66(2):221–244.

Li, Y. and Ghosh, SK. (2013). Technical report. North Carolina State University Department of Statistics. Efficient sampling methods for truncated multivariate normal and student-t distributions subject to linear inequality constraints.

Mallick, H., & Yi, N. (2014). A new Bayesian lasso. Statistics and its interface, 7(4), 571.

Park T, Casella G. The bayesian lasso. Journal of the American Statistical Association. 2008; 103:681– 686. 103.

Rianne, de Heide.(2016). the safe-bayesian lasso. Master thesis. mathematical institute.

Rianne, de Heide., Kirichenko,A., Mehta,N.A, and Grunwald, P.D. (2020). Safe-Bayesian Generalized Linear Regression. arXiv:1910.09227 **[math.ST]**. Cornell University.

Tibshirani R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological). 58:267–288.

Wu, P-S., and Martin,R. (2020). A comparison of learning rate selection methods in generalized Bayesian inference. arXiv:2012.11349. Cornell University.