إستخدام بعض طرق المربعات الصغرى الجزائية لتقدير واختيار متغيرات نموذج الإنحدار الخطى فى ظل وجود التعدد الخطى

د. خالد محمد محمد الغندور
مدرس بقسم الإحصاء والرياضة والتأمين
كلية التجارة – جامعة عين شمس

د. وائل سعد حسانين الدواخلى
مدرس بقسم الإحصاء والرياضة والتأمين
كلية التجارة – جامعة عين شمس

الملخص:

استخدمت هذه الدراسة النموذج الخطى المعمم من خلال طرق الإنكماش (Shrinkage) التى تُستخدم كبديل لطريقة المربعات الصغرى فى حالة حدوث مشكلة التعدد الخطى التى تظهر بشكل كبير فى حالة الفترات الزمنية القصيرة وكثرة المتغيرات المستقلة، ومع وجود هذه المشكلة يرتفع مقدار الخطأ المعيارى وتصبح المعلمات المقدرة غير معنوية، وبالتالى نحذف متغيرات هى فى الأصل متغيرات مؤثرة. ولذلك جاءت طرق الإنكماش كبديل فى التنبؤ والإستنتاج من خلال تخفيض قيمة الخطأ المعيارى وإظهار المتغيرات التى تؤثر حقيقة على المتغير التابع، ولذلك تم الإستعانة بطرق الإنكماش أو الطرق الجزائية التى تُستخدم فى تقدير واختيار متغيرات نموذج الإنحدار الخطى المتعدد آنيًا فى ظل وجود التعدد الخطى (Multicollinearity) والمتمثلة فى طرق إنحدار ريدج (Ridge) ولاسو (Lasso) والشبكة المرنة (Elastic-Net)، وقد أظهرت النتائج أن طريقتى إنحدار لاسو والشبكة المرنة هما الأفضل من بين طرق الإنكماش، وذلك من خلال دراسة

تأثير تكنولوجيا المعلومات والإتصالات على النمو الإقتصادى فى مصر خلال الفترة 2000-2020.

الكلمات المفتاحية:

التعـدد الخطـى، المربعـات الصـغرى الجزائيـة، الإنحـدار الجزائـى، طـرق الإنكمـاش، إنحـدار ريـدج، إنحـدار لاسـو، إنحـدار الشـبكة المرنـة، دالـة الجـزاء، تكنولوجيا المعلومات والإتصالات، النمو الإقتصادى.

Using some Penalized Least Squares Methods to Estimate and Select the Variables of Linear Regression Model in the Presence of Multicollinearity

Abstract

This study used the generalized linear model through the shrinkage methods that are used as an alternative to the least squares method in the event of multicollinearity problem which appears largely in case of short time periods and many independent variables. With this problem the standard error increases and the estimated parameters become insignificant, Thus we remove variables that are originally influential variables. Therefore shrinkage methods came as an alternative in prediction and conclusion by reducing the value of standard error and showing the variables that really affect the dependent variable. So the shrinkage or penalized methods that are used to estimate and select the variables of the multiple linear regression model were used simultaneously in the presence of multicollinearity, which is represented by ridge regression, Lasso regression, Elastic-Net regression. The results showed that Lasso regression and Elastic-Net regression are the best among the shrinkage methods. This is done by studying the impact of information and communication technology on economic growth in Egypt during the period 2000-2020.

Key Words: Multicollinearity, Penalized least squares, Penalized regression, Shrinkage methods, Ridge regression, Lasso regression, Elastic-Net regression, Penalty function, Information and communication technology, Economic growth.

المصطلحات

Elastic-Net	الشبكة المرنة
Least Absolute Shrinkage And Selection Operator	أقل إنكماش مطلق لإختيار العامل
Penalized likelihood	الإمكان الجزائى
Penalized Least Square	المربعات الصغرى الجزائية
Penalty Function	دالة الجزاء
Shrinkage	تقليص(إنكماش)

المقدمة:

غالباً ما تتضمن نماذج الإستجابة عدد كبير نسبياً من المتغيرات والتى عادة لا تكون ذات تأثير معنوى على ردود الفعل محل الإهتمام،ولذلك تُعتبر مشكلة اختيار متغيرات نموذج الإنحدار من المشاكل الإحصائية التى تحظى بالكثير من الإهتمام من قِبل الإحصائيين بشدة حتى الآن،وخاصة فى ظل تواجد عدد كبير من المتغيرات التفسيرية بالإضافة إلى مشكلة التعدد الخطى (Multicollinearity) بين هذه المتغيرات،وبالتالى يصعب التفرقة بين تأثير كل متغير من المتغيرات التفسيرية المرتبطة على متغير الاستجابة،وقد يحدث إسقاط لبعض المتغيرات التفسيرية الهامة،كما يُعد أسلوب اختيار الفئة الجزئية (Subset Selection)

أسلوب غير مستقر لأن تغيرات بسيطة فى البيانات يمكن أن تؤدى إلى تغيرات كبيرة فى النموذج المختار Zou (2006) و(1995, 1996) Breiman، ويكون ذلك غير ممكن حسابياً فى ظل وجود عدد كبير من المتغيرات التفسيرية.

كما أنه لا توجد طريقة واضحة لتحديد الأخطاء المعيارية المناسبة لغرض الإستدلال الإحصائى ونموذج التنبؤ، ويرجع ذلك إلى حقيقة تجاهل إجراءات اختيار النموذج إلى الطبيعة العشوائية للخطوات المختلفة لاختيار المتغيرات. كما توجد مشاكل أخرى بالنسبة إلى التجارب التى تتطلب قيدًا أو أكثر على التوزيع العشوائى مما يؤدى إلى ضرورة إستخدام تقدير المربعات الصغرى المعممة Generalized least squares (GLS) بالتوازى مع تقدير الإمكان الأكبر المقيد Restricted maximum likelihood estimation (REML) لمكونات التباين. وبالتالى فهناك حاجة لطريقة تختار المنحدرات التى يتم تضمينها فى النموذج وتقدر معاملات النموذج فى آن واحد، وهذا لن يُسهل عمل الباحثين فقط (وخاصة فى وجود الإستجابات المتعددة وفى حالة التوزيع العشوائى المقيد) ولكنه سيجعل من الممكن أيضا الحصول على أخطاء معيارية موثوق بها لمعالم النموذج المقدر، ولذلك سوف يتم إستخدام طريقة المربعات الصغرى الجزائية Penalized Least) (Squares لاختيار وتقدير نموذج متزامن وهى حالة خاصة من الإمكان الجزائى (Penalized likelihood). وتتمثل الفكرة الرئيسية للمربعات الصغرى الجزائية فى وجود جزاء لأى تقدير غير صفرى لمعاملات النموذج عند تخفيض مجموع مربعات البواقى، وبالتالى يوجد حافز لدى مدخل تقدير المربعات الصغرى الجزائية لتقدير معالم معينة للنموذج إلى الصفر Goos (2017)، أى أن هذه الطريقة تقوم بتقليص (Shrinkage) مقدرات بعض معالم الإنحدار وجعل البعض الآخر مساويًا للصفر، وبالتالى يمكن تقدير واختيار متغيرات نموذج الإنحدار الخطى آنيًا مع علاج مشكلة التعدد الخطى وذلك عن طريق المقدرات الجزائية وهى Ridge، Least Absolute Shrinkage and Selection Operator (Lasso)

Elastic-Net. وحيث أن قطاع تكنولوجيا المعلومات والإتصالات قد أصبح قاطرة النمو للاقتصاد الحديث أو ما يسمى بالإقتصاد الرقمى، كما أصبح معيارًا أساسيًا لقياس تقدم الدول حيث أنه يستحوذ على جانب كبير من الإستثمارات ويستوعب حجم كبير من الفنيين والعمالة الماهرة، فإنه سوف يتم إستخدام المقدرات الجزائية لتقدير النموذج القياسى للعلاقة بين تكنولوجيا المعلومات والإتصالات والنمو الإقتصادى فى مصر خلال الفترة 2000-2020، وبفضل البنية التحتية لتكنولوجيا المعلومات والإتصالات مثل الهواتف الثابتة والمحمولة والإنترنت والنطاق العريض فقد أصبح بإمكان الأفراد والشركات والحكومات الوصول إلى المعلومات والمعرفة بشكل أفضل من حيث الحجم والنطاق والسرعة، مما أدى إلى تحسين كفاءة تخصيص الموارد وخفض تكاليف الإنتاج بشكل كبير وتعزيز الطلب والإستثمار بشكل أكبر فى جميع القطاعات الاقتصادية Bahrini (2019)، وقد مهدت خدمات الإتصالات الطريق لمزيد من التقدم وانتشار التكنولوجيا التى غيرت المشهد الرقمى فى العديد من أنحاء العالم، ولا يزال قطاع الإتصالات قطاعًا متناميًا يخضع لإصلاحات كبيرة فى محاولة للتطور إلى مستوى يمكن من خلاله ربط الصناعات بفاعلية وتسريع عمليات الإنتاج بطريقة فعالة David (2019).

مشكلة البحث

أن دراسة نموذج الإنحدار فى حالة وجود عدد كبير من المتغيرات التفسيرية المؤثرة على المتغير التابع والتي غالبا ما يكون هناك ارتباط فيما بينها فتظهر مشكلة التعدد الخطى (Multicollinearity) بين المتغيرات، مما يؤدى إلى زيادة فى تباين المقدرات وتكون دقة التنبؤ غير مرضية، وقد تم إستخدام عدة طرق لتلافى مشكلة التعدد الخطى بين المتغيرات التفسيرية ولكنها لم تحل هذه المشكلة بصورة نهائية، كما أنه من الأمور الهامة فى أى دراسة تطبيقية تحديد المتغيرات الأكثر أهمية والتى يمكن تضمينها إلى النموذج، لذلك فإنه من الضرورى البحث عن طرق إحصائية مهمتها حل مشكلة التعدد الخطى واختيار عدد المتغيرات التفسيرية

الأهم، وبالتالى تقليص عدد معاملات الإنحدار والخروج بالنموذج الأفضل الذى تكون له قوة تفسيرية عالية، وذلك من خلال المقارنة بين بعض طرق تنظيم واختيار المتغيرات لتحديد الأفضل منها.

أى تتمثل مشكلة البحث فى اختيار أفضل طريقة جزائية يمكن إستخدامها لتقدير واختيار متغيرات نموذج الإنحدار الخطى آنيًا مع علاج لمشكلة التعدد الخطى وذلك بالتطبيق على قطاع تكنولوجيا المعلومات والإتصالات لبحث مدى تأثيره على النمو الإقتصادى فى مصر خلال الفترة 2000-2020، حيث يُعتبر هذا القطاع عاملاً هامًا فى تحديد قدرة الدولة على الإنتقال للإقتصاد العالمى الجديد المبنى على المعرفة والتكنولوجيا.

أهمية البحث

قبل ظهور الطرق الحديثة للإنحدار الجزائى ومنها إنحدار لاسو (Lasso) وإنحدار الشبكة المرنة (Elastic-Net) كانت الطريقة الأكثر إستخداما لاختيار المتغيرات التفسيرية التى يتم تضمينها للنموذج هى طريقة الاختيار المتدرج (Stepwise selection) والتى تعمل على تحسين دقة التنبؤ فى حالات معينة وخاصة عندما يكون لبعض المتغيرات التفسيرية علاقة قوية بمتغير الاستجابة، والذى يجعل التنبؤ غير دقيق، فضلاً عن طريقة إنحدار ريدج (Ridge regression) الأكثر إنتشاراً والتى تُستخدم لتحسين دقة التنبؤ للنموذج، حيث أنه يعمل على تحسين خطأ التنبؤ عن طريق تقليص معاملات الإنحدار الكبيرة من أجل تخفيض التكرار ولكنه لا يقوم بالاختيار المشترك وبالتالى لا يساعد على جعل النموذج أكثر قابلية للتفسير، بينما نجد فى المقابل أن طرق إنحدار لاسو والشبكة المرنة تستطيع تحقيق الهدفين عن طريق جعل مجموعة القيم المطلقة لمعاملات الإنحدار لها مقادير أقل من قيمة ثابتة، مما يُجبر بعض المعاملات لتكون مساوية للصفر، مع اختيار نموذج أبسط لا يتضمن تلك المعاملات، لذا كان لابد من دراسة

هذا الموضوع وتسليط الضوء عليه ومقارنته بالطرق شائعة الإستخدام مثل طريقة إنحدار ريدج ليتم التوصل إلى أكثر الطرق دقة فى تقدير نموذج الإنحدار المتعدد.

أهداف البحث

يهدف هذا البحث إلى الوصول لأفضل طريقة تقوم بتقدير واختيار مجموعة جزئية مختزلة من المتغيرات التفسيرية الأصلية وليس جميعها، حيث تقوم هذه المجموعة الجزئية من المتغيرات بتوضيح وتفسير معظم التباين الموجود فى البيانات بعد وضع بعض القيود الإضافية على عملية التقدير، ويتم عمل ذلك دون المساس بجوهر التحليل فيبقى مقدار التباين المفسر ضمن مداه المقبول، وتتمثل هذه الطرق الجزائية فى كل من Ridge، Lasso، Elastic-Net بالإضافة إلى طريقة المربعات الصغرى للمقارنة.

كما يهدف البحث أيضاً إلى علاج مشكلة التعدد الخطى من خلال نموذج قياسى لتوضيح تأثير قطاع تكنولوجيا المعلومات والإتصالات والمتغيرات المتعلقة به والتى تتمثل فى عدد مشتركى الهاتف المحمول، عدد مستخدمى الإنترنت، مؤشر الجاهزية الشبكية، الإنفاق على البحث العلمى والتطوير، الإستثمار فى البنية التحتية لقطاع الإتصالات وتكنولوجيا المعلومات، بالإضافة إلى مؤشر تطور تكنولوجيا المعلومات والإتصالات على النمو الإقتصادى فى مصر والذى يتمثل فى الناتج المحلى الاجمالى.

فروض البحث

1- يمكن صياغة الفرض الرئيسى لهذه الدراسة فى وجود علاقة طردية بين متغيرات قطاع تكنولوجيا المعلومات والإتصالات والمتمثلة فى عدد مشتركى الهاتف المحمول، عدد مستخدمى الإنترنت، مؤشر الجاهزية الشبكية، الإنفاق على

البحث العلمى والتطوير، الإستثمار فى البنية التحتية للقطاع، مؤشر تطور تكنولوجيا المعلومات والإتصالات والنمو الإقتصادى فى مصر.

2- وجود ارتباطات خطية بين المتغيرات المستقلة.

3- التقدير بطريقة المربعات الصغرى الإعتيادية يمكن إستخدامه كمرحلة أولية.

4- يختلف تقدير المربعات الصغرى الجزائية عن تقدير المربعات الصغرى الكلاسيكية فى أنها تعطى نماذج ذات متوسط مربعات أخطاء أقل، كما أنها تتضمن أقل عدد ممكن من المتغيرات وبالتالى تكون قابلة للتفسير بسهولة.

5- تعطى طرق الإنحدار الجزائى نتائج جيدة عند تطبيقها فى حالة البيانات التى تعانى من مشكلة التعدد الخطى.

حدود البحث

1- يغطى البحث الفترة الزمنية 2000-2020 وفقاً لما يتوافر من بيانات سنوية للمتغيرات، مع الإعتماد على مخرجات برنامج SPSS.

2- اعتمدت الدراسة على نشرة مؤشرات الإتصالات وتكنولوجيا المعلومات (سنوات مختلفة) والصادرة عن وزارة الإتصالات وتكنولوجيا المعلومات بالإضافة إلى قاعدة بيانات البنك الدولى والمنتدى الإقتصادى العالمى فى تجميع البيانات الخاصة بالدراسة.

الدراسات السابقة

يتمثل الهدف من عرض الدراسات السابقة التى تناولت موضوع تكنولوجيا المعلومات والإتصالات وتأثيرها على النمو الإقتصادى فى التعرف على المنهجية التى اتبعتها والمتغيرات التى درستها والنتائج التى توصلت إليها مما يساعد فى بلورة

الرؤيـة النظريـة لهـذه الدراسـة، وبالتـالى وضـع الـدعائم الأساسـية للنمـوذج المقتـرح لمعالجة مشكلة الدراسة.

1- دراسـة (مخزومـى، 2020): هـدفت إلـى دراسـة العلاقـة بـين البحـث والتطوير والنمو الإقتصادى لعدد (17) دولة صناعية متقدمة بإستخدام بيانات البانل للفترة 1996 - 2018، وأظهرت النتائج أن للمتغيرات التفسيرية (إجمالى تكوين رأس المال الثابت، وطلبات براءات الاختراع للمقيمين) فى الدول المختارة علاقة طردية ذات معنويـة إحصـائية تجـاه المتغيـر التـابع وهـو نصـيب الفـرد مـن النـاتج المحلـى الاجمـالى، بينمـا كـان لأعـداد البـاحثين فـى مجـال البحـث والتطـوير والإنفـاق الإستهلاكى النهائى للأسر تأثير عكسى على النمو الإقتصادى.

2- دراسـة (Bahrini, 2019): هـدفت هـذه الدراسـة إلـى تقيـيم تـأثير تكنولوجيا المعلومات والإتصالات على النمو الإقتصادى لمجموعة دول نامية مختارة مـن الشـرق الأوسـط ومنطقـة شـمال أفريقيـا (MENA) ومنطقـة جنـوب أفريقيـا بإستخدام الطريقة العامة للعزوم (GMM) لبيانات بانل خلال الفترة - 2016 2007، وتوصلت الدراسة لعدة نتائج منها تفوق دول منطقة الشرق الأوسط وشمال أفريقيا على دول منطقة جنوب أفريقيا فى مجالات إستخدام الإنترنت واعتماد النطاق العريض.

3- دراسـة (Fujino, 2018): حيث قام هو وآخرون بتطبيق إنحدار لاسو للتنبؤ بالمجال المرئى المستقبلى لمرض Glaucoma، وكذلك قاموا بتطبيق طريقة المربعات الصغرى العادية الخطية وإنحدار M الحصين وكذلك إنحدار MM الحصين، وتوصلوا إلى أن إنحدار لاسو أكثر دقة فى التنبة من بقية الطرق.

4- دراسة (Pradhan, 2018): هدفت هذه الدراسة إلى معرفة العلاقة السببية بين تكنولوجيا المعلومات والنمو الإقتصادى لدول مجموعة العشرين بإستخدام بيانـات البانـل خـلال الفترة 2001 - 2012 بإسـتخدام نمـوذج تصـحيح الخطـأ،

وتوصلت الدراسة إلى مجموعة من النتائج منها أن إستخدام تكنولوجيا المعلومات يعمل على زيادة النمو الإقتصادى فى دول عينة الدراسة.

5- دراسة (Aghaei, 2017): هدفت هذه الدراسة إلى قياس تأثير تكنولوجيا المعلومات والإتصالات على النمو الإقتصادى بإستخدام بيانات البانل لدول منظمة المؤتمر الإسلامى خلال الفترة 1990 - 2014 فى إطار نموذج النمو، وتوصلت الدراسة إلى وجود تأثير كبير للاستثمارات فى تكنولوجيا المعلومات والإتصالات على النمو الإقتصادى للدول محل الدراسة.

6- دراسة (Amiri, 2017): هدفت إلى قياس العلاقة بين التقدم التكنولوجى والنمو الإقتصادى واتجاهات العمالة عبر دول البريكس خلال الفترة 2007 - 2015، مع توضيح حجم المبادرات والحوافز الحكومية المختلفة والمتخذة من قبل تلك الدول لتعزيز التطور التكنولوجى بها، وتم إستخدام عدة مؤشرات منها: مؤشر تطور تكنولوجيا المعلومات والإتصالات والناتج المحلى الاجمالى ومعدل البطالة ومؤشر خاص بالاستعداد التكنولوجى لكل دولة، وتوصلت الدراسة إلى وجود علاقة إيجابية معنوية بين مؤشر تطور تكنولوجيا المعلومات والإتصالات والناتج المحلى الاجمالى وعلاقة سلبية معنوية بين معدل البطالة ومؤشر تطور تكنولوجيا المعلومات.

7- دراسة (Fonti, 2017): قام بشرح طريقة لاسو (Lasso) فى خاصية الاختيار وهى الميزة التى نختار بها عددًا أقل من المتغيرات التفسيرية لوصف متغير الاستجابة لجعل النموذج أسهل فى التفسير، وطبق طريقة Lasso على النماذج الخطية والنماذج الخطية المعممة عندما يكون عدد المتغيرات أكبر من عدد المشاهدات.

8- دراسة (Salahuddin, 2015): هدفت إلى بحث العلاقة بين إستخدام الإنترنت والنمو الإقتصادى والتنمية المالية والانفتاح التجارى لـ 11 دولة من جنوب أفريقيا خلال الفترة 1990 - 2012، وخلصت الدراسة إلى وجود علاقة إيجابية على المدى الطويل بين إستخدام الإنترنت وتحفيز النمو الإقتصادى فى المنطقة، كما أكدت الدراسة على أن إستخدام الإنترنت لا يؤثر على النمو الإقتصادى خلال فترة عينة الدراسة (1990 - 2012) فقط بل سيكون له تأثير متزايد على النمو الإقتصادى فى المنطقة فى المستقبل خلال الفترة - 2013) 2034).

9- دراسة (Tuna, 2015): ركزت هذه الدراسة على تحليل العلاقة بين نفقات البحث والتطوير والنمو الإقتصادى فى تركيا وذلك بإستخدام اختبارات جذر الوحدة واختبار التكامل المتناظر والسببية لجرانجر، وتوصلت النتائج إلى استقرار السلاسل الزمنية فى الدرجة الأولى وعدم وجود تكامل متناظر بينهم، بالإضافة إلى عدم وجود علاقة سببية بينهم.

ومن ثم تتمثل مساهمة الدراسة الحالية فى أنها تختبر العلاقة على مستوى الإقتصاد المصرى فقط حيث تندر الدراسات التطبيقية فى هذا المجال، وبإستخدام مؤشرات تعبر عن الجوانب المختلفة لإستخدام تكنولوجيا المعلومات والإتصالات والبنية الأساسية لها والبحث والتطوير فى أحدث فترة زمنية ممكنة متاحة للبيانات، حيث أنه مازال دور تكنولوجيا المعلومات والإتصالات على المستوى الكلى غير مكتشف بشكل واضح فى الأدبيات الاقتصادية ويتفاوت من دولة لأخرى. وذلك بالإضافة إلى إستخدام بعض طرق الاقتصاد القياسى الحديثة نسبيًا والمقارنة بينهم لاختيار أفضلهم.

متغيرات الدراسة

تم إستخدام عدة متغيرات لها علاقة كبيرة بموضوع الدراسة بالإعتماد على الدراسات السابقة لتأثير تكنولوجيا المعلومات والإتصالات على النمو الإقتصادى، وقد تم الإعتماد على لوغاريتمات هذه المتغيرات بهدف التقليص من حدة إختلاف التباين وبالتالى التخفيف من التقلبات سواء لنفس المتغير أو بين المتغيرات ، كما أنه يساعد فى تحويل بيانات السلسلة الزمنية إلى التوزيع الطبيعى إن لم تكن كذلك ، ويساعد أيضا فى تحويل التغيرات إلى النسب المئوية(المرونات)بدلاً من التغيرات بالوحدة . ويمكن تعريف المتغيرات المستخدمة فى الدراسة على النحو التالى:

- **لوغاريتم الناتج المحلى الاجمالى الحقيقى (Ln(GDP**: وهو يمثل المتغير التابع باعتباره مؤشرًا من مؤشرات النمو الإقتصادى ومقياسًا لحجم النشاط الإقتصادى. وقد اعتمدت الدراسة على بعض المؤشرات المعبرة عن الجوانب المختلفة لتكنولوجيا المعلومات والإتصالات مثل النفاذ (Access) والإستخدام (Use) وهما:
- **لوغاريتم عدد مشتركى الهاتف المحمول بالألف مشترك (X_1):** كمؤشر نفاذ.
- **لوغاريتم عدد مستخدمى الإنترنت بالألف مشترك (X_2):** كمؤشر إستخدام.
- **لوغاريتم الإستثمار فى قطاع تكنولوجيا المعلومات والإتصالات (X_3)**: تتوقع الدراسة وجود علاقة طردية بين هذا المتغير والنمو الإقتصادى، ويرجع الاهتمام بهذا المتغير إلى أنه يمثل أحد أهم القطاعات الواعدة فى الاقتصاد.
- **الإنفاق على البحث العلمى والتطوير كنسبة إلى الناتج المحلى الاجمالى (X_4):** إجمالى الإنفاق المحلى على البحث والتطوير معبرًا عنه كنسبة مئوية من اجمالى الناتج المحلى، حيث يغطى البحث والتطوير الأبحاث الأساسية والبحث التطبيقى والتطوير التجريبى، ويتوقع وجود علاقة طردية بينه وبين النمو الإقتصادى.

- **مؤشر الجاهزية الشبكية (X_5):** يُعتبر من بين أهم وأشمل المؤشرات المركبة المقترحة لقياس التحول إلى مجتمع المعلومات، حيث تم تقديره من قِبل المنتدى الإقتصادى العالمى ويصدر سنوياً فى تقرير مفصل يتضمن تحليل مقارن لكل الدول. ويُعرف هذا المؤشر بمدى استعداد البلد أو المجتمع للمشاركة والاستفادة من تطورات قطاع تكنولوجيا المعلومات والإتصالات.
- **مؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6):** يُعتبر أيضاً من بين المؤشرات المركبة المهمة جدًا فى قياس التحول نحو مجتمع المعلومات، ويستخدم بكثرة فى الدراسات المعنية بقياس التحول نحو مجتمع المعلومات، ويصدر من قِبل الاتحاد الدولى للاتصالات سنوياً بشكل تفصيلى مقارن لكل الدول، ويختلف عن مؤشر الجاهزية الشبكية فى منهجية الحساب وكذلك فى عدد المؤشرات الفردية المكونة له، ويُعتبر مؤشر تطور تكنولوجيا المعلومات والإتصالات هو الأشمل على الإطلاق.

الإطار القياسى المتبع فى التحليل

سوف يتم إستخدام بعض طرق المربعات الصغرى الجزائية Penalized Least Square) والتى تعالج مشكلة التعدد الخطى ومنها طرق إنحدار ريدج، إنحدار لاسو، إنحدار الشبكة المرنة، والتى تساعد الباحث أيضا فى أن يميز المتغيرات التى ينبغى أن يتضمنها النموذج لكى يصف متغير الاستجابة بصورة جيدة وكفوءة.

(1) إنحدار ريدج (Ridge Regression)

إن طريقة المربعات الصغرى الاعتيادية تعطى أفضل تقدير خطى غير متحيز BLUE وبأقل تباين لمعلمات نموذج الإنحدار، ومن المشكلات التى قد تظهر عند اتباع هذه الطريقة هى عند غياب أحد فروض التحليل ومنها عدم وجود

إرتباط خطى تام أو جزئى بين اثنين أو أكثر من المتغيرات التوضيحية والذى يؤدى إلى ظهور مشكلة التعدد الخطى التى تسبب إعطاء تقديرات غير دقيقة للمعلمات وذات تباينات كبيرة، ومن ثم الحصول على نتائج لاختبار الفروض لا يمكن الاعتماد عليها.

إن أول من أشار إلى خطورة تعدد العلاقة الخطية وتأثيرها فى نتائج تحليل الإنحدار هو العالم Fisher وكان ذلك عام 1934، وتبعه الكثير من الباحثين الذين وضعوا الجوانب المختلفة للمشكلة وطرائق حلها، حتى أضاف (Hoerl, 1970) مقداراً موجباً قيمته بين الصفر والواحد إلى عناصر قطر مصفوفة المعلومات X'X لحل هذه المشكلة، حيث أطلق على ذلك المقدار الثابت اسم معلمة التحيز (Biasing parameter)، وعلى الطريقة بإنحدار ريدج (Ridge regression) وتتميز طريقة إنحدار ريدج بإيجاد قيمة ثابتة k تُدعى بمعلمة التحيز، وهى كمية موجبة صغيرة تُضاف إلى عناصر قطر مصفوفة المعلومات X'X، وفائدة ذلك هو تقليل قيم عناصر قطر معكوس مصفوفة المعلومات الذى يؤدى إلى خفض قيم تباينات المعلمات المقدرة عند ابتعاد المتغيرات التفسيرية عن الإستقلالية أو عند ارتفاع قوة الارتباط بين أزواج المتغيرات التفسيرية، حيث أن إضافة الثابت k بقيم صغيرة تعمل على تغيير سريع فى قيم المعاملات المقدرة، ومع زيادة قيمة k تبدأ تلك القيم بالأستقرار تدريجياً إلى أن تصل إلى حد يكون التغير فيها طفيفاً وثابت الإشارة، وكلما كان إستقرار المعلمات سريعاً دل ذلك على أن المتغيرات التفسيرية قريبة من الاستقلالية (Ali, 2018).

أن مقدرات ريدج تعطى تقديرات للمعلمات بحيث تقلل مجموع مربعات الخطأ إلى جانب تحقيق القيد $\sum_{j=1}^{p} \beta j^2 \le s$ كالآتى:

$$\hat{\beta}=Min\sum_{j=1}^{n}(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij})^2 \qquad (1)$$

$$\text{subject to : } \sum_{j=1}^{p} \beta j^2 \le s$$

وباستعمال مضاعف لاجرانج (Lagrange Multiplier) يمكن أن نكتب المعادلة (1) كالآتى:

$$\hat{\beta}=Min\sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij})^2+k\,(\sum_{j=1}^{p}\beta_j^2-s) \qquad \textbf{(2)}$$

وحيث أن s ثابت ليس له أى تأثير على الحل وأن $k \geq 0$ هى معادلة الانكماش (Shrinkage)، فيمكن كتابة المعادلة (2) بالشكل الآتى:

$$\hat{\beta}=Min\sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij})^2+k\sum_{j=1}^{p}\beta_j^2 \qquad \textbf{(3)}$$

وتتكون المعادلة (3) من جزأين هما:

الأول : $\sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij})^2$ وهو الجزء الخاص بمجموع مربعات الأخطاء Sum Square Error (SSE).

الثانى : $k\sum_{j=1}^{p}\beta_j^2$ وهو يمثل دالة الجزاء (Penalty Function) أى أن:

$$\hat{\beta}=RSS+k\sum_{j=1}^{p}\beta_j^2 \qquad \textbf{(4)}$$

ويتم الحصول على القيم المقدرة لمعلمات النموذج بطريقة إنحدار ريدج عن طريق تدنية مجموعة مربعات الأخطاء الجزائية الآتية:

$$(Y-XB)'(Y-XB)+k\left\|\beta\right\|^2 \qquad \textbf{(5)}$$

وبإشتقاق المعادلة (5) بالنسبة إلى β ونساوى المشتقة بالصفر ينتج:

$$\begin{aligned}
&-2X'(Y-X\beta)+2k\beta=0\\
&-2X'Y+2X'X\beta+2k\beta=0\\
&-2X'Y+2\beta(X'X+kI_p)=0\\
&X'Y=\beta(X'X+kI_p)\\
&\hat{\beta}_R=(X'X+kI_p)^{-1}X'Y \qquad \textbf{(6)}
\end{aligned}$$

حيث أن:

$\hat{\beta}_R$: متجه المعلمات المقدرة بأسلوب إنحدار ريدج.

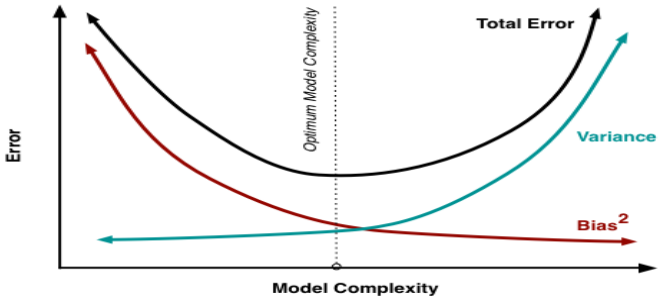
K : معلمة التحيز (معلمة ريدج).

وحيث أن مقدرات إنحدار ريدج هى تحويل خطى لمقدرات المربعات الصغرى، وأن قيمة متوسط مربعات الخطأ لمقدرات إنحدار ريدج هى:

$$MSE_R = \text{variance } (\hat{\beta}_R) + (\text{bias in } \hat{\beta}_R)^2 \qquad (7)$$

عند زيادة قيمة k فأن مقدار التحيز يزداد والتباين يقل، ولذلك يجب اختيار قيمة k بحيث يكون الانخفاض فى قيمة التباين أكبر من الارتفاع فى مقدار مربع التحيز، عند ذلك يكون متوسط مربعات الخطأ لإنحدار ريدج أقل من التباين لمقدرات المربعات الصغرى الإعتيادية، ويوضح ذلك شكل رقم (1).

Bias-Variance trade-off



شكل رقم (1): مقابلة التحيز بالتباين

المصدر: Arellano, C., (2021), "Lasso Regression", p (3)

كما أن زيادة قيمة k تؤدى إلى إنخفاض قيمة معامل التحديد R^2، ومن هنا يتضح بأن مقدرات إنحدار ريدج ليس من الضرورى أن تعطى أفضل نموذج ملائمة

للبيانات، إذ أننا نبحث عن أفضل معادلة ذات مقدرات ثابتة، أى تكون غير متحيزة عند زيادة قيمة k (Alkhamisi, 2007).

(2) إنحدار لاسو (Lasso Regression)

هى طريقة مقترحة من قبل الباحث (Tibshirani, 1996) وتم إجراء بعض التعديلات عليها من قِبل كل من (Lu, 2011)، (Tibshirani, 2013) ، (Zou, 2005). حيث أن Least Absolute Shrinkage and Selection Operator (Lasso) هى دالة جزاء لنموذج الإنحدار الخطى، وهى طريقة لتقدير معلمات نموذج الإنحدار وكذلك لاختيار وتنظيم المتغيرات الداخلة فى النموذج لزيادة الدقة التوضيحية لنماذج الإنحدار المستخدمة فى تحليل الظاهرة محل الدراسة من خلال عمليات ملائمة النموذج لاختيار مجموعة فرعية من المتغيرات المشتركة فى النموذج النهائى بدلاً من إستخدامها كلها، ففى طريقة لاسو يتم تصغير مجموع مربعات الأخطاء العشوائية لأعلى حد مع مجموع القيم المطلقة لمعاملات نموذج الإنحدار (Fonit, 2017) ، (Fujino, 2018) ، (Wang, 2007).

صممت Lasso أصلاً لنماذج المربعات الصغرى Least squares models، حيث أنها تكشف عن كمية كبيرة من سلوك المقدر عن طريق معامل لاسو أو ما يسمى Soft Thresholding بما فى ذلك علاقة مقدر لاسو مع مقدر إنحدار ريدج (Ridge Regression) ومقدر أفضل اختيار مجموعة جزئية من المتغيرات (Best subset selection) والتى تماثل طريقة الاختيار المتدرج (Stepwise selection)، ويكشف أيضا (كما فى الإنحدار الخطى) أنه لا يجب أن تكون تقديرات معامل لاسو وحيدة إذا كانت المتغيرات التفسيرية تعانى من مشكلة التعدد الخطى. وأن طريقة لاسو لها القدرة على اختيار مجموعة جزئية تعتمد على

صيغة القيد، وعلى الرغم من أنه تم تعريف لاسو للمربعات الصغرى إلا أنه يمكن بسهولة إستعمال طريقة لاسو فى مجموعة واسعة فى كثير من النماذج الإحصائية منها النماذج الخطية المعممة ومعاملات التقدير المعممة ونماذج المخاطر النسبية ومقدرات M (Mi Kim, 2018) ، (Li Fan, 2006) ، (Zhang, 2018). بينما يستطيع لاسو تحقيق كلا الهدفين عن طريق جعل مجموعة القيم المطلقة لمعاملات الإنحدار لها مقادير أقل من قيمة ثابتة، مما يُجبر بعض المعاملات لتكون مساوية للصفر، مع اختيار نموذج أبسط لا يتضمن تلك المعاملات (J.Fu Wenjiang, 1998)، (Tibshirani, 1997) ، (Dyar, 2012).

إن مبدأ طريقة إنحدار لاسو هو تصغير مجموع مربعات البواقى وفقاً إلى قيد يمثل المجموع المطلق للمعاملات والتى تكون أصغر من ثابت معين، إذ يضع لاسو قيداً على مجموع القيم المطلقة لمعاملات النموذج بحيث يجب أن يكون المجموع أقل من قيمة ثابتة (الحد الأعلى)، ومن أجل القيام بذلك تطبق لاسو عملية التقليص (تنظيم) إذ أنها تقوم بعمل جزاء لمعاملات الإنحدار وتقليص بعضها إلى الصفر، وأثناء عملية اختيار المتغيرات سيتم تحديد المتغيرات التى تكون معاملاتها غير صفرية بعد عملية التقليص (Shrinkage) وستكون جزءاً من النموذج والهدف من هذه العملية هو تقليل خطأ التنبؤ.

وتوجد فى طريقة لاسو معلمة ضبط (تسوية) هى التى تتحكم فى قوة معاقبة (جزاء) معاملات الإنحدار وتحتل أهمية كبيرة فى ذلك فعندما تكون معلمة الضبط كبيرة بشكل كافى تُجبر المعاملات إلى أن تكون مساوية للصفر، وتعتبر مقيدة فى تخفيض عدد المتغيرات فى النموذج، أى بمعنى كلما كانت قيمة معلمة الضبط كبيرة معناه عدد أكبر من المعاملات المساوية للصفر.

أما إذا كانت معلمة الضبط مساوية للصفر فسوف نحصل على إنحدار المربعات الصغرى الإعتيادية (Yi Congrui, 2016) ، (Tibshirani, 2013)، (Buhlmann, 2011).

وهناك العديد من المميزات فى إستخدام طريقة لاسو ومنها ما يلى:

- يمكن أن يوفر لاسو دقة تنبؤية جيدة للغاية لأن تقليص وإزالة المتغيرات يمكن أن يخفض التباين دون زيادة كبيرة فى التحيز، وهذا مفيد بشكل خاص عندما يكون لدينا عدد قليل من المشاهدات وعدد كبير من المتغيرات.
- يساعد لاسو على زيادة إمكانية تفسير النموذج من خلال القضاء على المتغيرات غير ذات الصلة التى لا ترتبط بمتغير الإستجابة.

وبذلك تعتبر طريقة لاسو طريقة لاختيار وتنظيم المتغيرات الداخلة فى النموذج الإنحدارى (Wang, 2007).

يتم تقدير معلمات إنحدار لاسو طبقاً لمبدأ المربعات الصغرى من الصيغة الأساسية كالاتى:

ليكن لدينا عينة مكونة من N حالات كل حالة مكونة من P من المتغيرات التفسيرية ومتغير تابع واحد y_i، وليكن X يمثل متجه المتغيرات التفسيرية للحالة j^{th}، فيكون هدف إنحدار لاسو هو حل المعادلة الآتية:

$$\min\left\{\frac{1}{N}\sum_{i=1}^{N}(y_i-\beta_0-x_i^T\beta)^2\right\} \qquad \textbf{(8)}$$

$$\text{Subject to } \sum_{j=1}^{p}\left|\beta_j\right|\le t$$

حيث أن:

t : تمثل معلمة تُحدد مسبقاً حيث تحدد مقدار التسوية (التقليص) (Li Fan, 2006)

وأن x_i^T هى الصف j^{th} من المصفوفة X

فيمكن كتابة صيغة لاسو بالشكل الأتى:

$$\min_{\beta_0,\beta}\left\{\frac{1}{N}\|y-X\beta\|_2^2\right\} \quad \textbf{(9)}$$

$$\text{Subject to } \|\beta\| \le t$$

حيث أن:

$\|\beta\|_p = (\sum_{i=1}^{N}|\beta_i|^p)^{1/p}$ وعندما يصبح 1 = P يصبح $\|\beta\|_1$ الطول القياسى ℓ^p وأن I_N متجه الوحدة (N x 1).

ويرمز $\bar{x}$ للمتوسط القياسى لنقاط البيانات x_i ، $\bar{y}$ متوسط متغير الإستجابة y_i والتقدير $\hat{\beta}_0 = \bar{y} - \bar{x}_i^T\beta$ حيث أن:

$$y_i-\beta_0-x_i^T\beta = y_i-(\bar{y}-\bar{x}_i^T\beta)-x_i^T\beta = (y_i-\bar{y})-(x_i-\bar{x})^T\beta$$

وبالتالى فإنه من الطبيعى العمل مع المتغيرات التى تم جعلها مركزية (جعل متوسطها يساوى صفر) إضافة إلى أن المتغيرات التوضيحية تكون معيارية مثالية Typically Standardizes بحيث أن

$$\frac{1}{N}\sum_{i=1}^{N}x_i = 0 \quad \text{و} \quad \frac{1}{N}\sum_{i=1}^{N}x_i^2 = 0$$

فيمكن إعادة كتابة الصيغة أعلاه بالشكل الآتى:

$$\min_{\beta_0,\beta}\left\{\frac{1}{N}\|y-X\beta\|_2^2\right\} \quad \textbf{(10)}$$

$$\text{Subject to } \|\beta\|_1 \le t$$

ويكون بصيغة مضاعف لاجرانج كالآتى:

$$\min_{\beta\in R^p}\{\frac{1}{N}\|y-x\beta\|_2^2+\lambda\|\beta\|_1\} \quad \textbf{(11)}$$

λ هى المعلمة التى تتحكم فى قوة الجزاء (الإنكماش) على معلمات الإنحدار (Yi Congrui, 2016)، (عبودى، 2017)، (Ranstam, 2018). وبالتالى فإن إنحدار لاسو يتميز عن إنحدار ريدج فى أنه يُحول القيم إلى صفر (جعلها صفر تماماً إذا كانت صغيرة بما فيه الكفاية) بدلاً من وضع قيم أقرب إلى الصفر وترك القيم الأكبر دون مساس كما يحدث فى إنحدار ريدج. أى أن تقديرات لاسو تجمع بين مميزات إنحدار ريدج وإنحدار أفضل اختيار جزئى والتى تحول جميع المعاملات إلى الصفر بقيمة ثابتة وتضبطها إلى الصفر إذا وصلت إليها (Malo, 2018).

(3) إنحدار Elastic Net

تعتبر طريقة إنحدار (Elastic Net) مشابهة لطريقة إنحدار لاسو ((Lasso)، حيث تُضيف إلى معامل التقليص $L_1 = \sum_{j=1}^{p} |\beta_j|$ معامل تقليص آخر $L_2 = \sum_{j=1}^{p} \beta_j^2$ المعرف كما يلى (Beran, 2015):

$$\hat{\beta}^{ridge} = \underset{\beta}{Argmin} \sum_{i=1}^{n} \left(Y_i - \beta_0 - \sum_{j=1}^{p} \beta_i x_{ij} \right) \quad / \quad \sum_{j=1}^{p} \beta_j^2 \le t \qquad \textbf{(12)}$$

وتكتب حسب دالة لاجرانج بالصيغة التالية:

$$\hat{\beta}^{ridge} = \underset{\beta}{Argmin} \left(\sum_{i=1}^{n} \left(Y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right) \qquad \textbf{(13)}$$

وعلى النقيض من طريقة إنحدار (Lasso) نجد أن طريقة إنحدار ريدج (Ridge) لا تختار أفضل المتغيرات المستقلة، بل هى طريقة إنكماش فقط وكلما كبرت قيمة (λ) يحدث إنكماش فى المتغيرات المستقلة إلى أن تصل قيمة المقدرات إلى الصفر. وعليه قام كل من (Zou Hui, 2005) باقتراح طريقة إنحدار (Elastic-Net) المعرفة كما يلى:

$$\hat{\beta}^{E.Net} = \underset{\beta}{Argmin}\left(\sum_{i=1}^{n}\left(Y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2+\lambda\sum_{j=1}^{p}((1-\alpha)\frac{1}{2}\beta_j^2+\alpha\left|\beta_j\right|\right) \quad (14)$$

أى أنه قام بمزج كل من طريقة إنحدار (Lasso) وطريقة إنحدار ريدج (Ridge)، من خلال معامل السيطرة (أو التقليص) $L_1=\sum_{j=1}^{p}\left|\beta_j\right|$ ومعامل السيطرة (أو التقليص) $L_2=\sum_{j=1}^{p}\beta_j^2$ ويتم الحصول على تقدير المعلمات حسب قيمة الضبط (λ)، ومراقبة أوزان كل من L_1 و L_2 حسب قيمة α ، فإذا كانت قيمة $(\alpha=0)$ نحصل على تقديرات إنحدار ريدج (Ridge)، وإذا كانت قيمة $(\alpha=1)$ نحصل على تقديرات إنحدار (Lasso)، أما إذا كانت $(0<\alpha<1)$ نحصل على تقديرات إنحدار (Elastic-Net).

الإحصاءات الوصفية لمتغيرات الدراسة:

يوضح جدول (1) أهم المقاييس الإحصائية الوصفية والمستخدمة فى تحليل متغيرات الدراسة.

جدول (1) : الإحصاءات الوصفية لمتغيرات الدراسة

X_6	X_5	X_4	X_3	X_2	X_1	Y	
3.4	3.56	0.46	9.53	9.44	10.39	14.02	الوسط الحسابى
3.44	3.62	0.43	9.77	9.74	11.16	14	الوسيط
1.4	3	0.2	7.88	6.23	7.28	12.74	الحد الأدنى
5.4	4.3	0.78	11.02	10.73	11.54	15.58	الحد الأعلى
1.28	0.37	0.21	0.9	1.24	1.38	0.92	الانحراف المعيارى
-0.01	0.32	0.14	-0.48	-1.27	-1.01	0.18	الالتواء
-1.5	-0.6	-1.79	-0.53	1.34	-0.33	-1.2	التفرطح
0.182	0.097	0.243	0.151	0.155	0.249	0.104	Kolmogorov-smirnov
0.067*	0.2*	0.002	0.2*	0.2*	0.001	0.2*	الاحتمال
0.922	0.959	0.851	0.945	0.871	0.797	0.944	Shapiro-wilk
0.097-	0.497*	0.004	0.272*	0.01	0.001	0.262*	الاحتمال
21	21	21	21	21	21	21	عدد المشاهدات

(*) الاحتمال أكبر من 5%: وبالتالى نقبل فرض العدم بأن البيانات تتبع التوزيع الطبيعى.

نلاحظ من جدول (1) أن القيمة الإحتمالية لإختبار التوزيع الطبيعى(Shapiro-Wilk)أكبر من ٠.٠٥ لمعظم المتغيرات،وبالتالى لا يمكن رفض الفرض العدم القائل بتبعية هذه المتغيرات للتوزيع الطبيعى فى مقابل الفرض البديل(عدم تبعيتها للتوزيع الطبيعى) ، ويوضح جدول (٢) مصفوفة الإرتباط بين المتغيرات التفسيرية .

جدول (2) : مصفوفة الإرتباط بين المتغيرات التفسيرية

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1					
X_2	0.953	1				
X_3	0.94	0.932	1			
X_4	0.835	0.845	0.869	1		
X_5	0.819	0.801	0.88	0.778	1	
X_6	0.901	0.9	0.913	0.973	0.857	1

حيث نلاحظ من جدول (2) أن بعض المتغيرات التفسيرية لها ارتباطات قوية أكثر من غيرها وهى: $X_1X_2 = 0.953$، $X_2X_3 = 0.932$ ، $X_4X_6 = 0.973$ ، مما يؤكد وجود مشكلة التعدد الخطى بين المتغيرات التفسيرية محل الدراسة.

أولاً: التقدير بطريقة المربعات الصغرى العادية (OLS)

يوضح جدول (3) تقدير نموذج الإنحدار الخطى بإستخدام طريقة المربعات الصغرى العادية.

جدول (3) : تقدير نموذج تكنولوجيا المعلومات والإتصالات بإستخدام المربعات الصغرى العادية

	المعاملات	الخطأ المعيارى	إحصاء t	الاحتمال
المقدار الثابت	10.023	0.544	18.42	0.0001*
عدد مشتركى الهاتف المحمول X_1	-0.152	0.074	-2.046	0.06
عد مستخدمى الإنترنت (X_2)	0.054	0.074	0.726	0.48
الإستثمار فى البنية التحتية X_3	0.146	0.115	1.264	0.2227
الإنفاق على البحث العلمى (X_4)	0.652	0.67	0.974	0.346
مؤشر الجاهزية الشبكية (X_5)	0.473	0.182	2.599	0.021*
مؤشر تطور تكنولوجيا (X_6)	0.495	0.152	3.255	0.006*

(*) معنوى عند مستوى 5%.

يتضح من جدول (3) أن المتغيرات الخاصة بكل من مؤشر الجاهزية الشبكية (X_5) ومؤشر تكنولوجيا المعلومات والإتصالات (X_6) كانت معنوية عند مستوى 5%. وباقى المتغيرات أتضح عدم معنويتها والتى تتمثل فى عدد مشتركى الهاتف المحمول (X_1) وعدد مستخدمى الإنترنت (X_2) والإستثمار فى البنية التحتية لقطاع الإتصالات (X_3) والإنفاق على البحث العلمى والتطوير (X_4)، وهذه النتيجة لا تتوافق مع المنطق ولا التوقعات حيث أنه يتوقع أن يكون لها تأثير معنوى على النمو الإقتصادى، وبالتالى نلاحظ أن طريقة المربعات الصغرى العادية أعطت تقديرات غير دقيقة لمعاملات الإنحدار، حيث أنها لم تمثل واقع الظاهرة المدروسة وأظهرت عدم أهمية بعض المتغيرات وعجز النموذج عن أظهار أهميتها نظراً لارتباط هذه المتغيرات بعضها ببعض، ولذلك سوف يتم اللجوء إلى الطرق الأخرى لتقدير النموذج. ولاختبار معنوية النموذج ككل تم إعداد جدول (4).

جدول (4) : جدول تحليل التباين لنموذج المربعات الصغرى

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	إحصاء F	الاحتمال
الإنحدار	16.881	6	2.813	223	0.0001*
البواقى	0.177	14	0.013		
الكلى	17.057	20			

(*) معنوى عند مستوى 5%.

ويتضح من الجدول (4) معنوية النموذج المقدر بطريقة المربعات الصغرى العادية عند مستوى معنوية 5%، وكانت قيمة معامل التحديد (R^2) للنموذج هى 0.99، ومتوسط مربعات الخطأ (MSE) هى 0.013، ويوضح جدول (5) القيم المختلفة لعامل تضخم التباين للمتغيرات بطريقة المربعات الصغرى.

جدول (5) : قيم عامل تضخم التباين (VIF) للمتغيرات بطريقة المربعات الصغرى

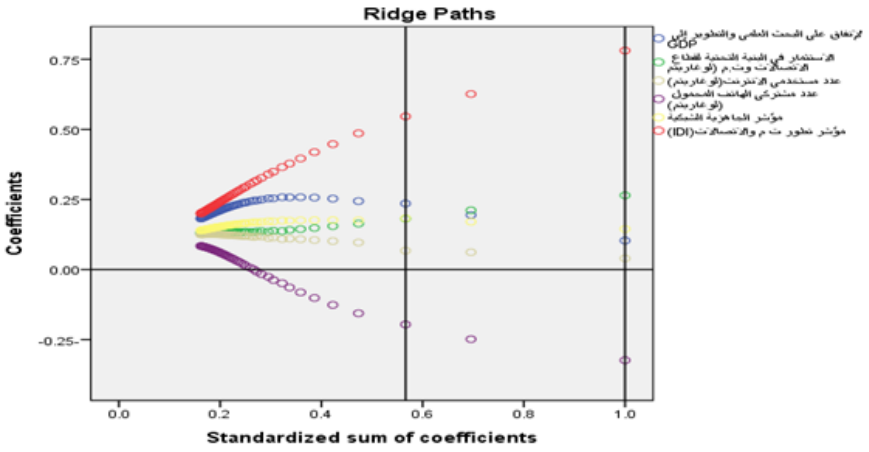
المتغير المستقل	Tolerance	تضخم التباين (Variance Inflation)
عدد مشتركى الهاتف المحمول (X_1)	0.06	16.74*
عدد مستخدمى الإنترننت (X_2)	0.074	13.53*
الإستثمار فى البنية التحتية لقطاع الإتصالات (X_3)	0.058	17.23*
الانفاق على البحث العلمى والتطوير (X_4)	0.03	33.11*
مؤشر الجاهزية الشبكية (X_5)	0.138	7.26
مؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6)	0.017	59.92*

(*) عامل تضخم التباين (VIF) أكبر من 10.

ويتضح من جدول (5) أن قيمة عامل تضخم التباين (VIF) لجميع المتغيرات فيما عدا مؤشر الجاهزية الشبكية (X_5) تزيد عن 10، وبالتالى فهى تشير جميعاً إلى وجود مشكلة التعدد الخطى بين هذه المتغيرات.

ثانياً: التقدير بطريقة إنحدار ريدج (Ridge Regression)

يستخدم أسلوب إنحدار ريدج عندما تعانى البيانات من التعدد الخطى حتى لو كان تقدير المربعات الصغرى (OLS) غير متحيز، فإن كبر حجم التباين يزيد الفجوة بين القيم التقديرية والقيم الحقيقية، وبالتالى فيمكن لإنحدار ريدج أن يقلل التباين بشكل فعال عن طريق إضافة درجة إضافية من الإنحراف إلى تقدير الإنحدار. وإشارة إلى نتائج تقدير طريقة المربعات الصغرى العادية (OLS) والتى تم اعتبارها غير واقعية، وبناءًا على وجود ارتباط بين المتغيرات التفسيرية، فإنه سوف يتم إستخدام طريقة إنحدار ريدج والذى يتضح من الشكل (2).



X-axis reference lines at optimal model and at most parsimonious model within 1 Std. Error.

شكل (2) : إحصاءات التقديرات النهائية لنموذج إنحدار ريدج الأمثل عند معلمة جزاء K = 0.04

ويوضح جدول (6) نتائج تقدير نموذج إنحدار ريدج

جدول (6) : تقدير نموذج تكنولوجيا المعلومات والإتصالات بإستخدام إنحدار ريدج عند معلمة جزاء K = 0.04

المتغيرات المستقلة	المعاملات	الخطأ المعيارى	Fإحصاء	الاحتمال
عدد مشتركى الهاتف المحمول (X_1)	-0.196	0.086	5.197	0.039*
عد مستخدمى الإنترنت (X_2)	0.067	0.065	1.084	0.315
الإستثمار فى البنية التحتية لقطاع الإتصالات (X_3)	0.182	0.083	4.737	0.047*
الإنفاق على البحث العلمى والتطوير (X_4)	0.236	0.06	15.237	0.002*
مؤشر الجاهزية الشبكية (X_5)	0.183	0.067	7.518	0.016*
مؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6)	0.546	0.196	32.665	0.0001*

(*) معنوى عند مستوى 5%.

نلاحظ من جدول (6) تحول جميع المتغيرات إلى المعنوية الإحصائية عند مستوى 5% فيما عدا المتغير الخاص بعدد مستخدمى الإنترنت (X_2)، وبالتالى يتضح وجود تحسن فى النتائج، ولاختبار معنوية النموذج ككل تم عمل جدول (7).

جدول (7) : جدول تحليل التباين لنموذج إنحدار ريدج

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	إحصاء F	الاحتمال
الإنحدار	20.644	6	3.441	135.217	0.0001*
البواقى	0.356	14	0.025		
الكلى	21	20			

(*) معنوى عند مستوى 5%.

ويتضح من الجدول (7) معنوية النموذج المقدر بطريقة إنحدار ريدج عند مستوى معنوية 5%، وكانت قيمة معامل التحديد (R^2) للنموذج هى 0.983 ومتوسط المربعات الخطأ (MSE) هى 0.025، ويوضح جدول (8) القيم المختلفة لعامل تضخم التباين للمتغيرات بطريقة إنحدار ريدج.

جدول (8) : قيم عامل تضخم التباين (VIF) للمتغيرات بطريقة إنحدار ريدج

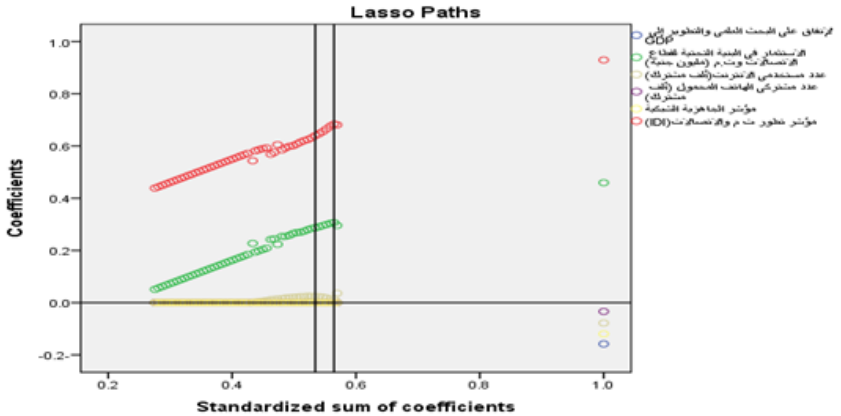
المتغير المستقل	Tolerance	تضخم التباين (Variance Inflation)
عدد مشتركى الهاتف المحمول (X_1)	0.121	8.26
عدد مستخدمى الإنترننت (X_2)	0.114	8.77
الإستثمار فى البنية التحتية لقطاع الإتصالات (X_3)	0.092	10.78*
الانفاق على البحث العلمى والتطوير (X_4)	0.1	10
مؤشر الجاهزية الشبكية (X_5)	0.282	3.55
مؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6)	0.061	16.39*

(*) عامل تضخم التباين (VIF) أكبر من 10.

يتضح من جدول (8) أن قيمة عامل تضخم التباين (VIF) للمتغيرين الإستثمار فى البنية التحتية لقطاع الإتصالات (X_3) ومؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6) أكبر من 10، وبالتالى فهى لازالت تعانى من مشكلة التعدد الخطى على الرغم من تحسن قيمتها عنه فى طريقة المربعات الصغرى العادية، وبالنسبة لبقية المتغيرات فقد تم علاج هذه المشكلة، الأمر الذى يُعنى تحسن فى نتائج إنحدار ريدج عنه فى المربعات الصغرى العادية بصفة عامة.

ثالثاً: التقدير بطريقة إنحدار لاسو (Lasso Regression)

أحد أنواع الإنحدار الخطى الذى يستخدم الإنكماش (Shrinkage)، حيث أن الانكماش هو ا لمكان الذى تقلصت فيه قيم البيانات باتجاه نقطة مركزية. ويناسب هذا النوع من الإنحدار النماذج التى تعرض مستويات عالية من التعدد الخطى أو عندما نريد اختيار أجزاء معينة فقط من النموذج، كما يقوم هذا النموذج بإضافة دالة جزاء (Penalty) تساوى القيمة المطلقة لحجم المعاملات، ومع هذا التنظيم يمكن أن تصبح بعض المعاملات صفرية ويتم إزالتها من النموذج للحصول على نماذج أبسط، ويوضح شكل (3) الاحصاءات النهائية لنموذج إنحدار لاسو الأفضل عند معلمة جزاء λ = **0.14**



X-axis reference lines at optimal model and at most parsimonious model within 1 Std. Error.

شكل (3): إحصاءات التقديرات النهائية لنموذج إنحدار لاسو الأمثل عند معلمة جزاء λ = 0.14

جدول (9) : تقدير نموذج تكنولوجيا المعلومات والإتصالات بإستخدام إنحدار لاسو (Lasso) عند معلمة جزاء λ = 0.14

المتغيرات المستقلة	المعاملات	الخطأ المعيارى	إحصاء F	الاحتمال
عدد مشتركى هاتف المحمول(X1)	0**	-	-	-
عدد مستخدمى الإنترنت (X_2)	0.024	0.061	0.074	0.789
استثمار قطاع الاتصالات (X_3)	0.287	0.082	12.254	0.003*
إنفاق البحث العلمى (X_4)	0**	-	-	-
مؤشر الجاهزية الشبكية (X_5)	0**	-	-	-
مؤشر تطور التكنولوجيا (X_6)	0.64	0.136	22.116	0.0001*

(*) معنوى عند مستوى 5%.

(**) متغيرات تم تقليصها.

ونلاحظ من الجدول (9) توافر المعنوية الاحصائية عند مستوى معنوية 5% للمتغيرين الإستثمار فى البنية التحتية لقطاع الإتصالات (X_3) ومؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6)، كما يتضح عدم معنوية المتغير الخاص بعدد مستخدمى الإنترنت (X_2)، بالإضافة إلى أن المتغيرات عدد مشتركى الهاتف المحمول (X_1) والانفاق على البحث العلمى والتطوير (X_4) ومؤشر الجاهزية الشبكية (X_5) قد قُلصت من النموذج (أصبحت معاملات إنحدارها تساوى صفر). ويوضح جدول (10) اختبار معنوية النموذج المقدر بطريقة إنحدار لاسو.

جدول (10): جدول تحليل التباين للنموذج المقدر بطريقة إنحدار لاسو (Lasso)

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	إحصاء F	الاحتمال
الإنحدار	20.706	3	6.902	398.6	0.0001*
البواقى	0.294	17	0.017		
الكلى	21	20			

(*) معنوى عند مستوى 5%.

يتضح من الجدول (10) معنوية النموذج المقدر بطريقة إنحدار لاسو (Lasso) عند مستوى معنوية 5%، وكانت قيمة معامل التحديد للنموذج هى 0.986، ومتوسط مربعات الخطأ (MSE) هى 0.017، ويوضح جدول (11) القيم المختلفة لعامل تضخم التباين للمتغيرات بطريقة لاسو.

جدول (11) : قيم عامل تضخم التباين (VIF) للمتغيرات بطريقة إنحدار لاسو

المتغير المستقل	Tolerance	تضخم التباين (Variance Inflation)
عدد مشتركى الهاتف المحمول (X_1)	0.059	16.95*
عدد مستخدمى الإنترننت (X_2)	0.553	1.81
الإستثمار فى البنية التحتية لقطاع الإتصالات (X_3)	0.12	8.33
الانفاق على البحث العلمى والتطوير (X_4)	0.06	16.67*
مؤشر الجاهزية الشبكية (X_5)	0.196	5.1
مؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6)	0.338	2.96

(*) عامل تضخم التباين (VIF) أكبر من 10.

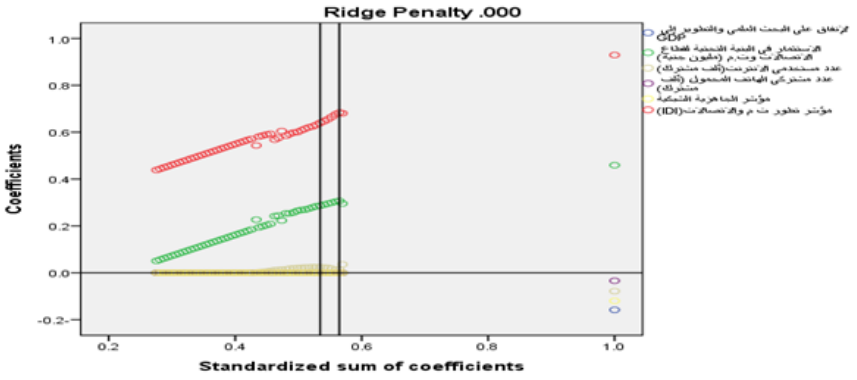
يتضح من جدول (11) أن قيمة عامل تضخم التباين (VIF) للمتغيرات الثلاثة المتضمنة فى نموذج لاسو وهى عدد مستخدمى الإنترنت (X_2) والإستثمار فى البنية التحتية لقطاع الإتصالات (X_3) ومؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6) أقل من 10، مما يؤكد على عدم وجود مشكلة التعدد الخطى فى

نموذج لاسو، وبالنسبة لباقى المتغيرات فقد تم تقليصها فى إنحدار لاسو، حيث نلاحظ أن طريقة إنحدار لاسو قد قلصت المتغيرات المستقلة الداخلة فى النموذج إلى ثلاثة متغيرات فقط واستبعدت الثلاثة متغيرات الباقية، حيث أجبرت معاملاتها أن تكون مساوية للصفر.

رابعًا: التقدير بطريقة إنحدار الشبكة المرنة (Elastic Net)

إنحدار الشبكة المرنة (Elastic Net) هو تقنية مختلطة من إنحدار ريدج وإنحدار لاسو، وبالتالى فهو يستخدم عاملين للإنكماش مما يجعله مناسب أكثر فى حالة المتغيرات عالية الارتباط، حيث أنه يدعم التأثيرات الجماعية.

وعند تقدير النموذج بطريقة الشبكة المرنة كان أفضل نموذج تم تقديره عند معلمة جزاء لإنحدار ريدج (λ_1) = صفر ومعلمة جزاء لإنحدار لاسو (λ_2) = 0.14، وبالتالى كانت النتائج مطابقة تماماً للنموذج المقدر بطريقة إنحدار لاسو، ويوضح شكل (4) الاحصاءات النهائية لنموذج إنحدار الشبكة المرنة الأفضل عند معلمة جزاء لإنحدار ريدج K = 0، ومعلمة جزاء لإنحدار لاسو λ 0.14 =



شكل (4): إحصاءات التقديرات النهائية لنموذج إنحدار الشبكة المرنة عند معلمة جزاء ريدج K = 0 ومعلمة جزاء لاسو λ = 0.14

مقارنة بين طريقة المربعات الصغرى العادية وطريقة إنحدار ريدج وإنحدار لاسو

جدول (12) : الأخطاء المعيارية وعامل تضخم التباين لطرق إنحدار OLS ، Ridge، Lasso

	المربعات الصغرى العادية		إنحدار ريدج		إنحدار لاسو	
	الخطأ المعيارى	VIF	الخطأ المعيارى	VIF	الخطأ المعيارى	VIF
X_1	0.074	16.74	0.086	8.26	-	-
X_2	0.074	13.53	0.065	8.77	0.061	1.81
X_3	0.115	17.23	0.083	10.87	0.082	8.33
X_4	0.67	33.11	0.06	10	-	-
X_5	0.182	7.26	0.067	3.55	-	-
X_6	0.152	59.91	0.196	16.39	0.136	2.96

يتضح من الجدول (12) أن الأخطاء المعيارية لطريقة لاسو كانت أقل من الأخطاء المعيارية للطرق الأخرى، ويرجع ذلك إلى أنها عالجت مشكلة التعدد الخطى والتى تُسبب تضخم الأخطاء المعيارية للمقدرات، ويؤكد ذلك أيضاً أن قيمة عامل تضخم التباين (VIF) لطريقة لاسو كانت هى الأقل أيضا بالنسبة للطرق الأخرى، ويمثل الجدول (13) مقارنة بين الطرق الثلاث بالاعتماد على معيار متوسط مربعات الخطأ ومعيار معامل التحديد والقيمة الاحتمالية المقابلة لاختبار F لكل نموذج.

جدول (13) : متوسط مربعات الخطأ ومعامل التحديد والقيمة الإحتمالية للطرق الثلاث

	متوسط مربعات الخطأ (MSE)	معامل التحديد (R^2)	القيمة الاحتمالية (P-value)
المربعات الصغرى العادية	0.013	0.990	0.0001
إنحدار ريدج	0.025	0.983	0.0001
إنحدار لاسو	0.017	0.986	0.0001

يتضح من الجدول (13) أن نموذج الإنحدار الخطى المتعدد اعتماداً على طريقة المربعات الصغرى العادية أعطى أكبر نسبة تفسير وهى 99% وكذلك أقل نسبة فى متوسط مربعات الخطأ 0.013، إلا أنها تضمنت متغيرات تفسيرية لها مشكلة التعدد الخطى وبدرجة عالية، وبالتالى تضعف القدرة التنبؤية للنموذج ولا يمكن الاعتماد عليه فى التنبؤ. أما عند المقارنة بين نموذجى إنحدار ريدج ولاسو فيتضح لنا أن طريقة لاسو هى الأفضل حيث أنها تتضمن معامل تحديد أكبر وهو 98.6% وكذلك متوسط مربعات خطأ أقل وهو 0.017، ويؤكد ذلك على أن طريقة إنحدار لاسو أفضل من طريقة إنحدار ريدج وكذلك المربعات الصغرى العادية فى تقدير نموذج الإنحدار الخطى المتعدد وفى ظل البيانات التى تعانى من مشكلة التعدد الخطى.

وتكون معادلة الإنحدار المقدرة بطريقة لاسو كالتالى:

$$\hat{y}=0.024X_2+0.287X_3+0.64X_6$$

حيث يتضح وجود تأثير طردى غير معنوى للمتغير الخاص بعدد مستخدمى الإنترنت (X_2) على النمو الإقتصادى، حيث تشير قيمة المعلمة المقدرة إلى أن زيادة هذا المتغير بنسبة 10% تؤدى إلى ارتفاع النمو الإقتصادى والممثل

فى الناتج المحلى الاجمالى بنسبة 0.24%، ويتضح أيضاً وجود تأثير طردى معنوى للإستثمار فى البنية التحتية لقطاع الإتصالات (X_3) على النمو الإقتصادى، حيث أشارت قيمة المعلمة المقدرة إلى أن زيادة قيمة هذا المتغير بنسبة 10% تؤدى إلى ارتفاع النمو الإقتصادى بنسبة 2.87%، كذلك الأمر بالنسبة إلى مؤشر تطور تكنولوجيا المعلومات والإتصالات (X_6) كان له أثر إيجابى ويعتبر أكثر المتغيرات تأثيراً على النمو، حيث تشير قيمة المعلمة المقدرة إلى أن زيادة هذا المؤشر بنسبة 10% يترتب عليها زيادة فى النمو الإقتصادى بنسبة 6.4%، كما وصلت القدرة التفسيرية للنموذج إلى 98.6%، أى أن حوالى 98.6% من التغيرات فى النمو الإقتصادى يمكن تفسيرها من خلال المتغيرات المدرجة فى النموذج. ويوضح جدول (14) نتائج إستخدام نموذج لاسو الأفضل ($\lambda=0.14$) فى التنبؤ بالإضافة إلى قيم الخطأ خلال السنوات الخمس الأخيرة.

جدول (14) : القيم التنبؤية لنموذج لاسو ($\lambda=0.14$)

السنة	القيمة الفعلية (Y)	القيم التنبؤية ($\hat{y}$)	البواقى
2016	14.8122	14.7352	0.0771
2017	15.0597	14.9153	0.1443
2018	15.3056	14.4874	-0.1817
2019	15.4874	15.5738	-0.0865
2020	15.5769	15.5013	0.0757

النتائج:

1- اتضح من مصفوفة الارتباط بين المتغيرات التفسيرية وجود ارتباط قوى بين هذه المتغيرات ومنها $X_1X_2 = 0.953$، $X_2X_3 = 0.932$، $X_4X_6= 0.973$، مما يشير إلى وجود مشكلة التعدد الخطى بينها.

2- كان عامل تضخم التباين (VIF) للمتغيرات بطريقة المربعات الصغرى العادية لجميع المتغيرات التفسيرية (فيما عدا مؤشر الجاهزية الشبكية X_5) أكبر من 10، ويؤكد ذلك على وجود التعدد الخطى فيما بينها.

3- إن نموذج الإنحدار الخطى المتعدد اعتماداً على طريقة المربعات الصغرى العادية قد أظهر معنوية احصائية لمتغيرين فقط من ستة متغيرات وهما مؤشر الجاهزية الشبكية (X_5) ومؤشر تكنولوجيا المعلومات والإتصالات (X_6) عند مستوى معنوية 5%، كما أعطى نسبة تفسير 99% ومتوسط مربعات الخطأ (MSE) يساوى 0.013. وبالتالى فإن عدم معنوية أربعة متغيرات فى النموذج رغم أهميتها الاقتصادية المتوقعة دليل على عجز النموذج والذى يرجع إلى تواجد مشكلة التعدد الخطى والتى تؤثر على معنوية بعض المتغيرات التفسيرية وتُظهر عدم معنويتها رغم أهميتها.

4- كان عامل تضخم التباين (VIF) للمتغيرات بطريقة إنحدار ريدج عند قيمة K = 0.04 للمتغيرات X_1 ، X_2 ، X_4، X_5 أقل من 10 مما يشير إلى عدم وجود تعدد خطى فيما بينها، وبالنسبة للمتغيرين X_3 ، X_6 فقد كانت قيمة عامل تضخم التباين (VIF) لهما هى 10.87 ، 16.39 على الترتيب، ويلاحظ انخفاض القيمتان عن نظيرتهما فى طريقة المربعات الصغرى 17.23 ، 59.92. مما يُعنى وجود تحسن كبير فى نتائج إنحدار ريدج عنه فى طريقة المربعات الصغرى العادية.

5- إن نموذج الإنحدار المتعدد اعتماداً على طريقة إنحدار ريدج قد أظهر معنوية جميع المتغيرات التفسيرية عند مستوى معنوية 5% فيما عدا المتغير الخاص بعدد مستخدمى الإنترنت (X_2)، كما أعطى نسبة تفسير 98.3% ومتوسط مربعات خطأ (MSE) يساوى 0.025، ورغم أن الفروق بينها وبين طريقة

المربعات الصغرى العادية بسيطة إلا أنه يُعتبر قد حدث تحسن كبير فى النتائج، والذى يرجع إلى علاج مشكلة التعدد الخطى فى نموذج إنحدار لاسو إلى حد كبير.

6- كان عامل تضخم التباين (VIF) لجميع المتغيرات المتضمنة فى النموذج بطريقة إنحدار لاسو (Lasso) وهى X_2 ، X_3 ، X_6 أقل من 10، مما يشير إلى أنه لا توجد مشكلة تعدد خطى فى النموذج، وبالتالى حدوث تحسن إضافى فى نتائج إنحدار لاسو عن الطرق السابقة.

7- إن نموذج الإنحدار المتعدد اعتمادًا على طريقة إنحدار لاسو قد أظهر معنوية جميع المتغيرات التفسيرية عند مستوى معنوية 5% فيما عدا المتغير الخاص بعدد مستخدمى الإنترنت (X_2)، كما أعطى نسبة تفسير 98.6% ومتوسط مربعات خطأ (MSE) يساوى 0.017.

8- بمقارنة طريقتى إنحدار ريدج وإنحدار لاسو يتضح أفضلية طريقة إنحدار لاسو حيث أنها أكبر فى قيمة معامل التفسير 98.6% وأقل فى قيمة متوسط مربعات الخطأ (MSE) 0.017، وبالتالى فإن طريقة إنحدار لاسو تعتبر أفضل من طريقة إنحدار ريدج فى معالجة مشكلة التعدد الخطى.

9- حيث أنه عند تقدير نموذج الإنحدار المتعدد إعتمادًا على طريقة إنحدار الشبكة المرنة (Elastic Net) والتى تدمج بين إنحدار ريدج وإنحدار لاسو كان النموذج الأفضل عند معلمة جزاء لإنحدار ريدج (λ_1) = 0 ومعلمة جزاء لإنحدار لاسو (λ_2) = 0.14، فكان النموذج الناتج مطابق تماماً للنموذج المقدر بإستخدام طريقة إنحدار لاسو.

10- يؤدى إستخدام طرق إنحدار ريدج وإنحدار لاسو وإنحدار الشبكة المرنة إلى تحسين فى معنوية معاملات الإنحدار المتعدد مقارنة بطريقة المربعات الصغرى العادية، وبالتالى تصبح المعاملات أكثر منطقية.

11- ليس بالضرورة أن يؤدى إستخدام طرق إنحدار ريدج وإنحدار لاسو وإنحدار الشبكة المرنة إلى زيادة فى قيمة معامل التفسير (R^2) أو تخفيض فى قيمة متوسط مربعات الخطأ (MSE) مقارنة بطريقة المربعات الصغرى العادية.

12- قلصت طريقتى إنحدار لاسو وإنحدار الشبكة المرنة المتغيرات المستقلة الأقل أهمية سواء فى نموذج المربعات الصغرى العادية أو نموذج إنحدار ريدج وهى: عدد مشتركى الهاتف المحمول (X_1) والإنفاق على البحث العلمى والتطوير (X_4) ومؤشر الجاهزية الشبكية (X_5).

13- فعالية طرق إنحدار لاسو وإنحدار الشبكة المرنة والتى جعلت النموذج أكثر قابلية للتفسير واختيار المتغيرات أكثر سهولة.

14- رغم أن مقدرات إنحدار لاسو وإنحدار الشبكة المرنة تزيد التحيز إلا أنها أفضل من ناحية المعنوية الاحصائية والاختبار الإقتصادى مقارنة بطريقة المربعات الصغرى العادية فى حالة وجود التعدد الخطى بين المتغيرات التفسيرية.

التوصيات :

1- ضرورة التحقق من توافر الفروض الخاصة بنظرية Gauss-Markov ومنها الفرض الخاص بعدم وجود ارتباط شبه تام بين المتغيرات التفسيرية بإستخدام الاختبارات الاحصائية للكشف عن وجود مشكلة التعدد الخطى.

2- النظرة المتعمقة للنتائج بطريقة المربعات الصغرى العادية وليس مجرد التعليق على النتائج الحسابية حتى ولو كانت غير متوافقة مع المنطق العلمى والواقعى.

3- تقدير معلمات نموذج الإنحدار الخطى المتعدد بإستخدام طرق التقليص المذكورة وهى إنحدار لاسو وإنحدار الشبكة المرنة كبديل لطريقة المربعات الصغرى العادية كلما اتجهت أحجام العينات إلى الصغر، حيث أن هذه الطرق تعطى نتائج أفضل فى التعامل مع زيادة حدة التعدد الخطى فى العينات الصغيرة.

4- إستخدام طرق إنحدار لاسو أو إنحدار الشبكة المرنة لاختيار المتغيرات لانها قادرة على تقليص (Shrinkage) تقدير المعلمات إلى الصفر بفرض قيد على مجموع مربعات الإنحدار، وبالتالى فإنها أفضل فى اختيار المتغيرات من طريقة إنحدار ريدج.

5- عدم الاعتماد على طريقة المربعات الصغرى العادية وطريقة إنحدار ريدج وذلك لعدم معالجتها لمشكلة التعدد الخطى بالكامل وخاصة عندما يكون عدد المتغيرات المستقلة كبير.

6- التوسع فى دراسة طريقة إنحدار لاسو وكذلك إنحدار الشبكة المرنة كونها من الطرق الحديثة فى تقدير النماذج الاحصائية والتى تلعب دوراً فعالاً فى تحديد المتغيرات المستقلة الأكثر أهمية فى النموذج.

7- التوسع فى إستخدام طرق إنحدار جزائى أخرى لتقدير واختيار متغيرات نموذج الإنحدار الخطى المتعدد آنياً فى حالة وجود التعدد الخطى ومنها Smoothly Clipped Absolute Deviation (SCAD) وكذلك طريق لاسو التكييفية (Adaptive Lasso).

المراجـع

أولاً: المراجـع العربيـة:

(1) عبودى، عماد حازم، على، حميد يوسف، (2017): مقارنة مقدرى Hurber Lasso و Hurber Elastic Net بإستخدام المحاكاة، مجلة الكويت للعلوم الاقتصادية والإدارية، العدد (28)، الجزء الأول.

(2) مخزومى، لطفى (2020): أثر البحث والتطوير على النمو الإقتصادى – دراسة قياسية لعينة من الدول الصناعية 1996-2018، المجلة العلمية المستقبل الإقتصادى – كلية العلوم الاقتصادية والتجارية وعلوم التسيير، الجزائر.

ثانياً: المراجع الأجنبية:

(1) Ali, Sadig Mohammed Bager, (2018): Ridge Regression for Addressing of the Multicollinearity Problem with Application in Cost of Production, 3rd Central & Eastern European LUMEN International Conference New Approaches in Social and Humanistic Sciences NASHS 2017 Chisinau, Republic of Moldova.

(2) Aghaei, Majid & Rezagholizadeh, Madhya, (2017): The Impact of Growth in the OLC Countries, Economic and Environmental Studies, Vol. 17, No. 2.

(3) Amiri, Shahram & Woodside, Joseph (2017): Emerging Markets, The Impact of ICT on the Economy and Society, Digital Policy, Regulation and Governance.

(4) Arellano, C. Aguilera (2021): Lasso Regression, UMN.

(5) Alkhamisi M.A., Shukur G., (2007): A Monte Carlo Study of Recent Ridge Parameters, Communications in Statistics - Simulation and Computation, Taylor & Francis, 36.

(6) Bahrini, Raef & Qaffas, Alaa, (2019): Impact of Information and Communication Technology on Economic Growth, Evidence from Developing Countries, Journal of Economies.

(7) Beran, J., Feng, Y., & Hebbel, H. (2015): Empirical economic and financial research, Theory, Methods and Practice, USA, Springer International Publishing.

(8) Buhlmann, Peter, van de Geer, Sara, (2011): Statistics for High-Dimensional Data, Methods, Theory and Applications, Springer, Heidelberg Dordrecht London New York.

(9) Dyar, M.D.; Carmosino, M.L.; Breves, E.A., Ozanne, M.V.; Clegg S.M.; Wiens, R.C., (2012): Comparison of partial least squares and lasso Regression techniques as applied to laser--induced breakdown spectroscopy of geological samples", Spectrochimica Acta Part B 70.

(10) Fonti Valeria, (2017): Feature Selection using LASSO, Research Paper in Business Analytics, VU Amsterdam.

(11) Fujino, Yuri, Murata, Hiroshi, Mayama, Chihiro Asaoka, Ryo, (2018): Applying Lasso Regression to Predict Future Visual Field Progression in Glaucoma Patients, The Association for Research in Vision and Ophthalmology, Inc.

(12) Hoerl, Arthur. E. and Kennard, Robert W., (1970): Ridge regression: Biased estimation for non-orthogonal Problems, T Econometrics Journal, Vol. 12, No. 1.

(13) J. Fu Wenjiang, (1998): Penalized Regressions: The Bridge Versus the Lasso, Journal of Computational and Graphical Statistics, Volume 7, Number 3.

(14) Li Fan; Yang Yiming; P. Xing, Eric, (2006): From Lasso regression to Feature vector machine, Pittsburgh, PA USA 15213 fhustlf,yiming, epxingg@cs.cmu.edu.

(15) Malo, Pekka, (2018): Modern Regression Analysis, Feature selection and regularization techniques, Aalto University, school of businesses, Aalto BIZ / Department of Information and Service Management.

(16) Mi Kim, Sun; Kim, Yongdai; Jeong, Kuhwan; Jeong, Heeyeong; Kim Jiyoung, (2018): Logistic LASSO regression for the diagnosis of breast cancer using clinical demographic data and

the BI-RADS lexicon for ultrasonography, Ultrasonography 37(1).

(17) Pradhan, Rudra & Mallik, Girijasankar & Bagchi, Tapan, (2018): information communication technology (ICT) infrastructure and economic growth, A causality evinced by cross-country panel data, IIMB Management Review, WWW.sciencedirect.com.

(18) Ranstam, J.; Cook, J.A., (2018): Lasso Regression, BJS Statistical Editors.

(19) Salahuddin, M., & Gow, J. (2015): The effect of the Internet on economic growth in Southern African Countries, A combination of panel and time series approaches cou.

(20) Tuna K, Kayacan E, Bectas H (2015): The relationship between research and development expenditures and economic growth, the case of Turkey, Social and Behavioral Science 195.

(21) Tibshirani, Robert, (1996): Regression Shrinkage and Selection via the Lasso, J.R. Statist. Soc. B 58, No. 1.

(22) Tibshirani, Robert, (1997): The LASSO method for variable selection in Cox Model, J.R. Statist. Soc. B 58, No. 1.

(23) Tibshirani, Ryan, (2013): Modern regression, Optional reading: ISL 6.2.1, SEL 3.4.1, Data Mining: 36-462/36-662.

(24) Wang, Hansheng; Li, Guodong, Tsai, Chih-Ling, (2007): Regression coefficient and autoregressive order shrinkage and selection via the lasso, J.R. Statist, Soc. B, 69, Part 1.

(25) Yi Congrui; Huang Jian, (2016): Semi smooth Newton Coordinate Descent Algorithm for Elastic-Net Penalized Huber Loss Regression and Quintile Regression, arXiv: 1509.02957v2 [stat. CO] 20.

(26) Zou Hui; Hastie Trevor, (2006): Regularization and variable selection via the elastic net J.R. Statist. Soc. B, 67, Part 2.