دراسة مقارنة لبعض معالجات مشكلتي الإزدواج الخطي والإرتباط الذاتي للبواقي[1]

أ.محمد عبدالله حسين خميس
باحث ماجستير
قسم الاحصاء والرياضية والتأمين
كلية التجارة - جامعة الاسكندرية
abdmohmmed995@gmail.com

أ د. مصطفى عبدالمنعم الخواجة
استاذ الإحصاء
كلية التجارة - جامعة الاسكندرية
جمهورية مصر العربية
Mostafa.elkhwaga@alexu.edu.eg

د. احمد محمد قاروصة
مدرس بقسم الاحصاء والرياضة والتامين
كلية التجارة - جامعة الاسكندرية
جمهورية مصر العربية
Ahmed.karosa@alexu.edu.eg

ملخص البحث

تناولت هذه الدراسة مقارنة إداء مقدرr-(k,d)class والذي يضم عدة مقدرات لمعالجة المشكلتي معاً ومقدر (r-k)class وطريقة المقدرات المجمعة Combined Estimators Method (CEM) والمربعات الصغرى العادية Ordinary Least Square (OLS) عند وجود مشكلة الازدواج الخطي والارتباط الذاتي للاخطاء في البيانات. وتمت مقارنة إداء المقدرات من أجل الحصول على أفضل مقدر للتغلب على المشكلتي معاً بإستخدام معيار جذر متوسط مربعات الخطأ وتبين أن مقدر r-(k,d)class أفضل مقدر ثم يله مقدر(CEM).

الكلمات الدالة

الانحدار الخطي المتعدد، المربعات الصغرى العادية، مقدر r-(k,d)class، مقدر (r-k)class، مقدر طريقة المقدرات المجمعة (CEM) .

1 بحث مشتق من رسالة ماجستير بعنوان: كفاءة بعض مقدرات نموذج الانحدار الخطي في ظل وجود مشكلتي الارتباط السلسلي والازدواج الخطي – دراسة قياسية.

- تم تقديم البحث بتاريخ 2023/5/25 وتم قبوله للنشر بتاريخ 2023/7/14

(1) المقدمة

أن تحليل الانحدار هو احد الاساليب الإحصائية لدراسة العلاقة بين المتغيرات حيث أن المتغير التابع هو داله في متغير واحد أو أكثر من المتغيرات المستقلة (Yan & Gang (2009، والهدف من تقدير القيمة المتوسطة للمتغير التابع القيمة المعروفة او الثابتة للمتغير المستقل بمعلومية قيمة المتغير المستقل. وتعد طريقة المربعات الصغرى العادية (OLS) أكثر أستخداماً لتقدير معالم نموذج الانحدار الخطي، في ظل فروض معينه (Gunst & Mason, 1979). ومن تلك الافتراضات أن المتغيرات المستقلة غير مرتبطة وأن حدود الخطأ غير مرتبطه. مع ذلك من الناحية العملية لاتتحقق هذه الافتراضات بالتالي حدوث مشكلة الازدواج الخطي والارتباط الذاتي. يعرف الازدواج الخطي بوجد علاقة خطيه بين المتغيرات المستقلة (Muhammad, Maria & Muhammad, 2013) من اثار الازدواج الخطي أن إداء مقدر طريقة المربعات الصغرى العادية (OLS) غير كفء وتحتوي معاملات الانحدار على إخطاء معيارية كبيرة ولها اشارت خاطئة (Gujarati, 2003) ولمعالجة هذه المشكلة هناك عديد من المقدرات مثل Principal Components(PC) من قبل Massey في عام 1965 ومقدر المربعات الجزئية من قبل Wold في عام 1966 ومقدر ريدج الذي قدمه Hoerl & Kennard(1970) وقام Liu(1993) بدمج مقدر Stein مع مقدر انحدار ريدج العادي (ORR) Ordinary Ridge Regression لغرض لمعالجة الازدواج الخطي.

تحدث مشكلة الارتباط الذاتي نتيجة عدم تحقق الافتراض الخاص بأستقلال حدود الأخطاء العشوائية في نموذج الانحدار الخطي و فرضية إستقلالية حدود الخطأ العشوائي تأخذ الشكل التالي:

$$\mathrm{E}\left(U_i, U_j\right) = 0 \qquad \forall\, i \neq j \qquad (1)$$

والذي يمكن تعريف الارتباط الذاتي بين الأخطاء على انه ارتباط بين قيم السلسلة الزمنية نفسها وعبر فترات زمنية مرتدة (Koutsoyiannis, 1977) .

أما عدم تحقق شرط الأستقلال المتبادل يؤدي إلى وجود الارتباط الذاتي للبواقي مما يجعل المقدر في هذه الحالة أقل كفاءة وذلك لأن تباين المقدر يتأثر بوجود الارتباط الذاتي وتصبح التقديرات فترة الثقة أقل مصداقية رغم الارتباط الذاتي للبواقي لايؤثر في عدم التحيز المقدر أو إتساقه. من أجل معالجة مشكلة الارتباط الذاتي غالبا ما يتم استخدام طريقة المربعات الصغرى المعممة Generalized least squares method (GLS) (Greene, 2003; Wooldridge, 2002)، ويمكن أن توجد المشكلتين بشكل مشترك في نموذج الانحدار (Hussein, 2009; Firinguetti,1989) ويتم تقدير نموذج الانحدار في ظل مشاكل القياس واستخدام طرق التقدير المختلفة ولكن كل مشكلة على حده، ولمعالجة الازدواج الخطي استخدام مقدر انحدار ريدج تعطي مقدرات متحيزة ويمكن أن يكون ادائها بشكل أفضل من الطرق الاخرى. قام Hussein (2005) بأستخدام مقدر ذات المرحلتين للتعامل مع الارتباط

الذاتي وتم أستخدام مقدر انحدار ريدج Ridge وانحدار المكونات الرئيسية Principal Components (PC) وانحدار الجذور الكامنة للتعامل مع مشكلة الازدواج الخطي. وبعد ذلك تم التوصل الى إيجاد طرق تقدير جديدة مدمجة لمعالجة المشكلتين الارتباط الذاتي مابين الاخطاء والازدواج الخطي بين المتغيرات المستقلة عند ظهور المشكلتين معاً، وقد تم التوصل بأن الدمج سوف ينتج عنه مقدرات ذات خصائص جيدة وبالتالي الحصول على أفضل مقدر. أن أول من قام بإيجاد طريقة لمعالجة هاتين المشكلتين Trenkler (1984) وذلك عن طريق دمج مقدر انحدار ريدج مع مقدر المكونات الرئيسية Principal Components Estimator مع مقدر المربعات الصغرى المعممة. وقدم Arowolo, Adewale, & Kayode (2016) مقدر المربعات الجزئية ذات المرحلتين لتعامل بشكل مشترك مع المشكلتين. وقدم Huang & Yang (2015) مقدراً لمعالجة مشكلة الازدواج الخطي والارتباط الذاتي من خلال تعميم مقدر المكونات الرئيسية ذو المعلمتينPCTP Principal Components Two – Parameter Generalaized في نموذج الانحدار الخطي في ظل وجود المشكلتين معاً حيث قاما بدمج مقدر المكونات الرئيسية (PC) مع مقدر ذات المعلمتين Two –Parameter ومقدر المربعات الصغرى المعممة Generalaized least square (GLS) ومن خلال المقارنة بين مقدر المربعات الصغرى المعممة ومقدر انحدار المكونات الرئيسية ومقدر (r-k) ومقدر (r-d) من خلال نتائج المحاكاة تبين ان مقدر (PCTP) أكثر كفاءة من المقدرات الاخرى وفقاً لمعيار متوسط مربعات الخطأ. قدم Oyewole & Agunbiode (2020) مقدر GLS- Ridge(GR) لمعالجة مشكلة والارتباط الذاتي في وقت واحد من خلال دمج المربعات الصغرى المعممة(GLS) مع مقدر انحدار ريدج Ridge Regression وتكون معلمة المقدر المقترح (k) وتم مقارنة المقدر مع مقدر المربعات الصغرى المعممة وانحدار ريدج وانحدار لاسو Lasso Regression ومن خلال دراسة المحاكاة بينت النتائج أن مقدر (GR) أفضل من المقدرات التي تم المقارنة بينهما وفقاً لمعيار متوسط مربعات الخطأ وكذلك من خلال معيار Akaike Information Criterion (AIC) هو أفضل مقدر. قدم Ahmed, Ali, & Amal (2021) مقدر Almost Unbiased Liu Principal Components Regression (AULPCR) حيث قاموا بدمج كل من مقدر المربعات الصغرى المعممة(GLS) ومقدر انحدار المكونات الرئيسية Principal Components (PCR) Regression ومقدر ليو غير المتحيز تقريباً Almost Unbiased Liu Estimator ومن خلال تطبيق الدراسة لإيجاد أفضل مقدر لمعالجة المشكلتين تبين أن المقدر (AULPCR) ومقدر انحدار المكونات الرئيسية(PCR) هو أفضل مقدرين من خلال المقارنة مع بقية المقدرات وفقاً لمعيار متوسط مربعات الخطأ. قدم Dawoud (2023) مقدر Generalized Biased (GB) لمعالجة مشكلة الارتباط الذاتي بين الاخطاء والازدواج الخطي بين المتغيرات المستقلة، ومن خلال مقارنة اداء المقدر (GB) مع مقدر المربعات الصغرى المعممة ومقدر انحدار ريدج المعمم ومقدر ليو المعمم وتم اختيار مقدرات للمعلمتين K و d للمقدر (GB) ومن خلال دراسة المحاكاة تبينت النتائج أن المقدر (GB) أفضل من المقدرات المعممة وفقاً لمعيار متوسط مربعات الخطأ.

(2) نموذج الانحدار Regression Model

يعد تحليل الانحدار الخطي من أساسيات علم الإحصاء وأسلوباً مهماً من أساليب الإحصاء التطبيقي، ويعد واحد من الاساليب الإحصائية الأكثر شيوعاً وأستخداماً حيث يعمل على تحليل الظواهر لاسيما مع توافر البرامج الإحصائية الجاهزة والتي جعلت امكانية تجهيز نتائج تحليل الانحدار سهلة وسريعة فعند دراسة اي ظاهرة يجب تحديد العلاقة بين المتغيرات المستقلة والمتغير التابع ويستخدم الانحدار للتنبؤ بقيمة المتغير التابع، ويعتبر نموذج الانحدار الخطي المتعدد من النماذج الخطية المستخدمة في تحليل البيانات والذي يأخذ بالشكل التالي:

$$Y_{n,1} = X_{n,p}\,\beta_{p,1} + U_{n,1} \qquad (2)$$

حيث

$Y_{n,1}$: يمثل متجه مشاهدات المتغير التابع من الرتبة ($n \times 1$) .

$X_{n,p}$: تمثل مصفوفة المتغيرات المستقلة من الرتبة ($n \times p$) .

$\beta_{p,1}$: تمثل متجه معلمات نموذج الانحدار من الرتبة ($p \times 1$) .

$U_{n,1}$: تمثل متجه حد الخطأ العشوائي من الرتبة ($n \times 1$).

يفترض وجود ارتباط بين المتغيرات المستقله ويوجد لها ايضا خطأ عشوائي التي تتبع نموذج الانحدار الذاتي (AR1) ويأخذ الشكل التالي

$$u_t = \rho u_{t\_1} + \epsilon_t \qquad (3)$$

وحيث تكون قيمه $-1 < \rho < 1$

ϵ_t : تمثل حد الخطأ المنقى والذي يتبع التوزيع الطبيعي بمتوسط صفر وتباين ($\sigma^2 I$).

ρ : معامل الارتباط الذاتي.

للكشف عن مشكلة الارتباط الذاتي إستخدام اختبار دربن واتسون Durbin-Watson ويعد هذا الاختبار من أكثر الإختبارات شيوعاً لسهولة تطبيقة، ويستخدم هذا الاختبار في تحديد الارتباط الذاتي للأخطاء من الدرجة الاولى (Gujarati & Porter, 2008).

فروض الاختبار حيث أن الفرض العدمي القائل

لايوجد ارتباط ذاتي من الدرجة الاولى بين البواقي $H_0: \rho = 0$

أما الفرض البديل

يوجد ارتباط ذاتي من الدرجة الاولى بين البواقي $H_1: \rho \neq 0$

من أجل أختبار فروض لابد من حساب إحصائية الاختبار والتي تأخذ الشكل التالي:

$$D.W = 2(1 - \hat{\rho}) \qquad (3)$$

نجد أن إحصائية الاختبار تنحصر بين(0 و4) عندما تكون قيمة إحصائية الاختبار دربن واتسون قريبة من الصفر دل ذلك على وجود ارتباط ذاتي موجب في حين كلما كانت إحصائية الاختبار قريبة من (4) دل ذلك على وجود ارتباط ذاتي سالب، أما إذا كانت قيمة إحصائية الاختبار أقتربت من (2) دل ذلك على عدم وجود ارتباط ذاتي.

وللكشف عن مشكلة الازدواج الخطي هناك عدة أساليب منها مقايس تضخم تباينات معاملات الانحدار Variance Inflatian Factor (VIF) ويعتبر هذا المقياس من الطرق الاساسية للكشف عن مشكلة الازدواج الخطي وهو يقيس مدى تضخم تباينات معاملات الانحدار المقدرة عند وجود ارتباط خطي بين المتقيرات المستقلة (Gujarati, 2003) ، ويمكن إيجاد من خلال الصيغة التالية:

$$VIF = \frac{1}{1-R_j^2} \qquad (4)$$

والاعتماد على معامل التحديد R كلما أقترب من الواحد الصحيح فإن VIF يقترب من المالانهاية أي أنه كلما زاد الازدواج الخطي كلما زاد تباين المقدر ونهايتة تؤول الى المالانهاية، فإذا كانت قيمة معامل التحديد كبيرة وهذا يعني أن المتغير المستقل X_i له ارتباط قوي بين المتغيرات المستقلة الأخرى، ولقد أشار Ronald & Jeffrey (2006) إلى زيادة قيمة VIF عن(10) يدل على أن هناك مشكلة الازدواج الخطي.

(3) مربعات الصغرى العادية Ordinary Least Square

تعد طريقة المربعات الصغرى من الطرق المهمة والاكثر شيوعاً لتقدير معالم نموذج الانحدار الخطي المتعدد، إذ تتميز هذه الطريقة بخصائص جيدة جعلتها من أفضل الطرق وأوسعها أستخداماً، وتعتمد هذه الطريقة على مبدأ تصغير مجموع مربعات انحرافات القيم الحقيقية عن القيم المقدرة إلى أقل مايمكن ويأخذ المقدر المربعات الصغرى العادية الشكل التالي:

$$\hat{\beta}_{p,1} = (\acute{X}_{n,p} X_{n,p})^{-1} \acute{X}_{n,p} Y_{n,1} \qquad (5)$$

عندما تكون هناك مشكلة الازدواج الخطي تكون مصفوفة $\acute{X}_{n,p} X_{n,p}$ في حالتها المعتلة ill- conditioned وتقدير معاملات الانحدار تمتلك خطأ معيارياً كبير. وبشكل عام يكون مقدر المربعات الصغرى العادية غير كفء عندما يكون هناك ازدواج خطي بين المتغيرات المستقلة، وعندما تكون هناك مشكلة الارتباط الذاتي فإن مقدر المربعات الصغرى العادية على الرغم من أنه خطي وغير متحيز لم يعد له حد أدنى من التباين بين جميع المقدرات الخطية غير المتحيزة (Gujarati, 2003) وبالتالي قد لايكون اختبار t واختبار f صحيحين.

(4) طريقة المقدرات مجمعة Combined Estimators Method (CEM)

للتعامل مع المشكلتين بشكل مشترك يتم الجمع بين مقدر المربعات الصغرى ذو مرحلتين مع تقنيات التقدير المتحيزة التالية.

حيث قاما بدمج مقدر ذو المرحلتين مع مقدر انحدار المكونات الرئيسية ومقدر المربعات الصغرى الجزئية ولتكوين مقدر انحدار المكونات الرئيسية ذو مرحلتين ومقدر المربعات الصغرى الجزئية ذو مرحلتين ويتم الحصول عليه كما ذكر(Arowolo *et al*, 2016) بالخطوات على النحو التالي:

- الحصول على نموذج الانحدار الخطي المقدر بأستخدام طريقة OLS.
- استخدم مقدر معامل الارتباط الذاتي ρ لتحويل متغيرات النموذج والحصول على المتغيرات الجديدة.
- تطبيق طريقة المربعات الصغرى العادية والحصول على مقدرOLS لتقدير النموذج بعد تحويل المتغيرات.
- استخدام اختبار ديربن واتسون للتأكد من وجود الارتباط الذاتي.
- اذا تم حل مشكلة الارتباط الذاتي يتم تطبيق أسلوب التقدير المتحيز على المتغيرات بعد التحويل.

(5) مقدر(r-k) class

اقترح Baye & parker (1948) مقدر(r-k) class وتتضمن هذه الطريقه دمج ثلاث مقدرات وهم: مقدر المربعات الصغرى العادية (OLS) ومقدر انحدار ريدج العادي (ORR) حيث أن انحدار ريدج يعمل على فك الارتباطات بين المتغيرات المستقلة لغرض التغلب على مشكلة الازدواج الخطي و يعتمد انحدار ريدج على اضافة مقدار ثابت (k) إلى عناصر القطر الرئيسي للمصفوفه $\left(\acute{X}_{n,p} X_{n,p}\right)$ قبل أخذ معكوسها ومقدر المكونات الرئيسية (PCR) من المقدرات المستخدمة لمعالجة مشكلة الازدواج الخطي بين المتغيرات المستقلة فهي تحول المتغيرات التي ترتبط مع بعضها بعلاقة خطية الى مكونات خطية مستقلة ويأخذ المقدر (r-k)class الشكل الاتي:

$$\hat{\beta}_r(k) = T_{r,r}(\acute{T}_{r,r}\acute{X}_{n,p}X_{n,p}\ T_{r,r} + kI_{r,r})^{-1}\acute{T}_{r,r}\acute{X}_{n,p}Y_{n,1}\ K \geq 0 \quad r \leq p \qquad (6)$$

حيث

$T_{r,r}$: تمثل مصفوفة متعامده من الدرجة$(r \times r)$ ، $0 \leq r \leq p$.

$I_{r,r}$: تمثل مصفوفة الوحدة من الدرجة $r \times r$.

k : يمثل مقدار ثابت ويعبر عن التحيز في المقدرات.

كما قام Şiray, Kaçıranlar, & Sakallıoğlu (2014) تعميم هذا المقدر باضافة مقدرات المربعات الصغرى المعممة (GLS) بدلاً من مقدر المربعات الصغرى العادية ، ليصبح المقدر الجديد قادراً على التخلص من مشكلة الارتباط الذاتي أيضاً. وبفرض وجود المصفوفة المتعامدة T ويأخذ المقدر r-k)class) الشكل التالي:

$$\hat{\beta}_r(k) = T_{r,r}(T'_{r,r}X'_{n,p}\, V^{-1}_{n,n}\, X_{n,p}\, T_{r,r} + kI_{r,r})^{-1}T'_{r,r}X'_{n,p}V^{-1}_{n,n}\; Y_{n,1} \qquad (7)$$

حيث

$V^{-1}_{n,n}$: تمثل مصفوفة الارتباط من الدرجة $n \times n$.

(6) مقدر r-(k,d) class

أقترح Chandra & Tyagi (2017) مقدر جديد يقوم بمعالجة مشكلتي الازدواج الخطي والارتباط الذاتي معاً، ويتكون المقدر من عدة مقدرات وهم: مقدر المربعات الصغرى العاديه (OLS) ومقدر انحدار ريدج العادي (ORR) ومقدر انحدار المكونات الرئيسيه (PCR) ومقدر ليو Liu مقدراً متحيزاً ويتميز مقدر ليو (LE) على مقدر انحدار ريدج في أن $\hat{\beta}(d)$ دالة خطية في d، وبالتالي من الاسهل اختيار d في $\hat{\beta}(d)$ مقارنة بأختيار k في $\hat{\beta}(k)$ ، والمقدر ذات المعلمتين (TP) يقوم هذا المقدر على معالجة مشكلة الازدواج الخطي الشبه تام ومقدر المربعات الصغرى المعممة (GLS) تقوم هذه الطريقة على معالجة الارتباط الذاتي ومقدر r-d)class) وهو مقدر بديل لمقدر r-k)class) من خلال دمج مقدر ليو مع مقدرات المكونات الرئيسية ومقدر(r-k) class هو عبارة عن مجموعة من المقدرات المدمجة للتغلب على المشكلتين معاً ويأخذ المقدر الشكل التالي:

$$\tilde{\beta}_r(k,d) = T_{r,r}(T'_{r,r}X'_{n,p}\; V^{-1}_{n,n}X_{n,p}T_{r,r} + kI_{r,r})^{-1}(T'_{r,r}\; X'_{n,p}\; V^{-1}_{n,n}\, X_{n.p}T_{r,r})^{-1}(T'_{r,r}\; X'_{n,p}\, V^{-1}_{n,n}\, X_{n,p}T_{r,r} + kdI_{r,r})T'_{r,r}X'_{n,p}\; V^{-1}_{n,n}Y_{n,1} \qquad (8)$$

حيث

d: تمثل معلمة ليو وتكون قيمتها $0 < d < 1$.

(7) نتائج الدراسة

تناولنا بعض المقدرات التي تتغلب على مشكلة الازدواج الخطي والارتباط الذاتي في البيانات والمقارنة بين المقدرات المستخدمة للتقدير يتم تقدير نموذج الانحدار الخطي المتعدد وفقاً لبيانات قطاع الزراعة في جمهورية العراق عن انتاج القمح في الفترة الزمنية من عام 2017 إلى عام 2021 وعن 15 محافظة. وتتمثل الدراسة التطبيقية بعدة خطوات تبدأ برصد خمسة متغيرات على النحو التالي:

- كمية انتاج القمح: وهو المتغير التابع يرمز له (Y) في الدراسة ويعبر عن الكمية بالطن.

- المساحة المحصودة: ويمثل المتغير المستقل الأول ويرمز له (X_1) في الدراسة ويعبر عنه بالدونم.
- المساحة الاجمالية: وتمثل المتغير المستقل الثاني ويرمز له(X_2) في الدراسة ويعبر عنه بالدونم.
- كمية انتاج الشعير: ويمثل المتغير المستقل الثالث ويرمز له(X_3) في الدراسة ويعبر عن الكمية بالطن.
- كمية انتاج الذرة: ويمثل المتغير المستقل الرابع ويرمز له (X_4) في الدراسة ويعبر عن الكمية بالطن.

ولتقدير معالم نموذج الانحدار بأستخدام البرنامج MATLAB 2020 وتم استخدام طريقة المربعات الصغرى العادية OLS وثلاث مقدرات مدمجة للتغلب على مشكلتي الازدواج الخطي والارتباط الذاتي فى الدراسة التطبيقية وهم: طريقة المقدرات مجمعة (CEM) ومقدر r-k)class) ومقدر r-(k,d)class. وذلك لمقارنة وتقييم أداء المقدرات المدمجة فى مدى التغلب على مشكلتي القياس محل الدراسة وذلك بالاعتماد على معيار جذر متوسط مربعات الخطأ وكذلك الخطأ المعياري للمقدرات. ويستدل على وجود مشكلة الازدواج الخطي من خلال معامل التضخم (VIf) فإذا كانت قيمة معامل التضخم المناظرة لمتغير مستقل ما أكبر من 5 دل ذلك على وجود مشكلة ازدواج خطي بسبب ذلك المتغير المستقل. كما يستدل على مدى وجود الارتباط الذاتي من خلال اختبار ديربن واتسون وقيمة احصائية الاختبار له فإذا كانت قيمة (DW) تساوي 2 دل ذلك على عدم وجود ارتباط ذاتي، واذا كانت كانت قيمة (DW) أكبر من 2 دل ذلك على وجود ارتباط ذاتي سالب بينما اذا كانت كانت قيمة (DW) تساوى أقل من 2 دل ذلك على وجود ارتباط ذاتي موجب. أو من خلال القيمة الاحتمالية (p-value) فإذا كانت القيمة الاحتمالية أقل من 5% قيمة مستوى المعنوية دل ذلك على قبول الفرض البديل القائل بأنه يوجد ارتباط ذاتي بين المشاهدات.

ويوضح الجدول التالي ملخص نتائج تقدير معاملات النموذج بأستخدام معادلة رقم (5) الخاصة بطريقة المربعات الصغرى العادية.

جدول 1: تقدير معاملات النموذج بإستخدام طريقة المربعات الصغرى العادية

Variables	Coefficient	Standard. Error	T-test	Sig.	VIF
X1	1.3597	0.1092	12.4527	0.0000	2960.5500
X2	-0.3878	0.1078	-3.5981	0.9997	348.2741
X3	-0.0143	0.0097	-1.4825	0.9288	2.8850
X4	0.0224	0.0049	4.5657	0.0000	1.1381
Rsq= 0.5303		DW =0.1723 (0.000)		Jarque-Bera= 204.1924(0.001)	
RsqAdj= 0.5104		White test =33.74 (0.000)		F(3,74)= 60.07262 (0.000)	
T-table(74,0.05) = 1.6657					

يبين جدول رقم (1) تطبيق طريقة المربعات الصغرى العادية في تقدير معاملات النموذج وللحكم على جودة توفيق النموذج تستخدم عدة معايير احصائية مثل: اختبارات الفروض الخاصة بمعنوية كل معامل من معاملات الانحدار، وكذلك اختبار معنوية نموذج الانحدار ككل ويتضح من خلال القيمة الاحتمالية وتكون أقل من 5% لإختبار F اي أن النموذج معنوي ، وبعض المقايس الاحصائية لتقييم نموذج الانحدار الخطى المتعدد. ويتضح ايضاً أن قيمة معامل التضخم (VIF) لكل من المتغير المستقل (X_1) والمتغير المستقل (X_2) أكبر من 5 مما يدل على وجود ازدواج خطي فعال بسبب المتغيرين (X_2, X_1). وقد بلغت قيمة احصائية دربن واتسون 0.17 وهى قيمة بعيدة نسبياً عن الرقم 2 مما يدل على وجود مشكلة الارتباط الذاتي، ومن خلال القيمة الاحتمالية لأحصائية Jarque-Beraيتضح أن البواقي تتبع التوزيع الطبيعي.

وباستخدام بعض المقدرات المدمجة التى ذكرت فى الدراسات السابقة للتغلب على مشكلتي الازدواج الخطي والارتباط الذاتي مثل طريقة المقدرات مجمعة (CEM) ومقدر class (r-k) ومقدر r-(k,d) class. والجدول التالي ملخص لمعاملات الانحدار المقدرة في ظل طريقة المقدرات مجمعة (CEM) المدمجة.

جدول 2: تقدير معاملات النموذج بإستخدام طريقة المقدرات المجمعة CEM

Variables	Coefficient	Standard. Error	T-test	Sig.	VIF
X1	0.4731	0.1505	3.1432	0.0012	802.3484
X2	0.4807	0.1486	3.2358	0.0009	916.5839
X3	0.0068	0.0133	0.5084	0.3063	2.8588
X4	0.0193	0.0068	2.8522	0.0028	1.1429
Rsq= 0.9952		DW = 1.6010(0.0792)		Jarque-Bera= 137.9322(0.001)	
RsqAdj= 0.9950		White test = 35.89(0.000)		F(3,74)= 2942.217 (0.000)	
T-table(74,0.05) = 1.6657					

من خلال نتائج جدول رقم (2) والذي يبين تطبيق طريقة المقدرات مجمعة (CEM) في تقدير معاملات النموذج, إذ يلاحظ أن المتغير المستقل (X_1) والمتغير المستقل (X_2) والمتغير المستقل (X_4) كان لهم تأثير معنوي على المتغير التابع. وذلك لأن القيمة الاحتمالية لهم أقل من 5% كما بلغت قيمة معامل (X_1) 0.4731 بمقدار خطأ معياري 0.1505، وكذلك بلغت قيمة معامل المتغير(X_2) 0.4807 بمقدار خطأ معياري 0.1486 كما أن قيمة معامل المتغير (X_4) 0.0193 بخطأ معياري 0.0068. بينما ما يخص المتغير (X_3) فليس له تأثير معنوي على النموذج

وذلك لأن القيمة الاحتمالية له أكبر من 5%. ويتضح كذلك من خلال القيمة الاحتمالية لإختبار F اصغر من 5% اي أن النموذج معنوي.

وبالنسبة لمشكلة الازدواج الخطي فإن قيمة معامل التضخم (VIF) لكل من المتغير المستقل (X_1) والمتغير المستقل (x_2) أكبر من 5 مما يدل على وجود ازدواج خطي فعال بسبب المتغيرين (X_2, X_1). إذن طريقة المقدرات مجمعة (CEM) حسنت من أداء المتغير (X_2) وأصبح له تأثير معنوي على المتغير التابع ويستدل على ذلك من خلال القيمة الاحتمالية لهذ المتغير وتساوي 0.0009 مما يؤكد على معنوية المعلمة β_2.

ويوضح الجدول التالي ملخص نتائج تقدير معاملات النموذج بأستخدام معادلة رقم (7)الخاصة بمقدر (r-k)class.

جدول 3: تقدير معاملات النموذج بإستخدام المقدر (r-k) class

Variables	Coefficient	Standard. Error	T-test	Sig.	VIF
X1	0.4720	0.1508	3.1305	0.0012	726.4681
X2	0.4796	0.1488	3.2225	0.0009	828.7030
X3	0.0083	0.0134	0.6228	0.2677	2.8601
X4	0.0215	0.0068	3.1711	0.0011	1.1452
Rsq= 0.9950		DW =1.6086	(0.0852)	Jarque-Bera=131.3685(0.001)	
RsqAdj= 0.9948		White test = 39.71(0.000)		F(3,74)= 2477.060488 (0.000)	
T-table(74,0.05) = 1.6657					

من خلال نتائج جدول رقم (3) والذي يبين تطبيق الطريقة المدمجة (r-k) class في تقدير معاملات النموذج، إذ يلاحظ أن المتغير المستقل (X_1) والمتغير المستقل (X_2) والمتغير المستقل (X_4) كان لهم تأثير معنوي على المتغير التابع. وذلك لأن القيمة الاحتمالية لهم أقل من 5% كما بلغت قيمة معامل (X_1) 0.4720 بمقدار خطأ معياري 0.1508، وكذلك بلغت قيمة معامل المتغير(X_2) 0.4796 بمقدار خطأ معياري 0.1488 كما أن قيمة معامل المتغير (X_4) 0.0215 بخطأ معياري 0.0068. بينما ما يخص المتغير (X_3) فليس له تأثير معنوي على النموذج وذلك لأن القيمة الاحتمالية له أكبر من 5%.

وبالنسبة لمشكلة الازدواج الخطي فإن قيمة معامل التضخم (VIF) لكل من المتغير المستقل (X_1) والمتغير المستقل (X_2) أكبر من 5 مما يدل على وجود ازدواج خطي فعال بسبب المتغيرين (X_2, X_1). إذن الطريقة المدمجة -r) k)class حسنت من أداء المتغير (X_2) وأصبح له تأثير معنوي على المتغير التابع بالمقارنة بمقدر (OLS) ويستدل على ذلك من خلال القيمة الاحتمالية لهذا المتغير وتساوي 0.0009 مما يؤكد على معنوية المعلمة β_2.

يوضح الجدول التالي تقدير معاملات نموذج الانحدار بأستخدام المعادلة رقم (8) الخاصة بمقدر r-(k,d) class.

جدول 4: تقدير معاملات النموذج بإستخدام المقدر r-(k,d) class

Variables	Coefficient	Standard. Error	T-test	Sig.	VIF
X1	0.4752	0.1501	3.1651	0.0011	831.3748
X2	0.4829	0.1482	3.2588	0.0008	950.0899
X3	0.0018	0.0133	0.1322	0.4476	2.8583
X4	0.0186	0.0068	2.7520	0.0037	1.1421
Rsq= 0.9955		DW = 1.6707 (0.1486)		Jarque-Bera=161.39326(0.001)	
RsqAdj= 0.9953		White test = 37.63(0.000)		F(3,74)= 9062.4908 (0.000)	
T-table(74,0.05) = 1.6657					

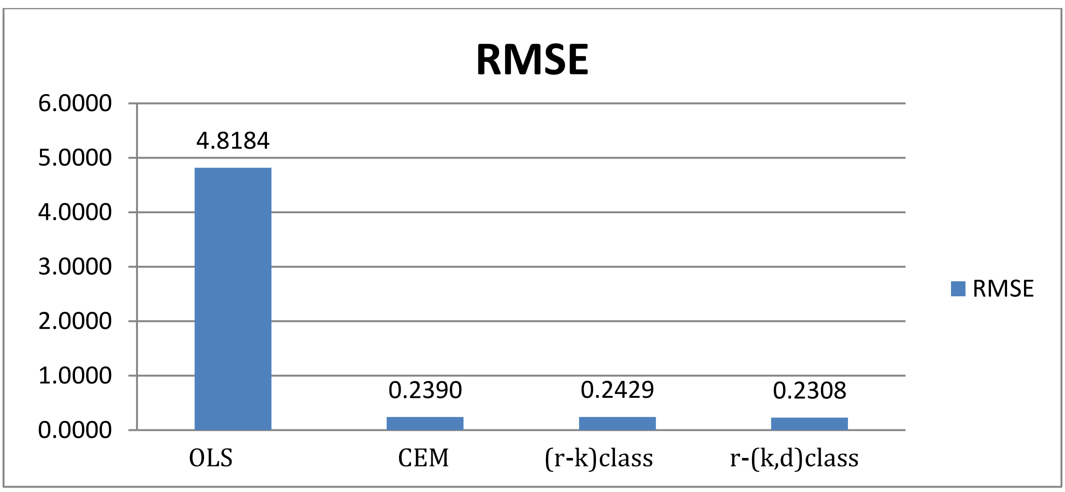
من خلال نتائج جدول رقم (4) والذي يبين تطبيق الطريقة المدمجة r-(k,d) class في تقدير معاملات النموذج، إذ يلاحظ أن المتغير المستقل (X_1) والمتغير المستقل (X_2) والمتغير المستقل (X_4) كان لهم تأثير معنوى على المتغير التابع. وذلك لأن القيمة الاحتمالية لهم أقل من 5% كما بلغت قيمة معامل (X_1) 0.4752 بمقدار خطأ معياري 0.1501، وكذلك بلغت قيمة معامل المتغير (X_2) 0.4829 بمقدار خطأ معياري 0.1482 كما أن قيمة معامل المتغير (X_4) 0.0186 بخطأ معياري 0.0068. بينما ما يخص المتغير (X_3) فليس له تأثير معنوي على النموذج وذلك لأن القيمة الاحتمالية له أكبر من 5%.

وبالنسبة لمشكلة الازدواج الخطى فإن قيمة معامل التضخم (VIF) لكل من المتغير المستقل (X_1) والمتغير المستقل (X_2) أكبر من 5 مما يدل على وجود ازدواج خطي فعال بسبب المتغيرين(X_2, X). إذن الطريقة المدمجة r-(k,d) class حسنت من أداء المتغير (X_2) وأصبح له تأثير معنوي على المتغير التابع بالمقارنة بمقدر (OLS) ويستدل على ذلك من خلال القيمة الاحتمالية لهذا المتغير وتساوى 0.0008 مما يؤكد على معنوية المعلمة β_2. كما أن استخدام طريقة r-(k,d)class للتقدير قللت من الازدواج الخطي المتسبب فيه المتغير المستقل (X_1) ويستدل على ذلك من خلال إنخفاض قيمة معامل التضخم (VIF) من 2960.5 إلى 831.3748. ولكن بالرغم من ظهور المتغير المستقل

(X_2) كمعنوي في دراسة العلاقة بينه وبين المتغير التابع إلا أن الازدواج الخطي المتسبب فيه ذلك المتغير زاد تأثيره وبشدة لأن قيمة (VIF) ارتفعت من 348.27 إلى 950.089 .

وفقاً لاختبار دربن واتسون للكشف عن مشكلة الارتباط الذاتي وجد أن قيمة الاحصائية DW 0.1723 وهى قيمة أقل من 2 مما يدل على وجود ارتباط ذاتى موجب ومن خلال القيمة الاحتمالية والتي تساوي صفر وهى قيمة أقل من 5% فإنه يتم رفض فرض العدم ومن ثم قبول الفرض البديل القائل بأنه يوجد ارتباط ذاتي بين حدود الخطأ فى النموذج. بينما فى ظل طريقة المقدرات مجمعة (CEM) وجد أن قيمة الاحصائية DW ارتفعت إلى 1.6010 والقيمة الاحتمالية والتى تساوى 0.0792 وهى قيمة أكبر من 5% مما يدل على قبول فرض العدم القائل بأنه لا يوجد ارتباط ذاتي بين حدود الخطأ العشوائى. وفى ظل المقدرات المدمجة r-k) class) وجد أن قيمة الاحصائية DW ارتفعت من 0.1723 إلى 1.6086 والقيمة الاحتمالية والتى تساوي 0.0852 وهى قيمة أكبر من 5% مما يدل على قبول فرض العدم القائل بأنه لا يوجد ارتباط ذاتي بين حدود الخطأ العشوائي. وفى ظل المقدرات المدمجة r-(k,d) class وجد أيضاً تحسناً فى قيمة الاحصائية DW والتي ارتفعت من 0.1723 إلى 1.6707 والقيمة الاحتمالية والتي تساوي 0.1486 وهى قيمة أكبر من 5% مما يدل على قبول فرض العدم القائل بأنه لا يوجد ارتباط ذاتي بين حدود الخطأ العشوائي (حسين، 2023).

للمقارنة بين اداء المقدرات التي تتغلبت على المشكلتين إستخدام معيار جذر متوسط مربعات الخطأ لغرض معرفة أفضل مقدر من بين المقدرات الدراسة من خلال الشكل ادناه.



شكل 1: مقارنة طرق التقدير المختلفة بواسطة جذر متوسط مربعات الخطأ (RMSE)

ويتضح من هذا الشكل أنه بمقارنة إداء المقدرات المدمجة مع مقدر المربعات الصغرى فإن مقدر المربعات الصغرى غير كفء للغاية فى ظل مشكلتي الازدواج الخطي والارتباط الذاتي، بينما بمقارنة أداء المقدرات المدمجة وبعضها البعض يلاحظ أن المقدر r-(k,d) class هو أفضل المقدرات المدمجة حيث يصاحب هذا المقدر قيمة (RMSE=0.2308) وهى أقل قيمة فى ظل المقدرات محل الدراسة. كما يقترب أداء المقدرين r-(k,d) class و (CEM) وفقاً لمعيار RMSE إلا أن المقدر r-(k,d) class هو الأفضل.

(8) الخلاصة

كدراسة مستخلصة من رسالة ماجستير الباحث ويعتبر وجود الازدواج الخطي والارتباط الذاتي من أكثر المشاكل تأثيراً في تحليل الانحدار الخطي. وعلى الرغم من تعدد الطرق المتاحة لعلاج كل مشكلة من المشكلتين بشكل منفصل، إلا في الواقع عادة ما تحدث المشكلتان آنيا. وكطريقة لعلاج المشكلتين يتم استخدام المقدرات المدمجة للتغلب على المشكلتين معاً وفي هذه الدراسة تم أخذ ثلاث مقدرات تمتلك القدرة على معالجة مشكلتين الازدواج الخطي والارتباط الذاتي معاً وتمت المقارنة اداء مقدر r-(k,d) class ومقدر (r-k) class مع طريقة المقدرات مجمعة (CEM) ومن خلال النتائج التي تم الحصول عليها من الدراسة التطبيقية ان أفضل مقدر هوr(k-d) class ومن ثم يأتي بعد (CEM) وفي المرتبة الاخيرة ومقدر (r-k) clas وفقاً لمعيار جذر متوسط مربعات الخطأ. وقد كان لهذه المقدرات صفات جيدة كونها تمتلك متوسط مربعات خطأ منخفض وقد كان تفوقها على بعضها بمقادير ضئيلة نسبياً ولكافة مستويات التحليل والذي يجعل من اختيار احدى هذه المقدرات في التطبيقات متساوية من حيث الكفاءة ويعود ذلك لكون عملية الدمج اعتمدت في اساسها على نفس المقدرات الرئيسية التي اشتقت منها.

(9) التوصيات

1. يوصي بضرورة اتباع تلك الطرق المدمجة المذكورة في الدراسة في حالة وجود مشكلة الازدواج الخطي والارتباط الذاتي معاً في البيانات.
2. يفضل الاعتماد على طريقة r-(k,d) class المدمجة حيث أنها الأفضل من بين مقدرات الدراسة الحالية.
3. يجب البحث واقتراح طرق تقدير جديدة تعتمد فكرتها على دمج المقدرات وخاصة المقدرات المستخدمة فى التغلب على كل مشكلة على حدة بخلاف طريقة المربعات الصغرى العادية.

المراجع

أولا :المراجع باللغة العربية

الجهاز المركزي للإحصاء (2022). بيانات سنوية عن كميات انتاج القمح من الفترة 2017 –2021 لخمسة عشر محافظة لإنتاج القمح. مديرية الإحصاء الزراعي, وزارة الزراعة, جمهورية العراق.

حسين، محمد (2023). كفاءة بعض مقدرات نموذج الانحدار الخطي في ظل وجود مشكلتي الارتباط السلسلي والازدواج الخطي – دراسة قياسية (رسالة ماجستير غير منشورة).كلية الاعمال,الاسكندرية, جمهورية مصر العربية.

ثانيا: المراجع باللغة الاجنبية

Ahmed, A. E., Ali, H. M., & Amal, H. A. (2021). Almost Unbiased Liu Principal Component Estimator in the Presence of Multicollinearity and Autocorrelation. *The Egyptian Statistical Journal*, *64*(1), 21-33.

Arowolo, T. O., Adewale, F. L., & Kayode, A. (2016). A Comparative Study of Some Method of Handling Multicollinearity in an Autocorrelated Error. *African Journal of Science and Technology (AJST)* ,13(2) ,68.

Baye, M. R. & Parker, D. F. (1984).Combining Ridge and Principal Component Regression: A money Demand Illustration. *Communications in Statistics - Theory and Methods*, Vol. 13, No. 2, pp. 197-205.

Chandra, S., & Tyagi, G. (2017). A General Class of Biased Estimators in the Presence of Multicolinearity with Autocorrelated Errors, *International Journal of mathematics and statistics*, Vol. 19, No. 2, pp. 30-49.

Dawoud, I. (2023). A Proposed Biased Regression Estimator for Treating the Existence of Autocorrelation and Multicollinearity. *The Journal of Natural Sciences*, *25*(01), 33-46.

Firinguetti, L. L. (1989). A Simulation Study of Ridge Regression Estimators with Autocorrelated Errors. *Communications in Statistics-simulation and computation*, 18(2), 673-702.

Greene, W. (2003). *Econometric Analysis*. New Jersey, Prentice Hall.

Gujarati, N. D. (2003). *Basic Econometrics (4th Ed.)*. New Delhi: Tata McGraw-Hill, 748, 807.

Gujarati, D. N., & Porter, D. C. (2008). *Basic Econometric (Edisi Kelima)*.

Gunst, R. F., & Mason, R. L. (1979). Some Considerations in the Evaluation of Alternate Prediction Equations. *Technometrics*, *21*(1), 55-63.

Hoerl, A. E. & Kennard, R.W. (1970). Ridge Regression: Biased estimation for non-orthogonal problems. *Technometrics*, 12, 55-67.

Hussein, Y. A. (2009). A Simulation Study of Ridge Regression Method with Autocorrelated Error. *Journal of Shendi University*, 7, 1-19.

Hussein, Y.A. (2005). Study of Biased Estimation Methods with Autocorrelated Errors using Simulation, Ph.D. Thesis, *Sudan University of Science and technology, Khartoum, Sudan*.

Huang, J., & Yang, H. (2015). On A Principal Component Two-Parameter Estimator in Linear Model with Autocorrelated Errors. *Statistical Papers*, *56*(1), 217-230.

Koutsoyiannis, A. (1977). *Theory of Econometrics*. Palgrave publishers, 2nd Edition.

Kejian, L.(1993).A New Class of Blased Estimate in Linear Regre-ssion. *Communications in Statistics-Theory and Methods*, 22(2), 393-402.

Muhammad, I., Maria, J. & Muhammad, A. (2013). Comparison of Shrinkage Regression Methods for Remedy of Multicollinearity Problem. *Middle-East Journal of Scientific Research, 14* (4), 570-579.

Oyewole, O., & Agunbiade, D. A.(2020). Regression Techniques in the Presence of Multicollinearity and Autocorrelation Phenomena: *monte carlo approach. Anale. Seria Informatica*. Vol. XVIII fasc.

Ronald P.C & Jeffrey K.S.(2006). Applied Statistics and the SAS Programing Language ,5th ed. ,*Pearson Prentice Hall, Pearson Education , Inc*. USA.

Trenkler, G. (1984). On the Performance of Biased Estimators in the Linear Regression Model with Correlated or Heteroscedastic Errors. *Journal of Econometrics,* Vol. 25, , No. 2, 179-190.

Üstündagˇ Şiray, G., Kaçıranlar, S., & Sakallıoğlu, S. (2014). r− k Class Estimator in the Linear Regression Model with Correlated Errors. *Statistical Papers*, *55*(2).

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Yan. X & Gang S. X.(2009). *Linear Regression Analysis, Theory and Computing*. puplished by world scientific puplishing. Co. Pte. Ltd, London.

A Comparative Study of Some Treatments For the Problems of Multicollinearity and Autocrrelation of Residuals

Mr. Mohammed Abdullah Hussein

Prof. Moustafa Abdel Moneim EL-Khawaga Dr.Ahmed Mohammed karosa

Abstract

This study deals with a comparison of the performance of an r-(k,d) class estimator which includes serveal estimators treatments for the problems, and (r-k)class estimator and combined estimators method(CEM) and 0rdinary least squares (OLS) estimator when there is a problem of multicollinearity and auto-correlated errors in the data. The performance of the estimators was compared in order to obtain the best estimator to overcome the two problems together using root mean square error.It was found that the estimator is the best r-(k,d) class estimator and then the (CEM) estimator.

Keywords

Multiple linear regression, Ordinary least squares, r(k,d) class estimator, (r-k) class estimator, Combined estimators method (CEM).

التوثيق المقترح للدراسة وفقا لنظام APA

خميس، محمد عبدالله حسين؛ الخواجة، مصطفى عبدالمنعم؛ قاروصة، احمد محمد (2023).دراسة مقارنة لبعض معالجات مشكلتي الإزدواج الخطي والإرتباط الذاتي للبواقي. **مجلة جامعة الإسكندرية للعلوم الإدارية**، كلية التجارة، جامعة الإسكندرية 60(4)،347-363.