تطبيق انحدار الحافة لمعالجة التعدد الخطي في تحليل العوامل المناخية المؤثرة على هطول الأمطار في مطار طرابلس الدولي.

عبدالحليم مولود الصويعي[1]
قسم الإحصاء، كلية العلوم
جامعة الزاوية، ليبيا
halimsuaiee@zu.edu.ly

نادية عبدالله محمد الذيب[2]
قسم العلوم الرياضية، مدرسة العلوم الأساسية
الأكاديمية الليبية، ليبيا
nadiaaldeeb71@gmail.com

ملخص Abstract

يتناول هذا البحث مشكلة التعدد الخطي في نماذج الانحدار المتعدد، وهي ظاهرة تنشأ عند وجود علاقات ارتباطية بين المتغيرات التفسيرية، مما يؤدي إلى تضخم الأخطاء المعيارية وتباينات التقديرات، وبالتالي إضعاف القدرة التنبؤية للنموذج. تم استخدام بيانات الأرصاد الجوية لمحطة مطار طرابلس الدولي (2000-2010) كحالة دراسية، حيث مثلت كمية الأمطار المتغير التابع. اعتمدت الدراسة على أدوات تشخيصية متعددة مثل مصفوفة الارتباط، معامل تضخم التباين(VIF) ، العدد الشرطي(CN) ، والجذور المميزة(EV) ، وطريقة فرار جلوبر. أظهرت النتائج وجود مشكلة التعدد الخطي في البيانات، وأثبتت المقارنة بين طريقتي التقدير أن طريقة انحدار الحافة تتفوق على طريقة المربعات الصغرى العادية من حيث استقرار التقديرات وانخفاض متوسط مربعات الخطأ(MSE) ، مما يجعلها الخيار الأمثل لمعالجة هذه المشكلة الإحصائية.

الكلمات المفتاحية: التعدد الخطي، انحدار الحافة، العوامل المناخية

1. مقدمة

تُعد نماذج تحليل الانحدار من الموضوعات الأساسية في علم الإحصاء، حيث تُستخدم على نطاق واسع في مختلف المجالات العلمية لدراسة العلاقات بين المتغيرات. ويُعَدُّ نموذج الانحدار الخطي المتعدد أحد أكثر النماذج شيوعًا، إذ يهدف إلى تحليل العلاقة بين متغير الاستجابة (Y) ومجموعة من المتغيرات التفسيرية(X) . لضمان دقة التقديرات الإحصائية، يتطلب تحليل الانحدار تحقيق مجموعة من الافتراضات الأساسية، غير أن انتهاك أي من هذه الافتراضات قد يؤدي إلى تدهور جودة التقديرات المستخلصة باستخدام طريقة المربعات الصغرى العادية. ومن بين المشكلات الإحصائية التي قد تظهر نتيجة لذلك: عدم تجانس التباين(Heteroscedasticity) ، والارتباط الذاتي بين الأخطاء العشوائية(Autocorrelation) ، والتعدد الخطي بين المتغيرات التفسيرية(Multicollinearity) ، مما قد يؤدي إلى تقديرات ضعيفة وغير موثوقة، تُعرف بما يسمى بالانحدار الزائف.

نظرًا لأهمية مشكلة التعدد الخطي في تحليل الانحدار وتأثيرها السلبي على دقة الاستنتاجات الإحصائية، اقترح العديد من الباحثين حلولًا لمعالجتها، حيث أشار فيشر (Fisher) عام 1934 إلى تأثير هذه المشكلة على نتائج التحليل الإحصائي. ومنذ ذلك الحين، قدمت العديد من الدراسات التطبيقية طرقًا مختلفة لمعالجة التعدد الخطي، ومن أبرزها: انحدار الحافة (Ridge Regression)، وأسلوب المكونات الرئيسية (Principal Components Method)، وطريقة المربعات الصغرى المقيدة (Restricted Least Square Method)، وانحدار الجذور الكامنة (Latent Root Regression)، ومقدرات بييز (Bayes Estimators).

يركز هذا البحث على دراسة مشكلة التعدد الخطي وتقديم حلول لمعالجتها من خلال استخدام طريقة انحدار الحافة(Ridge Regression) ، التي طُرِحَت لأول مرة من قبل Hoerl و Kennardعام 1970، نظرًا لأهميتها في الحد من تأثير التعدد الخطي وتحسين دقة تقديرات معاملات نموذج الانحدار. ومن ثم تطبيقها على بيانات واقعية للوصول إلى حل لمشكلة التعدد الخطي بين المتغيرات التفسيرية

2. انحدار الحافة Ridge Regression

يُعَدُّ انحدار الحافة (Ridge Regression) من الطرق الإحصائية الفعالة لمعالجة مشكلة التعدد الخطي (Multicollinearity) التي تؤثر على دقة تقدير معاملات الانحدار في نموذج المربعات الصغرى العادية .(OLS) تقوم هذه الطريقة على إضافة معلمة تحيز k إلى القطر الرئيسي لمصفوفة المعلومات $X'X$ ، مما يقلل من تضخم التباين ويمنح النموذج استقرارًا أكبر. وقد طُرِحَت هذه الطريقة لأول مرة من قبل Hoerl عام 1962 كحل لتقليل عدم استقرار المقدرات وتحسين دقتها في ظل وجود التعدد الخطي.

تتمثل الفكرة الأساسية وراء انحدار الحافة في تحقيق توازن بين التحيز والتباين، حيث يتم تقليل التباين على حساب إدخال مقدار صغير من التحيز، مما يؤدي إلى الحصول على تقديرات أكثر استقرارًا وأقل عرضة للتغيرات الكبيرة عند استخدام عينات مختلفة من البيانات. وفقًا لدراسة حديثة أجراها El-Habilو(2023) Almoghari، فقد أظهرت المقارنة بين انحدار الحافة وأساليب أخرى، مثل تحليل الانحدار الجزئي (Partial Regression)والانحدار السببي(Causal Regression) ، أن انحدار الحافة يتفوق من حيث استقرار المقدرات ودقتها. كما أكد Kuhn و Johnson (2022)في كتابهما "Applied Predictive Modeling"على أن انحدار الحافة يُعَدُّ إحدى أكثر الطرق فعالية وموثوقية في معالجة مشكلة التعدد الخطي.

تعتمد تقنية انحدار الحافة على إضافة مقدار موجب صغير $k \geq 0$ 0 يُعرف بمعلمة التحيز إلى مصفوفة المعلومات، مما يؤدي إلى تقليص قيم المقدرات وتقليل متوسط مربعات الأخطاء .(MSE) ومع ذلك، فإن هذه المقدرات تصبح متحيزة مقارنة بمقدرات المربعات الصغرى العادية. وبالتالي، عند استخدام انحدار الحافة، يجب الموازنة بين مقدار التحيز والتباين، بحيث يتم اختيار القيمة المثلى لمعلمة التحيزk التي تحقق أدنى مستوى لمتوسط مربعات الأخطاء. وكما أوضح(2012) Younker ، فإن تقليل التباين الناتج عن إدخال معلمة التحيز يجب أن يكون أكبر من الزيادة في التحيز لضمان تحسين جودة التقديرات.

عند تطبيق انحدار الحافة، يتم تحويل المتغيرات إلى الصورة القياسية من خلال طرح الوسط الحسابي لكل متغير والقسمة على انحرافه المعياري، مما يؤدي إلى تحويل مصفوفة $X'X$ إلى مصفوفة ارتباط المتغيرات التفسيرية، وفقًا لما ذكره محمد (2011). وبخلاف طريقة المربعات الصغرى العادية التي تركز على تقليل مجموع مربعات البواقي(RSS) ، فإن انحدار الحافة يسعى إلى تقليل المقدار $(RSS + k\sum_{j=1}^{p}\beta_j)$، حيث تعمل معلمة التحيز كعقوبة انكماش (Shrinkage Penalty) تؤدي إلى تحسين استقرار النموذج وتقليل الأخطاء التقديرية .ZACH (2020)

للتحكم في التباين والحد من تضخم معاملات الانحدار، يتم فرض قيود على مقدرات المربعات الصغرى بحيث لا تؤدي إلى قيم كبيرة وغير مستقرة. ومن بين هذه القيود، يُعد قيد الحافة أحد الأساليب المستخدمة لتقليل مجموع مربعات الأخطاء. في هذا السياق، قدَّم Hoerl و Kennard (1970)

تعديلًا على نموذج الانحدار الخطي المتعدد من خلال إعادة تشكيل مصفوفة التصميم $X'X$، وذلك بهدف تقليل دالة الخسارة وفقًا للصيغة التالية:

$$l_{ridge}(\beta,k) = \|Y - X\beta\|_2^2 + k\|\beta\|_2^2 = \sum_{i=1}^{n}(Y_i - X_i\beta)^2 + k\sum_{j=1}^{p}\beta_j^2 = RSS + k\sum_{j=1}^{p}\beta_j^2 \quad (1)$$

حيث:

Y_i: القيم الفعلية لمتغير الاستجابة

X_i: المتغيرات التفسيرية

β: معالم النموذج

β_j: ميل X_i على Y

k: معلمة التحيز

للحصول على مقدرات الحافة $\hat{\beta}$ نقوم باشتقاق المعادلة (1) بالنسبة للمعلمة β، ونساوي المشتقة بالصفر، (سلامي، حجاب،2018؛ الكفيشي،2019؛ Saleh et al.،2019) :

$$\min_{\beta}[(Y - X\beta)'(Y - X\beta) + k\sum_{j=1}^{p}\beta_j^2] = \min[f(\beta)]$$

حيث:

$$f(\beta) = Y'Y - 2\beta'X'Y + \beta'X'X\beta + k\beta'\beta$$

$$\therefore \hat{\beta}_{rr} = (X'X + kI)^{-1}X'Y \quad ; k \geq 0 \quad (2)$$

حيث:

I : مصفوفة الوحدة ذات ابعاد $(p \times p)$

Y : متجه في أبعاد $(n \times 1)$ لمتغير الاستجابة

X: مصفوفة ذات أبعاد $n \times (p+1)$ للمتغيرات التفسيرية عددها p

k: معلمة التحيز

نلاحظ أن معلمة التحيز تتحكم في مقدار مصطلح العقوبة، حيث أنه: عندما تكون $k = 0$ فإن المقدرات التي نتحصل عليها هي مقدرات المربعات الصغرى الاعتيادية، وعندما تكون $k > 0$ فإن مقدرات الحافة تميل للاستقرار عند قيمة معينة ل k ويقل الارتباط الخطي بين المتغيرات التفسيرية، وعندما تكون $k \to \infty$ فإن مقدرات الحافة تؤول إلى الصفر، (Tibshirani,2019).

3. افتراضات انحدار الحافة Assumption of Ridge Regression

نفس الافتراضات المستخدمة في الانحدار الخطي المتعدد إلا أن انحدار الحافة لا يوفر حدودا للثقة، فلا داعي لافتراض الطبيعة (normality):

1. الأخطاء العشوائية تكون مستقلة أي أن:
$$cov(e_i, e_j) = 0، \ i \neq j$$
2. المتغيرات التفسيرية تكون مستقلة أي أن:
$$cov(X_i, X_j) = 0 \ ،i \neq j$$
3. الأخطاء العشوائية تكون مستقلة عن المتغيرات التفسيرية:
$$cov(X_i, e_i) = 0$$
4. متوسط الخطأ العشوائي = 0 أي أن: $E(e_i) = 0$
5. تباين الأخطاء العشوائية ثابت أي أن: $Var(e_i) = \sigma^2$

4. خصائص مقدرات انحدار الحافة

Properties of Ridge Regression Estimators

مقدرات الحافة لها العديد من الخصائص، والتي نلخصها في النقاط التالية (Ahmed,2020).

4.1 متوسط مقدرات انحدار الحافة

نعلم أن:

$$Y = X\underline{\beta} + \underline{e} \quad , \quad E(e_i) = 0 \quad , \quad E(Y) = X\beta$$

$$\hat{\beta}_{rr} = (X'X + kI)^{-1}X'Y$$

بأخذ التوقع للطرفين:

$$E(\hat{\beta}_{rr}) = E[(X'X + kI)^{-1}X'Y]$$

$$\therefore E(\hat{\beta}_{rr}) = (X'X + kI)^{-1}X'X\beta \quad (3)$$

بفرض أن:

$$Z = (X'X + kI)^{-1}X'X$$

$$\therefore E(\hat{\beta}_{rr}) = Z\beta \quad (4)$$

من المعادلة (4) نجد أن مقدرات الحافة متحيزة لمقدرات المربعات الصغرى الاعتيادية.

4.2 تباين مقدرات انحدار الحافة

$$var(\hat{\beta}_{rr}) = var[(X'X + kI)^{-1}X'Y]$$
$$= (X'X + kI)^{-1}X'var(Y)X'(X'X + kI)^{-1}$$
$$= \sigma^2(X'X + kI)^{-1}X'X(X'X + kI)^{-1}$$
$$\therefore var(\hat{\beta}_{rr}) = \sigma^2 Z(X'X)^{-1}Z' \quad (5)$$

4.3 التحيز في مقدرات انحدار الحافة

من المعادلة (3)، (AlNasser,2014) بفرض أن:

$$W = (X'X + k\,I_p)$$
$$E(\beta) = \beta$$
$$\therefore E(\hat{\beta}_R) = W^{-1}X'X\beta$$

$$bias(\hat{\beta}_R) = E[\hat{\beta}_R - \beta] = W^{-1}X'X\beta - \beta$$
$$= W^{-1}(X'X + kI_p - kI_p)\beta - \beta = (I_p - kW^{-1})\beta - \beta$$
$$= -kW^{-1}\beta$$
$$= -k(X'X + kI_p)^{-1}\beta \quad (6)$$

4.4 متوسط مربعات الخطأ لمقدرات انحدار الحافة

يعتبر متوسط مربعات الاخطاء من أهم المقاييس التي تدل على جودة التقدير، لذلك غالبا ما يستخدم لمقارنة أداء نماذج الانحدار، وهو عبارة عن التباين مضافا إليه مربع مقدار التحيز، (El-Dereny & Rashwan 2011؛ Batah,Gore.2014؛ Kibria,Banik,2016؛ حمود،2018).

$$MSE(\hat{\beta}_{rr}) = var(\hat{\beta}_{rr}) + (bias\ in\ \hat{\beta}_{rr})^2$$
$$\therefore MSE((\hat{\beta}_{rr}) = \sigma^2 tr[(X'X + kI_p)^{-2}X'X] + k^2\beta'(X'X + kI_p)^{-2}$$
$$= \sigma^2\sum_{i=1}^{p}\frac{\lambda_i}{(\lambda_i + k)^2} + k^2\beta'(X'X + kI_p)^{-2}\beta \quad (7)$$

من الواضح أنه كلما زاد مقدار التحيز أدى ذلك إلى انخفاض مقدار التباين وذلك عند زيادة قيمة معلمة التحيز، لهذا علينا اختيار قيمة k التي تجعل الانخفاض في التباين أكثر من الزيادة في مقدار مربع التحيز، لنتحصل على متوسط مربعات أخطاء لمقدرات الحافة أقل من متوسط مربعات الاخطاء لمقدرات المربعات الصغرى الاعتيادية، انظر (Montgomery & Peck 1982).

5. العلاقة بين مقدرات المربعات الصغرى ومقدرات الحافة

The Relationship Between OLS and RRE Estimators

تتناسب مقدرات الحافة مع مقدرات المربعات الصغرى الاعتيادية تحت شرط التعامد أي:

$$X'X = I_p = (X'X)^{-1}$$
$$\hat{\beta}_{rr} = (X'X + kI)^{-1}X'Y = (1 + k)^{-1}I_p\,X'Y$$
$$= (1 + k)^{-1}(X'X)^{-1}\,X'Y = (1 + k)^{-1}\hat{\beta}_{ols}$$
$$= \frac{1}{1 + k}\hat{\beta}_{ols} \quad (8)$$

كبداية لانحدار الحافة، يوصي Saleh وآخرون (2019)، بالإضافة إلى Yahya وOlaifa (2014)، وSharif وDuzan (2015)، وFan وآخرون (2016)، وSharif وDuzan (2018)، بتوحيد المتغيرات (التابع والمستقلة). وذلك بإجراء تحويلة الارتباط (Correlation

Transformation)، لغرض توحيد قياسات المتغيرات. يتم ذلك في الخطوات التالية:

1. حساب معامل الارتباط بين المتغير التابع والمتغيرات المستقلة.
2. استخراج الجذر التربيعي لمعاملات الارتباط.
3. ضرب المتغيرات المستقلة في معاملات الارتباط المستخرجة.

هناك عدة مبررات علمية لتوحيد المتغيرات قبل تطبيق انحدار الحافة، منها:

- تحسين دقة تقديرات المعاملات :تؤدي عملية توحيد المتغيرات إلى تحسين دقة تقديرات المعاملات، وذلك لأن قيمة التباين المعياري للمتغيرات بعد توحيدها تكون أكثر ملاءمة لمقارنة أهمية المتغيرات المختلفة.
- زيادة الاستقرار الإحصائي للمعاملات :تؤدي عملية توحيد المتغيرات إلى زيادة الاستقرار الإحصائي للمعاملات، وذلك لأن قيمة التباين المعياري للمتغيرات بعد توحيدها تكون أقل عرضة للتأثر بقيم الملاحظات الفردية.
- سهولة تفسير النتائج :تؤدي عملية توحيد المتغيرات إلى تسهيل تفسير النتائج، وذلك لأن المتغيرات بعد توحيدها تكون لها نفس وحدة القياس.

$$Y_i^* = \frac{1}{\sqrt{n-1}} \cdot \frac{(Y_i - \bar{Y})}{S_y}$$

$$X_{ri}^* = \frac{1}{\sqrt{n-1}} \cdot \frac{(X_{ri} - \bar{X}_r)}{S_r}$$

حيث:

$$\bar{Y} = \frac{\sum Y_i}{n} \quad ، \quad \bar{X} = \frac{\sum X_{ri}}{n}$$

$$S_y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}}$$

$$S_r = \sqrt{\frac{\sum (X_{ri} - \bar{X}_r)^2}{n-1}} \quad , r = 1,2,\ldots\ldots,p \quad ، \quad i = 1,2,\ldots\ldots,n$$

بعد اجراء تحويل الارتباط فإن المصفوفة $(X'X)$، تتحول إلى مصفوفة الارتباط بين أزواج المتغيرات التفسيرية، فتصبح على النحو التالي:

$$R_{xx} = X^{*\prime} X^*$$

أيضا المتجه $(X'Y)$، يتحول إلى متجه الارتباط بين المتغير التابع والمتغيرات التفسيرية فيصبح على النحو التالي:

$$R_{xy} = X^{*\prime} Y^*$$

وبالتالي نتحصل على مقدرات الحافة من الصيغة التالية:

$$\hat{\beta}_{rr}^* = (R_{xx} + kI)^{-1} R_{xy} \qquad (9)$$

حيث:

$\hat{\beta}_{rr}^*$: متجه معاملات انحدار الحافة

R_{xx}: مصفوفة معاملات الارتباط البسيط بين المتغيرات التفسيرية

R_{xy}: متجه معاملات الارتباط البسيط بين المتغير التابع والمتغيرات التفسيرية

k: معلمة التحيز $1 \geq k \geq 0$

I : مصفوفة الوحدة برتبة $(p \times p)$

ويكون عامل تضخم التباين لانحدار الحافة عبارة عن قيم عناصر المصفوفة التالية:

$$(R_{xx} + kI)^{-1} R_{xx} (R_{xx} + kI)^{-1} \qquad (10)$$

ويحسب مجموع مربعات البواقي من الصيغة التالية:

$$\mathrm{RSS_{rr}} = \sum (\mathrm{Y_i^*} - \hat{\mathrm{Y}}_\mathrm{i}^*)^2 \qquad (11)$$

ويحسب معامل التحديد لانحدار الحافة من الصيغة التالية:

$$R_{rr}^2 = 1 - \mathrm{RSS_{rr}} \qquad (12)$$

بعد تطبيق المعادلات السابقة نتحصل على النموذج التالي:

$$Y_i^* = \beta_1^* X_{i1}^* + \beta_2^* X_{i2}^* + \cdots + \beta_p^* X_{ip}^* + e_i^* \quad , i = 1,2,\ldots n \qquad (13)$$

يمكن اجراء التحويل السابق باستخدام مصفوفة التمركز centring matrix ويتم ذلك باتباع الخطوات التالية:

1- توسيط البيانات Centring the data:

يتم توسيط البيانات باستخدام مصفوفة التمركز الموضحة بالمعادلة التالية:

$$\mathrm{M} = \left(I - \frac{1}{n} 11'\right)$$

هذه المصفوفة تلعب دورا مهما حيث تستخدم في توسيط مصفوفة البيانات، أي تجعل جميع الأعمدة في المصفوفة لها متوسط يساوي صفر.

حيث:

M: مصفوفة متماثلة كاملة القوة أي أن: $M = M^2$ ، $M = M'$

I: مصفوفة الوحدة

بوضع: $\mathrm{Y}_c = \mathrm{MY}$ ، $X_c = MX$ ، $\varepsilon_c = M\varepsilon$

عليه فإن: $MY = MX\beta + M\varepsilon$

$$\therefore \; Y_c = X_c\beta + \varepsilon_c \qquad (14)$$

2- تحجيم البيانات scaling the data:

كذلك عندما تكون المتغيرات لها وحدات قياس مختلفة مما يؤدي إلى صعوبة اختيار معامل تنظيم مناسب لتقليل التباين المتضخم بسبب وجود التعدد الخطي لذلك نقوم بتحويل المتغيرات على النحو التالي:

نعرف مصفوفة القياس للمتغيرات بالمعادلة:

$$\mathrm{X}_c \Rightarrow \mathrm{D} = \left(diag\left(\frac{1}{n} X_c' X_c\right)\right)^{-\frac{1}{2}}$$

$$\mathrm{Y}_c \Rightarrow d_y = \sqrt{\frac{1}{n} 1' Y_c}$$

$$\therefore \; Y_{cs} = \left(\frac{1}{d_y}\right) Y_c \quad ، \quad X_{cs} = D^{-1} X_c$$

ونحصل على مقدرات الحافة من الصيغة التالية:

$$\hat{\beta}_{cs}(k) = (X_{cs}' X_{cs} + kI)^{-1} X_{cs}' Y_{cs} \qquad (14)$$

6. معلمة الحافة *Ridge Parameter k*

أول من تناول معلمة الحافة في دراساته كلا من Hoerl & Kennard,(1970)، ثم قام Vinod في عام (1976) بإجراء دراسة تطبيقية على انحدار الحافة لاختيار معلمة التحيز، معلمة التحيز هي عبارة عن قيمة موجبة غالبا يتم اختيارها بين [0,1]، تضاف إلى القطر الرئيسي للمصفوفة $(X'X)$ قبل إيجاد المعكوس لها، هذا المقدار يقلل تباين مقدرات الحافة بالتالي يتم الحصول على مقدرات أكثر استقرار.

6.1 اختيار معلمة التحيز k (معلمة الحافة) Choosing of Ridge Parameter

تلعب معلمة التحيز (k) دورا مهما في تقدير معلمات الحافة، حيث أنه بالنظر إلى معادلة تقدير معالم الانحدار نجد أن انحدار الحافة دالة في معلمة التحيز، اقترح الباحثون عدة طرق لاختيار معلمة التحيز منها الطرق البيانية ومنها الطرق التحليلية وهي:

6.1.1 طريقة أثر الحافة The Ridge Trace

اقترح Hoerl and Kennard طريقة تعرف بأثر الحافة، وهي تعتبر الطريقة الأكثر استخداما لاختيار معلمة الحافة، حيث تستخدم هذه الطريقة مخطط أثر الحافة Ridge Trace (تتبع الحافة) في اختيار معلمة التحيز k، تعتمد هذه الطريقة على الرسم البياني لتحديد القيمة المثالية لمعلمة التحيز، فبفرض أنه لدينا عدد p من مقدرات الحافة $\hat{\beta}_{RR}$ كدوال في معلمة التحيز k، فإن هذه المقدرات ترسم مقابل قيم k في شكل بياني، حيث يمثل المحور الرأسي مقدرات الحافة ويمثل المحور الأفقي القيم المختلفة لمعلمة التحيز k، وذلك ضمن مدى معين لقيم k يقع بين [0,1]، يقوم الباحث بتتبع الرسم ومع زيادة قيم k إبتداء من الصفر، يقوم الباحث باختيار $k = k_{trace}$ التي تشهد عندها المقدرات استقرارا واضحا، ويفضل اختيار قيمة صغيرة والتي تعطي

تحيز أصغر، ونلاحظ أنه كلما زادت قيمة معلمة التحيز كلما اقتربت مقدرات الانحدار من الصفر، تعتمد هذه الطريقة على خبرة الباحث في اختيار القيمة المثالية، والتي يلاحظ عندها انخفاض في عوامل تضخم التباين، ومتوسط مربعات الاخطاء، وذلك مقابل أدنى تحيز، Mardikyan, Cetin (2008؛ Bas et al. ,2017؛ السباح، مزهر،2018؛ CROPPER and PHILIP,2020).

اقترح العديد من الباحثين طرقاً مختلفة لاختيار معلمة الحافة نذكر منها ما يلي:

6.1.2 طرق تقدير k Methods Of Estimate

نظرا لأن انحدار الحافة دالة في معلمة التحيز، فقد قامت العديد من الدراسات بتقدير معلمة التحيز بعدة طرق وذلك للوصول إلى القيمة المثالية التي عندها يتم التخلص من مشكلة التعدد الخطي، والحصول على متوسط مربعات أخطاء لمقدرات الحافة أقل من متوسط مربعات أخطاء مقدرات المربعات الصغرى الاعتيادية، العديد من الباحثين والمؤلفين طوروا واقترحوا تقديرات مختلفة لمعلمة التحيز، على سبيل المثال لا الحصر نذكر منهم، Hoeral and Kennar (1970), Hoerl et al. (1975), McDonald and Galarneau (1975), Lawless and Wang (1976), Khalaf and Shukur(1977), Alkhamisi and Shukur (2007), Gruber et al. (2010), Aslam (2014), and Kibria and Banik (2016), وغيرهم، سنذكر بعض الطرق المستخدمة في تقدير معلمة التحيز k:

1- طريقة Hoerl & Kennard

$$k_{HK} = \frac{\hat{\sigma}^2}{\hat{\beta}_{max}} \qquad (15)$$

حيث:

$\hat{\sigma}^2$: التباين المقدر بطريقة المربعات الصغرى

$\hat{\beta}_{max}$: أكبر قيمة في المعالم المقدرة بطريقة المربعات الصغرى

$$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{n-p-1}$$

$$\hat{\beta}_{ols} = (X`X)^{-1}X`Y$$

2- طريقة Hoerl,Kennard &Baldwin:

يرمز لها بالرمز [HKB]، وضعت هذه الطريقة في عام 1975 حيث تحسب قيمة k بالصيغة الآتية:

$$k_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}_{ols}{}'\,\hat{\beta}_{ols}} \qquad (16)$$

حيث: p: عدد المتغيرات التفسيرية

3- طريقة Lawless & Wang

نشر كلا من *Lawless & Wang* عام (1976) في ورقتهما طريقة لتقدير معلمة الحافة وهي تعتمد على طريقة k_{HKB} وفقا للصيغة التالية:

$$k_{LW} = \frac{p\hat{\sigma}^2}{\sum_{j=1}^{p} \lambda_j \hat{\beta}_j^2} \qquad (17)$$

حيث: λ_j: الجذور المميزة للمصفوفة $X'X$

4- طريقة Khalaf & Shukur

في عام (2005) اقترح كلا من *Khalaf* & Shuku صيغة جديدة تعتبر تعديل لصيغة k_{HK} لتقدير معلمة الحافة وفقا للصيغة التالية:

$$\hat{k}_{KS} = \frac{\lambda_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + \lambda_{max}\hat{\beta}_{max}^2} \qquad (18)$$

حيث: λ_{max}: أكبر جذر مميز لجذور للمصفوفة $X'X$

5- طريقة Alkhamisi & Shukur

قدم كلا من Alkhamisi & Shukur في ورقتهما عام (2007) عدة طرق لتقدير معلمة الحافة، نذكر منها:

$$\hat{k}_{AS} = \frac{\lambda_{max}\hat{\sigma}^2 + \hat{\beta}_{max}^2}{\lambda_{max}\hat{\beta}_{max}^2} = \frac{\hat{\sigma}^2}{\hat{\beta}_{max}^2} + \frac{1}{\lambda_{max}}$$
$$= k_{HK} + \frac{1}{\lambda_{max}} \qquad (19)$$

7. الجانب التطبيقي Applied Aspect

7.1 المقدمة Introduction

أجريت هذه الدراسة باستخدام بيانات شهرية تمثل عوامل مناخية مأخوذة من محطة مطار طرابلس الدولي خلال الفترة الممتدة من عام 2000 إلى 2010. وتهدف الدراسة إلى تحليل تأثير العوامل المناخية المختلفة على كمية سقوط الأمطار (Y)، والتي تُعد المتغير التابع في النموذج الإحصائي. شملت المتغيرات التفسيرية مجموعة من العوامل المناخية التي يُفترض أن تؤثر في كمية الأمطار المتساقطة، حيث تضمنت درجة الحرارة الصغرى (X_1)، ودرجة الحرارة العظمى (X_2)، والرطوبة النسبية (X_3)، وسرعة الرياح بالعقدة (X_4)، وكمية التبخر (X_5)، بالإضافة إلى مدة سطوع الشمس (X_6).

تتم صياغة جميع متغيرات الدراسة السابق ذكرها في النموذج التالي:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + e_i \;, i = 1,2,\dots,132$$

وباستخدام البرنامج الاحصائي لغة R تم الحصول على نتائج الدراسة، حيث كان حجم العينة n=132.

7.2 تقدير معالم نموذج الانحدار الخطي المتعدد باستخدام طريقة المربعات الصغرى الاعتيادية

Parameters Estimation of Multiple Linear Regression Model using the Ordinary Least Squares method

تم تقدير معالم نموذج الانحدار باستخدام طريقة المربعات الصغرى الاعتيادية والتي تمتلك خاصية أفضل مقدر خطي غير متحيز، حيث تم تقدير معالم نموذج الانحدار $\underline{\beta}$ باستخدام المعادلة الموضحة بالصيغة التالية:

$$\hat{\beta}_{ols} = (X'X)^{-1}X'Y$$

تم الحصول على النتائج الموضحة بالجدول (1):

جدول (1): تحليل الانحدار باستخدام المربعات الصغرى الاعتيادية

	Coeff	SE	T value	P value
Intercept	123.524	33.69	3.667	<0.001
X1	-8.903	1.58	-5.627	<0.001
X2	7.813	1.48	5.287	<0.001
X3	0.294	0.36	0.807	0.4213
X4	0.537	1.31	0.409	0.6829
X5	-0.535	1.76	-0.304	0.7615
X6	1.146	1.48	0.772	0.4414
ResidSE = 18.54	R^2=0.546	R^2Adj. =0.524	F(6,125) = 25.07	Pvalue < 001

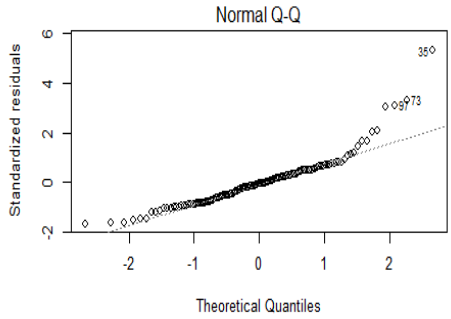
من نتائج تحليل الانحدار في الجدول (1) نجد أن معاملات المتغيران $(X5, X1)$ تمتلكان إشارة سالبة وهذا يعني أنها ذات تأثير سالب (عكسي) على متغير الاستجابة، في حين أن معاملات المتغيرات $(X6, X4, X3, X2)$ ذات إشارة موجبة أي أن لها تأثير موجب (طردي) على متغير الاستجابة.

نلاحظ أيضا أن المتغيران $(X2, X1)$ فقط لها تأثير معنوي لأن قيمة p-value لها أقل من مستوى المعنوية 0.05، وباقي المتغيرات كانت غير معنوية.

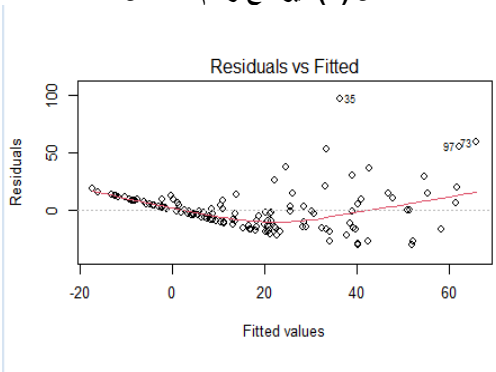
وتكون معادلة الانحدار المقدرة بطريقة المربعات الصغرى على النحو التالي:

$$\hat{Y} = 123.524 - 8.903X_1 + 7.813X_2 + 0.294X_3 + 0.537X_4 - 0.535X_5 + 1.146X_6$$

لاحظنا أنه باستخدام طريقة المربعات الصغرى الاعتيادية تحصلنا على تقديرات غير دقيقة لأنها لم تمثل الواقع، حيث أنه من خلال نتائج التحليل وجدنا أن بعض المتغيرات غير ذات أهمية، في حين أنها في الواقع مهمة، وهذا ما يثير الشك في اختلال أحد فروض الانحدار، عليه سنقوم بإجراء اختبارات التحقق من فروض تحليل الانحدار باستخدام طريقة المربعات الصغرى الاعتيادية.



الشكل (1): يوضح رسم الاحتمال



شكل (2): يوضح رسم البواقي مقابل

يشير تحليل الخطأ العشوائي إلى أن e_i يتبع توزيعًا طبيعيًا وفقًا للعلاقة $e_i \sim (0, \sigma^2)$، حيث يُظهر الشكل البياني (1) تجمع النقاط حول الخط مما يدل على أن البواقي تتبع توزيعًا طبيعيًا. فيما يتعلق باستقلالية الأخطاء العشوائية، فإن استخدام اختبار Durbin-Watson أظهر أن قيمة الاختبار بلغت DW=1.98≈2، مما يشير إلى غياب الارتباط الذاتي بين البواقي، كما أن قيمة p-value كانت 0.351، وهي قيمة مرتفعة تدل على عدم وجود ارتباط ذاتي بين البواقي. أما فيما يخص اختبار تجانس التباين، فقد أظهر اختبار Breusch-Pagan أن (prob> chi$^{2)}$ = 8.833749e-15، مما يدل على تحقق تجانس التباين، وهو ما أكده أيضًا رسم البواقي مقابل القيم المقدرة في الشكل (2). وبالنسبة لاستقلالية المتغيرات التفسيرية، فإن تحليل التباين المشترك بين المتغيرات $Cov(X_i, X_j) = 0,\ i \neq j$ يشير إلى ضرورة التحقق من وجود مشكلة التعدد الخطي. وقد تم الاعتماد على عدة اختبارات للكشف عن هذه المشكلة، مثل مصفوفة الارتباطات بين المتغيرات التفسيرية، معامل تضخم التباين(VIF) ، العدد الشرطي(CN) ، الجذور المميزة(ev) ، واختبار Farrar & Glauber. وقد تبين وجود مشكلة التعدد الخطي بين المتغيرات التفسيرية، وهي القضية التي سيتم دراستها ومعالجتها في هذا البحث باستخدام أسلوب انحدار الحافة.

جدول (2): مصفوفة الارتباط بين المتغيرات التفسيرية

Var	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1	0.97	-0.67	0.14	0.82	0.83
X_2		1	-0.62	0.097	0.77	0.74
X_3			1	-0.35	-0.86	-0.61
X_4				1	0.41	0.22
X_5					1	0.76
X_6						1

يوضح الجدول (2) وجود ارتباط قوي بين بعض المتغيرات التفسيرية، حيث بلغ معامل الارتباط بين المتغيرين (X_2, X_1)، قيمة عالية بلغت 0.97، كما لوحظ ارتباط موجب قوي بين المتغير X_1 وكل من المتغيرين X_5، X_6 بقيم 0.82 و0.83 على التوالي. يشير هذا الارتباط المرتفع إلى وجود تداخل خطي بين بعض المتغيرات التفسيرية. ويُستخدم مؤشر الجذور المميزة للكشف عن مشكلة التعدد الخطي وتحديد درجة التداخل الخطي بين المتغيرات. إذا تراوحت قيمة العدد الشرطي بين 5 و10، فإن ذلك يشير إلى ضعف التداخل، بينما تشير القيم بين 10 و30 إلى تداخل خطي متوسط إلى مرتفع، في حين أن القيم التي تتجاوز 30 تدل على وجود تداخل خطي كبير. في هذه الدراسة، تم حساب العدد الشرطي فكان 20.23، وهي قيمة تفوق 10 مما يؤكد وجود تداخل خطي بين المتغيرات التفسيرية. كذلك، يُستخدم معامل تضخم التباين (VIF) كأداة للكشف عن التداخل الخطي وتحديد المتغير المستقل المسؤول عن هذه المشكلة، وفقًا للمعادلة التالية:

$$VIF_j = \frac{1}{1-R_j^2} \qquad j = 1,2,\dots,p$$

تبين من خلال النتائج أن كلا من المتغيرين X_2, X_1 لهما معاملات تضخم تباين VIF>10 وهذا يعني أن هذين المتغيران هما المسئولان عن وجود التداخل الخطي. وعند تحليل الجذور المميزة لمصفوفة المعلومات $(X'X)$ بعد تحويلها إلى صيغة مصفوفة الارتباط، تبين أن أصغر قيمة لجذر مميز تبلغ 0.011، وهي قيمة قريبة جدًا من الصفر، مما يشير بوضوح إلى وجود مشكلة التداخل الخطي في النموذج. وباستخدام اختبار Farrar-Glauber تم حساب إحصاءة الاختبار χ^2 وفقا للعلاقة التالية:

$$\chi^2 = -\left[(n-1) - \frac{1}{6}(2k+7)\right] ln|D| = 1006.56$$

وبمقارنة قيمة الاختبار الحسابية بالقيمة الجدولية عند درجة حرية

$$\left(\frac{1}{2}p(p-1), \alpha\right)$$

$$\chi^2\left(\frac{1}{2}p(p-1), \alpha\right) = \chi^2\left(\frac{1}{2} * 6(6-1), 0.05\right)$$

$$= \chi^2(_{15,\,0.05}) = 25$$

نجد أن القيمة المحسوبة أكبر من القيمة الجدولية مما يعني وجود تعدد خطي بين المتغيرات التفسيرية.

7.3 خطوات تطبيق انحدار الحافة:

باستخدام لغة البرمجةR ، سيتم تطبيق انحدار الحافة من خلال مجموعة من الخطوات المنهجية التي تبدأ بتحميل بيانات الدراسة، حيث يتم استدعاء البيانات من برنامج Excel إلى برنامج R. بعد ذلك، يتم استخدام الدالة *glmnet()* لتقدير نموذج انحدار الحافة، مع الأخذ في الاعتبار أن هذه الطريقة تتطلب توحيد البيانات، وهو ما تقوم به الدالة *glmnet()* تلقائيًا. ومع ذلك، في حال تم إجراء التوحيد القياسي مسبقًا، يمكن تعطيل هذه الخاصية بإضافة الأمر *Standardize=F*.

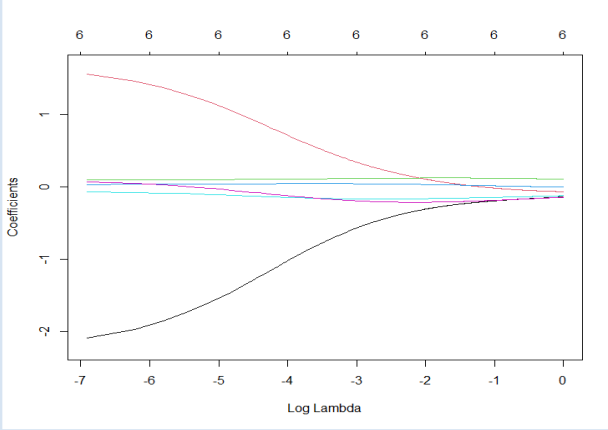
بعد تقدير النموذج الأولي، يتم اختيار القيمة المثلى لمعلمة التحيز k، وهي القيمة التي تحقق الحد الأدنى لمتوسط مربعات الأخطاء، وذلك باستخدام أسلوب التحقق المتقاطع (k-fold cross-validation) لضمان دقة النتائج. يمكن أيضًا استخدام مخطط أثر الحافة لمراقبة كيفية تغير مقدرات انحدار الحافة مع زيادة قيمة معلمة التحيز k ، مما يساعد في تفسير تأثير التحيز على النموذج. بمجرد تحديد القيمة المثلى لـ k ، يمكن تحليل النموذج النهائي واستخراج مقدرات انحدار الحافة التي تتيح فهماً أعمق للعلاقات بين المتغيرات التفسيرية والمتغير التابع.

7.4 تحديد قيمة معلمة التحيز

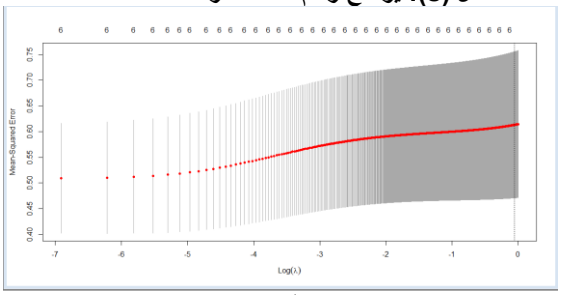
طريقة انحدار الحافة تعتمد بالدرجة الأولى على معلمة الحافة، وبما أنه يوجد العديد من الطرق لتقدير هذه المعلمة، فقد اعتمدنا في اختيارنا على طريقة التقاطع المتحقق (Cross Validation)، يتم اختيار قيمة k من بين القيم [0,1] بمقدار زيادة (0.001)، حيث أنه كل قيمة من قيم k تقابلها قيمة من قيم متوسط مربعات الأخطاء (MSE)، وقيم (VIF)، وأيضا باستخدام رسم أثر الحافة حيث تم رسم قيم مقدرات الحافة مقابل قيم k، نلاحظ من الشكل (3) أن المتغيران (X_1, X_2) أول متغيرين بدءا في

الاستقرار وهذا يؤكد أنهما أكثر المتغيرات تأثيرا على كمية سقوط الأمطار، أيضا من خلال رسم (CV) حيث تم رسم قيم k مقابل قيم MSE ، انظر الشكل (4).

من الشكل (3) نرى بوضوح أن المقدرات تبدأ في التحرك نحو الاستقرار كلما زادت قيم k بالتحديد عند الوصول إلى قيمة $k = 0.001$، وهذا مؤشر على تلاشي مشكلة التعدد الخطي، حيث تقترب العلاقة بين المتغيرات التفسيرية من التعامد ابتداءً من $(k \geq 0.001)$، وعلى ضوء هذه النتائج تم اختيار قيمة معلمة التحيز $k = 0.001$ هي القيمة المثالية التي يكون عندها متوسط مربعات الأخطاء أقل ما يمكن، انظر الشكل (4).



شكل (3): يوضح رسم مخطط أثر الحافة



شكل (4): يوضح التقاطع المتحقق (CV)

7.5 تحليل انحدار الحافة Ridge Regression Analysis

من خلال تطبيق انحدار الحافة، وذلك باعتماد قيمة معلمة الحافة (k=0.001)، تم الحصول على النتائج الموضحة بالجدول (3)

جدول (3): يوضح نتائج تحليل انحدار الحافة عند (k=0.001)

	Estimate	St. Error	VIF
X_1	-2.089	0.444	0.333
X_2	1.552	0.371	0.209
X_3	0.097	0.113	0.029
X_4	0.032	0.067	0.011
X_5	-0.065	0.124	0.063
X_6	0.070	0.115	0.034

استنادا على النتائج التي تم الحصول عليها في الجدول (3)، وباستخدام معلمة التحيز $k = 0.001$ تم توفيق النموذج التالي كمعادلة تقديرية لكمية سقوط الأمطار:

$$\hat{Y}_i = -2.089X_1 + 1.552X_2 + 0.097X_3 + 0.032X_4 - 0.065X_5 + 0.07X_6$$

من الجدول (3) نلاحظ أنه تم تقليص معلمات النموذج بعد تطبيق انحدار الحافة، كما نلاحظ انخفاض في عامل تضخم التباين ليصبح $(VIF < 10)$ لجميع المتغيرات التفسيرية، وهذا دليل على تلاشي مشكلة التعدد الخطي، ونتج عن ذلك تقليل تباين المعلمات مما أدى إلى الحصول على مجموع مربعات أخطاء أقل من طريقة المربعات الصغرى الاعتيادية.

7.6 المقارنة بين طريقتي OLS وRR:

بعد اجراء تحليل الانحدار باستخدام كلا من طريقة المربعات الصغرى الاعتيادية وطريقة انحدار الحافة تم تدوين النتائج المتحصل في الجدول (4)، وذلك لمقارنة أداء الطريقتين:

من الجدول (4) نلاحظ أنه باستخدام طريقة انحدار الحافة تم تقليص مقدرات النموذج مما أدى إلى التخلص من مشكلة التعدد الخطي، حيث انخفض عامل تضخم التباين ليصبح $(VIF < 10)$ لجميع المتغيرات التفسيرية، كذلك نلاحظ أن الخطأ المعياري لمقدرات الحافة كان أقل من الخطأ المعياري لمقدرات المربعات الصغرى الاعتيادية، وبالنظر لمعيار المقارنة (متوسط مربعات الخطأ) للطريقتين، نجد أن متوسط مربعات الاخطاء لمقدرات الحافة أقل من متوسط مربعات الاخطاء لمقدرات المربعات الصغرى الاعتيادية، عليه نجد أن طريقة انحدار الحافة هي الأفضل لتقدير نموذج الانحدار عندما تكون المتغيرات التفسيرية تحت تأثير مشكلة التعدد الخطي.

جدول (4): المقارنة بين نتائج (OLS) و (RR)

	OLS			RR		
	Coeff k=0	SE	VIF	Coeff k=0.001	SE	VIF
X_1	-8.903	1.58	43.66	-2.089	0.444	0.333
X_2	7.813	1.48	27.47	1.552	0.371	0.209
X_3	0.294	0.36	3.81	0.097	0.113	0.029
X_4	0.537	1.31	1.41	0.032	0.067	0.011
X_5	-0.535	1.76	8.21	-0.065	0.124	0.063
X_6	1.146	1.48	4.44	0.070	0.115	0.034
MSE	**18.54**			**0.5012**		
R^2	**0.55**			**0.55**		

8. الاستنتاجات:

بعد إجراء الدراسة وتحليل النتائج، تم التوصل إلى مجموعة من الاستنتاجات المهمة. أولًا، أثبتت طريقة انحدار الحافة فعاليتها في التخلص من مشكلة التعدد الخطي بين المتغيرات التفسيرية، حيث تبين أن عدم معنوية بعض المتغيرات التفسيرية في طريقة المربعات الصغرى الاعتيادية يعود بشكل رئيسي إلى وجود التعدد الخطي بينها. كما أن مقدرات المربعات الصغرى الاعتيادية تصبح غير دقيقة في ظل هذه المشكلة، نظرًا لتحيزها وارتفاع تبايناتها، مما يؤثر سلبًا على دقة التقدير.

من خلال تطبيق انحدار الحافة، تم تقليص معاملات الانحدار للحصول على مقدرات أكثر استقرارًا، مما يجعل النموذج الناتج باستخدام هذه الطريقة أكثر موثوقية للتنبؤ. وقد تبين أن القيمة المثلى لمعلمة التحيز k التي تحقق الحد الأدنى لمتوسط مربعات الأخطاء وتقلل من تضخم التباين لجميع المتغيرات التفسيرية هي $k = 0.001$. كذلك، أظهرت المقارنة باستخدام معيار متوسط مربعات الخطأ (MSE) أن تقدير النموذج باستخدام انحدار الحافة كان أكثر دقة من تقديره بطريقة المربعات الصغرى الاعتيادية، حيث كان متوسط مربعات الخطأ أقل في الحالة الأولى.

9. التوصيات:

بناءً على الاستنتاجات التي تم التوصل إليها من خلال دراسة مشكلة البحث وحلها، يتم التوصية بما يلي: **أولًا**، يُنصح باستخدام طريقة انحدار الحافة في تحليل الانحدار المتعدد عندما تكون المتغيرات التفسيرية تحت تأثير مشكلة التعدد الخطي، حيث أثبتت هذه الطريقة فعاليتها في معالجة تلك المشكلة. **ثانيًا**، يجب تجنب استخدام طريقة المربعات الصغرى الاعتيادية في مثل هذه الحالات، حيث أنها لا توفر حلاً لمشكلة التعدد الخطي وتؤدي إلى تقديرات غير دقيقة. **ثالثًا**، يوصى بالكشف عن وجود مشكلة التعدد الخطي باستخدام مجموعة متنوعة من الطرق، إذ أن بعض الأساليب قد لا تكون كافية للكشف عن المشكلة بشكل دقيق.

المراجع:

1. حمود أحمد جبار (2018). مقارنة بعض مقدرات انحدار الحرف الاعتيادية وانحدار الحرف البيزية مع تطبيق عملي. مجلة الاقتصاد والعلوم الإدارية 112(25)، 444-460، جامعة بغداد.

2. سلامي، أحمد، وحجاب، عيسى (2018). كيفية تقييم واختبار نماذج الانحدار في القياس الاقتصادي (دراسة تطبيقية). مجلة البديل الاقتصادي، المجلد الخامس، العدد الأول، الجزائر.
3. الكفيشي،سارة ماجد (2019). تقدير معلمات نموذجي انحدار لاسو وانحدار الحرف مع تطبيق عملي. رسالة ماجستير، جامعة كربلاء، كربلاء المقدسة، جمهورية العراق.
4. محمد، رواء صالح (2011). استخدام انحدار الحرف (Ridge) لدراسة أثر بعض العوامل على المؤشر العام لسوق الأوراق المالية. مجلة القادسية للعلوم الادارية والاقتصادية، المجلد 13، العدد 1، الجامعة المستنصرية.
5. Ahmed, Layla Aziz (2020). Using ridge regression to analysis the meteorological data in Sulaimani. Iraq Journal of Statistical Science (32), 76-86.
6. Bas, Eren, Egriooglu, Erol, & Uslu, Vedide, Rezan (2017). Shrinkage Parameters for Each Explanatory Variable Found Via Particle Swarm Optimization in Ridge Regression. Trends Comput Sci Inf Technol 2(1): 012-020. Dol: http://dx.doi.org/10.17352/tcsit.000005, Turkey.
7. Batah, Feras Sh. M., and Gore, Sharad Damodar (2009). Ridge Regression Estimator: Combining Unblased and Ordinary Ridge Regression Methods of Estimation. Article in Surveys in Mathematics and Applications, ISSN 1842-6298 (electronic), 1843-7265 (print), Vol. 4, 99-109.
8. Cropper, John Philip (1985). Tree-Ring Response Functions. An Evaluation by Means of Simulations [Dendrochronology Ridge Regression, Multicollinearity]. Text: Dissertation-Reproduction [electronic], http://hdl.handle.net/10150/187946, The University of Arizona.
9. El-Dereny, M., and Rashwan N. I. (2011). Solving Multicollinearity Problem Using Ridge Regression Models. Int. J. Contemp. Math. Sciences, Vol. 6, no. 12, 585-600, Tanta University, Egypt.
10. Duzan, Hanan, and Shariff, Nurul (2015). Simulation Study to Verify the Appropriate k Value for Ridge Regression in Two-variable Regression Model. Indian Journal of Science and Technology, Vol. 8(12), ISSN(Print): 0974-6846, ISSN(Online):0974-5645, Universiti Sains Islam Malaysia.
11. Fan, Chao, Rey, Sergio, J., and Myint, Soe, W. (2016). Spatially filtered ridge regression (SFRR): A regression framework to understanding impacts of land cover patterns on urban climate. DOI 10.1111/tgis.12240, Arizona State university.
12. Hoerl, E. Arthur, and Kennard, Ropert W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics, 12, 55-67.https://doi.org/10.1080/00401706.1970.10488634.
13. Hoerl, A.E., Kennard, R.W. & Baldwin, K.T. (1975). Ridge regression; Some simulation, Commun, Stat, 4, 105-123.
14. Khalaf, G. & Shukur, G. (2005). Choosing Ridge Parameter for Regression Problems. Communications in Statistics: Theory and Methods, 34, 1177-1182.
15. Alkhamisi, M. A., & Shukur, G. (2007). A Monte Carlo Study of Recent Ridge Parameters. Communications in Statistics-Simulation and Computation, 36: 535-547, ISSN:0361-0918, DOI: 10.1080/03610910701208619, Taylor &Francis Group.
16. Kibria, B. M. Golam, and Banik, Shipra (2016). Some Ridge Regression Estimators and Their Performance. Journal of Modern Applied Statistical Methods, Vol. 15, No. 1, 206-238.
17. Kuhn, M., & Johnson, K. (2022). Applied predictive modelling (2nd ed.). New York: Springer.
18. Lawless, J.F. & Wang, P. (1976). A simulation study of Ridge and other regression estimators, Commun. Stat. 5, 307-323.
19. McDonald, G. C., Galarneau, D. I. (1975). A Monte Carlo Evaluation of Some Ridge-Type Estimators. Journal of the American Statistical Association 70(350):407-416. DOI: 10.1080/01621459.1975.10479882.
20. Saleh, A. K., Arashi, Mohammad, & Kibria, B. M. Golam (2019). Theory of Ridge Regression Estimation with Applications. First Edition, John Wiley & Sons, USA.
21. Tibshirani, Ryan (2013). Modern Regression 1: Ridge Regression. Data Mining: 63-462/36-662.
22. Yahya, Waheed Babatunde, and Olaifa, Julius Babatunde (2014). A note on ridge regression modelling techniques. Electronic Journal of Applied Statistical Analysis, Vol. 07, Issue 02, 343-361, University of Ilorin, Ilorin, Nigeria.
23. Younker, James (2012). Ridge Estimation and its Modifications for Linear Regression with Deterministic or Stochastic Predictors. University of Ottawa, Ottawa, Canada.