رفع كفاءة معادلات التنبؤ بمعالجة الازدواج الخطي في الزيوت النباتية

د/ سحر عبد السلام إبراهيم

باحث بقسم بحوث التقييم الاقتصادي والبيئي– المعمل المركزي لبحوث التصميم والتحليل الإحصائي – مركز البحوث الزراعية

مقدمة:

تحليل الانحدار الخطي المتعدد هو طريقة إحصائية لديها الكثير من الافتراضات كغيره من التحليلات الإحصائية وعادة ما يتم التحقق من هذه الافتراضات قبل إجراء هذا التحليل فإذا أسقطت واحدة أو أكثر من هذه الافتراضات غالباً ما يتم معالجة هذه المشاكل بتطبيق أساليب إحصائية متقدمة علي البيانات ومن أبرز هذه المشكلات التي يتعرض لها الباحث عند استخدام تحليل الانحدار المتعدد مشكلة الازدواج الخطي، والذي يترتب عليه الكثير من الآثار مثل زيادة الخطأ القياسي وتحيز ثوابت المعادلة نتيجة زيادة التباين لهذه الثوابت فضلاً عن النتائج المضللة وتأثر أسلوب تحليل الانحدار المتعدد بالقيم الشاذة والمتطرفة وانخفاض درجة الثقة في نتيجة التحليل وما يترتب عليه من قرارات وتوصيات. ويعتبر أسلوب *Ridge Regression* من أهم الأساليب المستخدمة في معالجة الازدواج الخطي للوصول إلى أفضل تعبير قياسي للعلاقة الاقتصادية والذي يمكن أن يعطي حلاً منطقياً عند تحليل البيانات من خلال تلافي الآثار المترتبة علي تحليل البيانات باستخدام طريقة المربعات الصغرى العادية (*OLS*)، هذا ويمثل تقدير دالة استهلاك الزيوت النباتية في مصر أحد أهم التحديات التي تواجه الباحثين عند تقديرها نظراً لتوقع وجود مشكلة الازدواج الخطي بها، الأمر الذي يتطلب الاستعانة ببيانات أهم المتغيرات المؤثرة علي استهلاك الزيوت النباتية خلال الفترة (٢٠٠٠– ٢٠١٥) لتوضيح كيفية معالجة هذه المشكلة عند تحليل هذه البيانات، والذي يؤدي الخلل في تقديرها إلي مشاكل في التخطيط للإنتاج والاستهلاك والاستيراد والتصنيع تلبية احتياجات السوق المصري من هذه الزيوت النباتية.

تحتل مشكلة نقص إنتاج الزيوت النباتية في مصر أهمية قصوى، وتعد هدفاً هاماً وذلك لكونها غذاء هاما للإنسان وسلعة من السلع الغذائية الإستراتيجية وترجع الأهمية الغذائية للزيوت النباتية إلى أنها تحتوى على الفيتامينات الهامة الذائبة في الدهون، كما أنها مصدر رئيسي للطاقة وتنحصر مشكلة الزيوت النباتية في مصر في قلة الكميات المنتجة من الزيوت النباتية بالنسبة للكميات المطلوبة منه، إذ يعجز الإنتاج المحلى عن تغطية احتياجات الاستهلاك المحلى ومن ثم اتسعت الفجوة الغذائية.

مشكلة الدراسة:

تعتبر مشكلة الازدواج الخطي (*Multicollinearity*) من أهم مشاكل تحليل الانحدار المتعدد وما يترتب عليها من آثار سلبية علي نتيجة التحليل الإحصائي خاصة في ظل وجود عدد كبير من المتغيرات التفسيرية، وهو ما يمكن أن يتسبب في بناء نموذج قياسي للانحدار المتعدد لظاهرة ما مثل العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر وتقدير معالمه دون دراية أو دون التحقق من الفروض الأساسية الواجب توافرها عند تطبيقه والذي يقود إلي نتائج غير صحيحة تتسبب في قرارات خاطئة أو خطط لا تتناسب وطبيعة المشكلة أو الظاهرة التي يسعي الباحث لحلها.

أهداف الدراسة:

تستهدف الدراسة عرض ومناقشة مشكلة الازدواج الخطي من حيث طبيعتها وأسبابها وآثارها وطرق اكتشافها والبحث في أساليب معالجة البيانات التي تتضمن هذه المشكلة، لبيان حجم المشكلة التي يغفل عنها كثير من الباحثين الإحصائيين والاقتصاديين مع التركيز علي أسلوب انحدار *Ridge Regression* (*RR*) كأسلوب مميز وفعال لمعالجة هذه المشكلة ومن ثم مناقشة هذا الأسلوب وما يترتب عليه من نتائج لتحليل البيانات مع مقارنة هذه النتائج بأسلوب تحليل الانحدار المتعدد بطريقة المربعات الصغرى العادية (*OLS*)

في ضوء مميزات وعيوب كل طريقة، وذلك من خلال تطبيق عملي علي بيانات استهلاك الزيوت النباتية في مصر والعوامل المؤثرة عليها خلال الفترة (٢٠٠٠- ٢٠١٥) ووضع التفسيرات للاختلافات بينهما.

الطريقة البحثية ومصادر البيانات:

تستخدم الدراسة أسلوب تحليل الانحدار المتعدد بطريقة انحدار ريدج (*RR*) وطريقة المربعات الصغرى العادية (*OLS*) عند تحليل متغيرات الدراسة لتوضيح مشكلة الازدواج الخطي والمقارنة بينهما من خلال عدة معايير للمقارنة مثل معامل تضخم التباين (*VIF*) والدليل الشرطي (*CI*) وحدود التسامح (*Tolerance*) والجذور الكامنة (*Eigenvalues*) بجانب استخدام مصفوفة الارتباط البسيط بين المتغيرات التفسيرية موضع الدراسة، وكيفية معالجة ازدواج الخطي وذلك باستخدام حزم التحليل الإحصائي لبرنامج NCSS هذا إلي جانب توضيح الإطار النظري والرياضي للأساليب المستخدمة في التحليل الإحصائي وقد اعتمدت الدراسة علي بيانات الجهاز المركزي للتعبئة العامة والإحصاء خلال الفترة (٢٠٠٠- ٢٠١٥).

الازدواج الخطي[٧]:

إن أحد أهم الفروض الأساسية في تحليل الانحدار المتعدد يتمثل في عدم وجود علاقة خطية تامة بين المتغيرات المستقلة بمعني عدم تداخل أثنين أو أكثر من المتغيرات المستقلة، ويترتب علي إسقاط هذا الفرض حدوث مشكلة الازدواج الخطي (*Multicollinearity*)[٩] بحيث لا يمكن للتحليل الإحصائي أن يميز تماماً بينها أو عزل تأثيرها المستقل، ذلك لأن تفسير معادلة الانحدار الخطي يفترض أن متغيرات التنبؤ ليست مترابطة بقوة وبالتالي الازدواج الخطي يسقط هذا الفرض وهذا يعني عدم وجود علاقة خطية بين المتغيرات أي عدم استقلالية المتغيرات المستقلة، ومن ثم تعد مشكلة الازدواج الخطي من أهم المشاكل التي تواجه الباحث عند تحليل الانحدار الخطي المتعدد والتي يمكن أن تؤثر علي دقة تقديرات معاملات الانحدار.

اكتشاف الازدواج الخطي[٢]:

لا يمكن اكتشاف الازدواج الخطي قبل إجراء عملية التقدير فقد يوجد الازدواج الخطي وقد لا يوجد لذلك نفترض عدم وجوده في البداية، ونقوم بإجراء العلاقة الإحصائية بين المتغير التابع والمتغيرات المستقلة بالطريقة العادية ثم نختبر الفرض القائل بوجود الازدواج الخطي أو عدمه، حيث يتم الاعتماد في ذلك علي معامل تضخم البيانات (*Variance Inflation Factor* (*VIF*)، الدليل الشرطي (*Condition Index*)، والجذور الكامنة (*Eigenvalues*)، نسب التباين (*Variance Proportion*)، حدود التسامح (*Tolerance*).

أولاً: إيجاد مصفوفة الارتباط البسيط بين المتغيرات المستقلة:

١- يشير معامل الارتباط البسيط الذي يساوي صفر إلي عدم وجود ارتباط خطي بين المتغيرين المستقلين.
٢- يشير معامل الارتباط البسيط الذي يساوي واحد إلي وجود ارتباط خطي تام بين المتغيرين المستقلين.
٣- يشير معامل الارتباط البسيط الأعلى من ٠,٩٥ إلي وجود ارتباط خطي قوي بين المتغيرين المستقلين.
٤- يشير معامل الارتباط البسيط الذي يتراوح بين (٠,٩٠- ٠,٩٥) إلي احتمال وجود ارتباط خطي بين المتغيرين المستقلين.
٥- يشير معامل الارتباط البسيط الأعلى من ٠,٥٠ إلي وجود ارتباط خطي متوسط بين المتغيرين المستقلين.
٦- يكون هناك شك في وجود ازدواج خطي إذا كانت قيمة R^2 عالية بحيث تقارب ٠,٩٥ أو أكبر بينما مقدرات الانحدار الجزئية غير معنوية إحصائياً نظراً لتدني وانخفاض قيم t فإذا كانت قيمة R^2 عالية فإن قيمة (F) المحسوبة ستكون عالية لأنها تعتمد على R^2، وبالتالي سوف يكون هناك تناقض بين قيمة F والتي توضح لنا أن النموذج معنوي وقيمة t والتي توضح لنا أن معاملات الانحدار غير معنوية،

وبالتالي نستطيع أن نجزم فعلاً بوجود ارتباط خطي بين المتغيرات علماً بان وجود الازدواج لا يعتبر مشكلة وإنما المشكلة تتمثل في درجة الازدواج الخطي.

ثانياً: حساب معامل تضخم التباين (*VIF*) *Variance Inflation Factor*:

$$V.I.F\left(\hat{\beta}_j\right)=\frac{1}{1-R_j^2}$$

تدل قيمة معامل تضخم التباين[١٥] (*VIF*) *Variance Inflation Factor*" عندما تكون أكبر من ١٠ علي وجود الازدواج الخطي بين المتغيرات المستقلة موضع الدراسة ويحسب معامل تضخم التباين كما يلي:

$$V.I.F\left(\hat{\beta}_j\right)=\frac{1}{1-R_j^2}=\frac{1}{tolerance}$$

R^2 = معامل التحديد للمتغير X المستخرج من انحدار Yعلي المتغيرات المستقلة حيث $j=1,2,.....k$

يستخدم لقياس درجة التباين كمؤشر للدقة علي زيادة أو انخفاض التباين وذلك إذا زادت قوة الارتباط بين مجموعة المتغيرات المستقلة فإن R^2 ترتفع وبالتالي ينخفض المقام ويرتفع معامل تضخم التباين ويرتفع التباين وبالتالي تقل دقة القياس، فإذا كانت قيمة النموذج (*VIF*) أكبر من ١٠ دل ذلك علي وجود مشكلة الازدواج الخطي[٧].

ثالثاً: الدليل الشرطي *CI* (*Condition Index*):

هو عبارة عن الجذر التربيعي لحاصل قسمة أكبر جذر كامن (*Eigenvalue*) علي كل جذر كامن مقابل للإبعاد $D_J, J=1,2,.....8$ ويحسب هذا الدليل لأكبر جذر كامن مثلاً وفق العلاقة:

$$K \equiv \frac{Maximum\ eigenvalue}{Minimum\ eigenvalue}$$

$$CI \equiv \sqrt{\frac{Maximum\ eigenvalue}{Minimum\ eigenvalue}}=\sqrt{K}$$

لإيجاد الدليل الشرطي لابد أولاً من احتساب الجذور الكامنة (*Eigenvalues*) حيث توضح كمية الاختلافات الكلية بين المتغيرات، فعندما تكون الجذور المميزة مساوية للصفر فإنه يدل على التعدد الخطي التام أما إذا كانت قريبة من الصفر فهذا مؤشر على وجود تعدد خطي عالي، أما إذا كانت مساوية إلى (*1*) فتعتبر الحالة الأمثل في عدم وجود مشكلة التعدد الخطي ومن ثم يمكن إيجاد العدد الشرطي(K).

إيجاد الدليل الشرطي *CI* (*Condition Index*) ونسب التباين (*Variance Proportion*) حيث يؤخذان سوياً ويستفاد منهما في بيان درجة التعدد الخطي والمتغيرات المرتبطة مع بعضها البعض فإذا كانت قيمة الدليل الشرطي في حدود القيمة (٥– ١٠) وأن اثنين أو أكثر من نسب التباين أقل من ٠,٥ فهذا يدل علي أن الارتباط ضعيف، إما إذا كانت القيمة تقع بين (١٠– ٣٠) $10 \le CI \le 30$ فهذا يعني أن هناك ازدواج خطي من المعتدل إلي العالي أما إذا تجاوزت القيمة عن (٣٠) فهذا مؤشر علي أن الازدواج الخطي بدرجة أكبر وعلي وجود المشكلة وخطورتها.

رابعاً: حدود التسامح[٨] (*Tolerance*):

يعتبر هذا المقياس أحد المقاييس الدالة علي وجود أو عدم وجود مشكلة الازدواج الخطي، فإذا كانت قيمة *Tolerance* أكبر من ٠,١ دل ذلك علي عدم وجود مشكلة الازدواج الخطي، أما إذا كانت قيمة *Tolerance* أقل من ٠,١ دل ذلك علي وجود مشكلة الازدواج الخطي:

$$Tolerance = \frac{1}{VIF}$$

طرق علاج مشكلة الازدواج الخطي[١٤]:

وجود الازدواج الخطي لا يعني مشكلة وإنما المشكلة تتمثل في درجة هذا الازدواج فإذا كانت درجة الازدواج الخطي منخفضة فمن الممكن قبول هذا الازدواج، أما إذا كانت درجة الازدواج الخطي مرتفعة فيجب العمل علي معالجة هذا الازدواج بواحدة أو أكثر من الطرق الآتية:

١- جمع بيانات إضافية: كلما كبر حجم العينة عن طريق إضافة بيانات جديدة كلما ساعد ذلك علي تخفيض حجم التباينات وهذا يقلل أثر الارتباط الخطي فإنه لا ينصح في البحوث القياسية أن لا يقل حجم العينة عن ٢٥ مشاهدة وألا يزيد عدد المتغيرات المستقلة عن خمسة.

٢- الاستعانة بمعلومات خارجية: إذا كان هناك تقدير لمعلمة أحد المتغيرات الذي يتصف بكونه مرتبطاً ارتباطاً متعدداً فيمكن استخدام هذا التقدير الذي تم خارج إطار البحث مع نتائج دراسة البحث قيد الدراسة.

٣- تحويل العلاقات الدالية: يتم ذلك عن طريق استخدام الأدوات والمفاهيم الرياضية كأن يكون المتغير X_2 موجود في المقام فتضرب المعادلة في YX_2 وبهذا نحصل علي علاقة دالية جديدة، والخطورة في هذا الإجراء هو وجوب ملاحظة النتائج عند تحليلها وتفسيرها ومدي تطابقها مع منطوق النظرية الاقتصادية.

٤- حذف أو إضافة متغير جديد: قد يلجأ البعض لاستخدام الأسلوب القياسي لحذف المتغير الذي يمتاز بالارتباط العالي مع المتغيرات المستقلة الأخرى أو قد يضاف متغير جديد أخر ذو أهمية بالنسبة للنموذج.

٥- تعويض المتغيرات المستقلة ذات فترات الإبطاء الزمني بدالة للمتغير التابع.[١٣]

طريقة انحدار ريدج *RR* (*Ridge Regression*):

تتمثل أحد الطرق المستخدمة في التخلص من الازدواج الخطي في إزالة المتغير التفسيري الذي يتسبب في إحداث الازدواج الخطي إلا أنه في ظل وجود طرق جديدة يمكنها إزالة الازدواج الخطي دون حذف أي من المتغيرات موضع الدراسة، وأهم هذه الطرق طريقة انحدار ريدج *RR* (*Ridge Regression*) وهي عبارة عن طريقة لمعالجة مشكلة الازدواج الخطي للنموذج العام *GLM* (*General Linear Model*) وتتلخص هذه الطريقة في إضافة قيمة صغيرة موجبة تقع قيمتها بين الصفر وأقل من الواحد إلى العناصر القطرية لمصفوفة المعلومات ($\chi'\chi$) للحصول على مقدرات أكثر دقة، حيث تعمل هذه الطريقة على فك الارتباطات بين المتغيرات التوضيحية وتستخدم الصيغة الآتية في إيجاد تقديرات قيم (B) باستخدام طريقة انحدار ريدج على أن يتم تحويل المتغير التابع والمتغيرات التفسيرية إلى صيغة معيارية:

$$\hat{\beta}_R = \left(X' X + CI_P\right)^{-1} X' Y$$

تعبر I عن مصفوفة الوحدة (*Identity Matrix*) وعندما تكون قيمة $C = CI = 0$ فإن تقديرات معلمات طريقة انحدار ريدج تساوي تقديرات معلمات طريقة المربعات الصغرى العادية وعندما تكون ($C > 0$) فإن مقدرات انحدار ريدج تميل إلى الاستقرار عند قيمة معينة نسبة للتغيرات في البيانات ولكنها تكون متحيزة.

تحديد أفضل قيمة *K* في طريقة انحدار *Ridge*:

يتم تحديد أفضل قيمة K من خلال طريقة التكرار التي تستخدم لتقدير عدة معالم للنموذج الاحصائي مع كل تكرار لقيمة K بحيث تتحدد قيمة K في النهاية وفقاً للمعادلة التالية:

$$K = \frac{P\hat{\sigma}^2}{\sum_{i=1}^{P} \hat{\alpha}_{\circ}^{\,2}}$$

تعبر $\hat{\alpha}^2$ عن متوسط مربع الخطأ وهي من القيم المتحصل عليها من طريقة المربعات الصغرى العادية بينما تعبر P عن عدد المتغيرات المستقلة، وتتراوح قيمة K بين صفر وأقل من الواحد فإذا كانت $K=0$ فإن معامل انحدار طريقة (OLS) يساوي معامل انحدار طريقة (RR) وعندما تكون قيمة K أكبر من الصفر نحصل علي قيم مقدرة متحيزة إلا أنها أكثر استقراراً من القيم المقدرة بطريقة (OLS) وبذلك تعكس قيمة K مقدار التحيز في القيم المقدرة، وبصفة عامة يعاب علي طريقة (RR) صعوبة تحديد قيمة K المثلي ولتحديد قيمة K التي تعطي أفضل نموذج يستخدم عادة الرسم البياني لقيم معاملات انحدار (RR) علي المحور الرأسي مع قيم مختلفة لثابت التحيز K علي المحور الأفقي ويعرف الشكل حينئذ *Ridge trace* ويؤخذ في الاعتبار قيمة (VIF) للتأكد من حل مشكلة الارتباط الخطي فإذا أظهر الشكل استقراراً في قيم معاملات الانحدار وانخفاض قيم (VIF) عند قيمة محددة لثابت التحيز K يتم اختيار أحد النماذج المناظرة لها بصورة تحكمية.

العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠– ٢٠١٥):

يشير جدول (١) إلي أهم المتغيرات التي تؤثر علي كمية استهلاك الزيوت النباتية في مصر وهي الإنتاج المحلي من الزيوت بالألف طن (X_1)، واردات الزيوت بالألف طن (X_2)، متوسط سعر استيراد الزيوت دولار/ طن (X_3)، الدخل الفردي الحقيقي جنيه/ سنة (X_4)، متوسط سعر التجزئة الحقيقي جنيه/ طن (X_5)، سعر صرف الدولار مقابل الجنيه (X_6)، عدد السكان مليون نسمة (X_7)، قيمة الدعم الحقيقي للزيوت النباتية مليون جنيه (X_8).

جدول (١) العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠– ٢٠١٥)

السنوات	الاستهلاك ألف طن (Y)	الإنتاج المحلي من الزيوت النباتية ألف طن (X_1)	واردات الزيوت النباتية بالألف طن (X_2)	متوسط سعر استيراد الزيوت النباتية دولار/ طن (X_3)	الدخل الفردي الحقيقي جنيه/ سنة (X_4)	سعر التجزئة الحقيقي جنيه/ طن (X_5)	سعر صرف الدولار مقابل الجنيه (X_6)	عدد السكان مليون نسمة (X_7)	قيمة الدعم الحقيقي للزيوت النباتية بالمليون جنيه (X_8)
٢٠٠٠	٧٤٨	١١٥	٦٤٨	٧٢١,٧	٤٦٥٤	٣٧٠٥	٣,٥٠	٦٣,٩٧٤	٦٦٣
٢٠٠١	٥٧٤	١٢١	٤٧١	٣٤٥,٤	٥٠٣٢	٤٠٧٩	٤,٠٠	٦٥,٢٩٧	٥٨٦
٢٠٠٢	٤٩٤	١٦٩	٣٤٨	٤٧٩,٧	٥٠٦٤	٣٩٧٨	٥,٩٧	٦٦,٦٢٧	٤٩٠
٢٠٠٣	٥٠٨	١٣٥	٢٦٥	٥٤٥,٣	٥٧٣٤	٤١٨٤	٦,٢٧	٦٧,٩٦٤	٥٥٨
٢٠٠٤	١٠٠٧	١٤١	٨٧٧	٤٠٦,٦	٦٤١٦	٤٣٨٩	٥,٧٩	٦٩,٣٠٣	٤٧٨
٢٠٠٥	١٢٤٨	٢٠٤	١٠٧١	٤١٩,٥	٦٨٥٩	٤١٥٤	٥,٧٩	٧٠,٦٥٣	٤٢٢
٢٠٠٦	١٣٨٩	٢٠٧	١٢٠٧	٣٩٥,٠	٧٥٤٣	٤٣٢٦	٥,٧٥	٧٢,٠٠٨	٥٩٣
٢٠٠٧	٧٨٦	٢٥١	٥٨٠	٦٥٠,١	٨١٧٧	٤٥١٠	٥,٧٩	٧٣,٦٤٤	٤٩٦
٢٠٠٨	٧٣٦	١٧٥	٥١٤	٢١٠١,٥	٧٠٠٤	٤٢٩٣	٥,٤٥	٧٥,١٩٤	٤١٨
٢٠٠٩	٧١٤	١٧٦	٥٠٤	١٦٩٦,٨	٨٣٣٣	٥٠٣٤	٥,٧٦	٧٦,٩٢٥	٣٦٦
٢٠١٠	٦٧١	١٧٤	٤٧١	٢١٧٨,٧	٨٥٥٠	٥٠٩٠	٥,٦٧	٧٨,٦٨٥	٤٨٤
٢٠١١	١٢٧٠	٢٠٠	١٢٢١٢	١٦٩٨,٧	٧٩٨٣	٣٣١٦	٥,٩٧	٨٠,٥٣٠	٢٣١٤
٢٠١٢	١٢٠٧	١٧٦	١١٩٨	١٥٦٨,٣	٩٢٤٧	٣٣٢٤	٦,١٠	٨٢,٣٠٥	٢٤٣٤
٢٠١٣	١٥٣٠	٢٢٨	١٥٣١	١٠٥١,٠	٩٤٨٢	٣٣٠٧	٦,٨٨	٨٤,٦٢٨	١٩٧٥
٢٠١٤	١٣٢٤	١٩٩	٦٤٠	١٦٤٣,٧	١١٩٩٤	٢٢٨٩	٧,٨٠	٨٦,٨١٤	٢٥٣٨
٢٠١٥	١٤٥٠	٢١١	٧٢٠	١٧٥٨,٢	١٣٥٢٤	٢٩٥٠	٨,٩٠	٩٢,٥٨٢	٢٤٥٦

تم تقدير القيم الحقيقية باستخدام الرقم القياسي لأسعار الجملة لسنة الأساس ٢٠٠٠ = ١٠٠

المصدر: جمع وحسب من:

١– الجهاز المركزي للتعبئة العامة والإحصاء، المركز القومي للمعلومات، أعداد متفرقة.

٢– الجهاز المركزي للتعبئة العامة والإحصاء، الكتاب السنوي، أعداد متفرقة.

٣– الجهاز المركزي للتعبئة العامة والإحصاء، النشرة السنوية للتجارة الخارجية، أعداد متفرقة.

٤– الجهاز المركزي للتعبئة العامة والإحصاء، النشرة السنوية لأسعار المواد والمنتجات الغذائية والخدمات، أعداد متفرقة.

يوضح جدول (٢) معادلات الاتجاه الزمني العام لتطور العوامل المؤثرة علي استهلاك الزيوت في مصر خلال الفترة (٢٠٠٠- ٢٠١٥)، حيث تبين معادلة (١) أن كمية الاستهلاك من الزيوت النباتية بالألف طن تتزايد سنوياً بنحو ٥١,٢٥ ألف طن، وهذا التزايد معنوي إحصائياً عند مستوى معنوية ١%، وتعبر قيمة معامل الارتباط البالغة نحو ٠,٦٨ عن وجود ارتباط معنوي موجب بين كمية الاستهلاك من الزيوت النباتية وعامل الزمن، كما يتبين أن عامل الزمن يفسر نحو ٤٦% من إجمالي التغيرات في استهلاك الزيوت النباتية خلال فترة الدراسة.

تبين معادلة (٢) أن الإنتاج المحلي من الزيوت النباتية بالألف طن تتزايد سنوياً بنحو ٥,٢٤ ألف طن، وهذا التزايد معنوي إحصائياً عند مستوى معنوية ١%، وتعبر قيمة معامل الارتباط البالغة نحو ٠,٦٥ عن وجود ارتباط معنوي موجب بين الإنتاج المحلي من الزيوت النباتية بالألف طن وعامل الزمن، كما يتبين أن عامل الزمن يفسر نحو٤٣%من إجمالي التغيرات في الإنتاج المحلي من الزيوت النباتية خلال فترة الدراسة.

تبين معادلة (٤) أن متوسط سعر استيراد الزيوت دولار/ طن تتزايد سنوياً بنحو ١١١,٢ دولار/ طن، وهذا التزايد معنوي إحصائياً عند مستوى معنوية١%، وتعبر قيمة معامل الارتباط البالغة نحو ٠,٧٦ عن وجود ارتباط معنوي موجب بين متوسط سعر استيراد الزيوت وعامل الزمن، كما يتبين أن عامل الزمن يفسر نحو ٥٨% من إجمالي متوسط سعر استيراد الزيوت دولار/ طن خلال فترة الدراسة.

تشير المعادلة (٥) إلي أن الدخل الفردي الحقيقي جنيه/ سنة يتزايد سنوياً بنحو ٤٧٧,٠٥ جنيه/ سنة ويعتبر هذا التزايد معنوي إحصائياً ويفسر عامل الزمن نحو ٨٧% من إجمالي التغيرات في الدخل الفردي الحقيقي خلال فترة الدراسة.

جدول (٢) معادلات الاتجاه الزمني العام لتطور العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠- ٢٠١٥)

رقم المعادلة	المؤشر	المعادلة	R	R^2	المعنوية
١	كمية الاستهلاك من الزيوت النباتية بالألف طن	$\hat{Y}_i = \underset{(3.75)^{**}}{542.9} + \underset{(3.43)^{**}}{51.25}\, X_i$	٠,٦٨	٠,٤٦	**
٢	الإنتاج المحلي من الزيوت النباتية بالألف طن	$\hat{Y}_i = \underset{(8.66)^{**}}{135.55} + \underset{(3.24)^{**}}{5.24}\, X_i$	٠,٦٥	٠,٤٣	**
٣	واردات الزيوت النباتية بالألف طن	$\hat{Y}_i = \underset{(2.68)^{-}}{493.1} + \underset{(1.69)^{-}}{32.12}\, X_i$	٠,٤١	٠,١٧	-
٤	متوسط سعر استيراد الزيوت دولار/ طن	$\hat{Y}_i = \underset{(0.58)^{-}}{139.84} + \underset{(4.44)^{**}}{111.2}\, X_i$	٠,٧٦	٠,٥٨	**
٥	الدخل الفردي الحقيقي جنيه/ سنة	$\hat{Y}_i = \underset{(7.92)^{**}}{3794.85} + \underset{(9.62)^{**}}{477.05}\, X_i$	٠,٩٣	٠,٨٧	**
٦	متوسط سعر التجزئة الحقيقي جنيه/ طن	$\hat{Y}_i = \underset{(12.79)^{**}}{4566} - \underset{(-2.02)^{-}}{74.47}\, X_i$	٠,٤٧	٠,٢٣	-
٧	سعر صرف الدولار مقابل الجنيه	$\hat{Y}_i = \underset{(10.09)^{**}}{3.95} + \underset{(5.62)^{**}}{0.23}\, X_i$	٠,٨٣	٠,٦٩	**
٨	عدد السكان مليون نسمة	$\hat{Y}_i = \underset{(80.93)^{**}}{60.81} + \underset{(22.16)^{**}}{1.72}\, X_i$	٠,٩٩	٠,٩٧	**
٩	قيمة الدعم الحقيقي للزيوت النباتية مليون جنيه	$\hat{Y}_i = \underset{(-0.47)^{-}}{-144.23} + \underset{(4.51)^{**}}{143.96}\, X_i$	٠,٧٧	٠,٥٩	**

(**) معنوية عند مستوي ١% (-) غير معنوية

المصدر: حسب من جدول (١).

توضح معادلة (٧) أن سعر صرف الدولار مقابل الجنيه يتزايد سنوياً زيادة معنوية إحصائياً بنحو ٠,٢٣ دولار/ جنيه حيث يفسر عامل الزمن نحو ٦٩% من إجمالي التغيرات في سعر صرف الدولار مقابل الجنيه خلال فترة الدراسة.

تبين معادلة (٨) أن عدد السكان بالمليون نسمة تتزايد سنوياً بنحو ١,٧٢ مليون نسمة، وهذا التزايد معنوي إحصائياً عند مستوى معنوية ١%، وتعبر قيمة معامل الارتباط البالغة نحو ٠,٩٩ عن وجود ارتباط معنوي موجب بين عدد السكان بالمليون نسمة وعامل الزمن، كما يتبين أن عامل الزمن يفسر نحو٩٧% من عدد السكان بالمليون خلال فترة الدراسة.

تشير معادلة (٩) أن قيمة الدعم الحقيقي للزيوت النباتية بالمليون جنيه يتزايد سنوياً بنحو ١٤٣,٩٦ مليون جنيه، وهذا التزايد معنوي إحصائياً عند مستوى معنوية ١%، وتعبر قيمة معامل الارتباط البالغة نحو ٠,٧٧ عن وجود ارتباط معنوي موجب بين قيمة الدعم الحقيقي للزيوت النباتية مليون جنيه وعامل الزمن، كما يتبين أن عامل الزمن يفسر نحو ٥٩% من قيمة الدعم الحقيقي للزيوت النباتية مليون جنيه خلال فترة الدراسة.

دراسة العلاقة بين العوامل التي تؤثر علي استهلاك الزيوت النباتية في مصر:

نظراً لما صاحب عملية تقدير أهم العوامل التي تؤثر علي استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠– ٢٠١٥) من ظهور مشكلة الازدواج الخطي بين العوامل من خلال الاعتماد علي معامل تضخم البيانات (*VIF*)، الدليل الشرطي (*Condition Index*)، نسب التباين (*Variance Proportion*) ومن ثم تم استخدام *RR* في تقدير المعلمات باعتبارها أحد طرق معالجة تعدد العلاقات الخطية وذلك من أجل الوصول إلي مقدرات معبرة بشكل أكبر من أثر المتغيرات المؤثرة علي استهلاك الزيوت النباتية في مصر وقد تم استخدام البيانات الأصلية وتوظيف برنامج (*NCSS*) في الحصول علي النتائج.

أولاً: مصفوفة الارتباط البسيط بين المتغيرات موضع الدراسة:

يوضح جدول (٣) تقدير مصفوفة الارتباط البسيط لأهم المتغيرات المستقلة التي تؤثر علي استهلاك الزيوت في مصر وأوضحت النتائج أنه تبين وجود ارتباط بين كمية الاستهلاك من الزيوت (Y) والإنتاج المحلي من الزيوت بالألف طن بنحو (X_1) بنحو ٠,٦٠ بينما كانت قيمة الارتباط بين (X_1) و(X_4) تقدر بنحو ٠,٦٣ بينما تقدر قيمة معامل الارتباط بين (X_1) و (X_8) بنحو ٠,٦١ وبذلك يكون الارتباط بين X_1 وكلاً من X_4، X_8 أعلي من الارتباط بين X_1 وY.

تبين وجود ارتباط غير معنوي بين كمية الاستهلاك من الزيوت (Y) متوسط سعر استيراد الزيوت دولار/ طن (X_3) بينما يوجد ارتباط بين (X_3) و (X_7) تقدر بنحو ٠,٧٣ أي يوجد ارتباط معنوي، كما تبين وجود ارتباط بين كمية الاستهلاك من الزيوت (Y) الدخل الفردي الحقيقي (*X4*) بنحو ٠,٦٨ بينما كانت قيمة الارتباط بين (X_4) و (X_6) و (X_7) و (X_8) تقدر بنحو ٠,٩١ ، ٠,٩٦ ، ٠,٧٤ علي التوالي أي يوجد ارتباط قوي جداً بين (X_4) و (X_6) و(X_7) وأقوي من ارتباط (X_4) بالمتغير (Y).

أظهرت النتائج أيضاً وجود ارتباط بين كمية الاستهلاك من الزيوت (Y) متوسط سعر التجزئة الحقيقي (X_5) بنحو –٠,٦٠ بينما كانت قيمة الارتباط بين (X_5) و(X_8) تقدر بنحو –٠,٨٧ أي يوجد ارتباط قوي سالب بين (X_5) و (X_8) وأقوي من ارتباط (X_5) بالمتغير (Y).

كما أظهرت النتائج وجود ارتباط بين كمية الاستهلاك من الزيوت (Y) سعر صرف الدولار (X_6) بنحو ٠,٦٨ بينما كانت قيمة الارتباط بين (X_6) و (X_4) و (X_7) أقوي من الارتباط بين (Y) و (X_6) حيث تقدر بنحو ٠,٩١ ، ٠,٨٦ علي التوالي أي يوجد ارتباط قوي جداً بين (X_6) و (X_4) و (X_7)، وتظهر علاقة الارتباط بين (Y) وعدد السكان بالمليون نسمة (X_7) حيث تقدر بنحو ٠,٦٨ بينما تقدر علاقة الارتباط بين (X_7) وكلاً من (X_4) و (X_6) و (X_8) بنحو ٠,٩٦، ٠,٨٦، ٠,٨١ علي التوالي وتكون قوة العلاقة (X_7) مع (Y) أقل من هذه المتغيرات الثلاثة السابقة.

جدول (٣) مصفوفة الارتباط البسيط لأهم المتغيرات الكمية التي تؤثر علي كمية استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠- ٢٠١٥)

المتغيرات	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Y	١								
X_1	٠,٦٠	١							
X_2	٠,٨٤	٠,٤٦	١						
X_3	٠,١٧	٠,٢٦	٠,١٢-	١					
X_4	٠,٦٨	٠,٦٣	٠,٢٨	٠,٦٢	١				
X_5	٠,٦١-	٠,١٩-	٠,٣٥-	٠,١٣-	٠,٥٣-	١			
X_6	٠,٦٨	٠,٥٨	٠,٢٨	٠,٤٦	٠,٩١	٠,٤٩-	١		
X_7	٠,٦٨	٠,٦١	٠,٣٧	٠,٧٣	٠,٩٦	٠,٥٤-	٠,٨٧	١	
X_8	٠,٧٠	٠,٣٦	٠,٤٦	٠,٤٥	٠,٧٧	٠,٨٧-	٠,٦٨	٠,٨١	١

المصدر: **جمعت وحسبت من نتائج تحليل جدول (١).**

تبين أيضاً وجود ارتباط بين كمية الاستهلاك من الزيوت (Y) قيمة الدعم الحقيقي للزيوت النباتية (X8) بنحو ٠,٧٠ بينما كانت قيمة الارتباط بين (X_8) و (X_4) ، (X_5) ، (X_7) وتقدر بنحو ٠,٧٧ ، -٠,٨٧ ، ٠,٨١ علي التوالي أي يوجد ارتباط عالي بين (X_8) و (X_4) و (X_5) و (X_7) أقوي من ارتباط (X_8) مع (Y) وعليه ظهرت مشكلة الازدواج الخطي.

ثانياً: معامل تضخم التباين (*VIF*):

يوضح جدول (٣) تقدير معامل تضخم التباين (VIF) والذي يتضح أن قيمته قد زادت عن ١٠ لكل من المتغيرات (X_3) = ١٠,٣٦ ، (X_4) = ٣٥,٩٦ ، (X_7) = ١٠٥,٧٨ ، (X_8) = ٢١,٥٨ وهو ما يشير إلي وجود مشكلة الازدواج الخطي بين هذه المتغيرات التفسيرية، أما بقية المتغيرات فقد انخفضت قيمة (VIF) عن ١٠ وهي (X_1) = ٢,٩٣٣ ، (X_2) = ٢,٩٥٥ ، (X_5) = ٨,١٦٧ ، (X_6) = ٨,٤٣٤ وهو ما يشير إلي أن هذه المتغيرات التفسيرية ليست المتسببة في مشكلة الازدواج الخطي.

ثالثاً: الدليل الشرطي (*CI*):

يوضح جدول (٣) تقدير الدليل الشرطي (CI) ومنه يتبين أن قيمة (CI) أقل من ٥ للمتغير (X_1) = ٤,٥٥ وهو ما يشير إلي عدم وجود مشكلة الازدواج الخطي في هذا المتغير، أما قيمة (CI) للمتغير (X_2) = ٥,٨٣ وهي بذلك محصورة بين (٥- ١٠) مما يدل علي وجود ازدواج خطي ضعيف، وبالنسبة للمتغيرات التفسيرية (X3) و (X_4) و (X_5) فقد تراوحت قيمة (CI) لهذه المتغيرات بين (١٠- ٣٠) حيث قدرت بنحو ١٠,٠٩٨ ، ١٥,٩٦٥ ، ٢٥,٠٢٣ علي الترتيب مما يشير إلي وجود ازدواج خطي معتدل إلي عالي، في حين أن قيمة (CI) للمتغيرات (X_6) و (X_7) و (X_8) زادت عن ٣٠ وهو ما يشير إلي وجود ازدواج خطي بدرجة كبيرة لهذه المتغيرات حيث بلغت قيمة (CI) لهذه المتغيرات نحو ٤٨,٢٦٩ ، ٦٣,٣٥٣ ، ٣٣٠,٣١٥ علي الترتيب.

تشير قيم الجذور الكامنة (*Eigenvalues*) في جدول (٣) إلي أن قيمة هذه الجذور اقتربت من الصفر لكل من المتغيرات التفسيرية (X_6) و (X_7) و (X_8) مما يشير إلي وجود ازدواج خطي عالي لهذه المتغيرات دون غيرها من المتغيرات موضع الدراسة، في أن بقية المتغيرات التفسيرية كانت قيم الجذور الكامنة (*Eigenvalues*) المقدرة لها بعيدة نسبياً عن الصفر مما يشير إلي أن هذه المتغيرات لا تتضمن مشكلة الازدواج الخطي.

رابعاً: حدود التسامح (*Tolerance*):

يبين جدول (٤) تقدير قيمة حدود التسامح (*Tolerance*) في قيم المتغيرات التي تؤثر علي استهلاك الزيوت تبين أن هناك أكثر من متغير تظهر فيها مشكلة الازدواج الخطي وهي كلاً من المتغيرات X3 ،

X4 ، X7 ، X8 حيث قدرت قيمته بنحو ٠,٠٨ ، ٠,٠٢ ، ٠,٠٠٨ ، ٠,٠٤ علي التوالي وكل هذه القيم أقل من ٠,١ وهو دليل علي وجود مشكلة الازدواج الخطي.

تحليل الانحدار المتعدد باستخدام طريقة (١٠)(*OLS*):

بإجراء تحليل الانحدار الخطي المتعدد بطريقة المربعات الصغرى (*OLS*) لانحدار المتغيرات التفسيرية موضع الدراسة علي كمية استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠– ٢٠١٥) تبين أن المعادلة كانت علي النحو التالي:

$$\hat{Y}_I = \underset{(2.81)^{*}}{2500} - \underset{(-0.33)^{-}}{0.232}X_1 + \underset{(11.54)^{**}}{0.84}X_2 + \underset{(1.43)^{-}}{0.10}X_3 + \underset{(3.76)^{**}}{0.15}X_4 - \underset{(-3.43)^{**}}{0.206}X_5 + \underset{(1.82)^{-}}{63.62}X_6 - \underset{(-1.90)^{-}}{36.83}X_7 - \underset{(-1.77)^{-}}{0.145}X_8$$

$R = 0.99$ $\quad R^2 = 0.98$ $\quad F = (65.67)^{**}$

Y = كمية الاستهلاك من الزيوت بالألف طن **X_1 = كمية الإنتاج المحلي من الزيوت بالألف طن**

X_2 = كمية واردات الزيوت بالألف طن **X_3 = متوسط سعر استيراد الزيوت بالدولار/ طن**

X_4 = الدخل الفردي الحقيقي بالجنيه/ سنة **X_5 = متوسط سعر التجزئة الحقيقي بالجنيه/ طن**

X_6 = سعر صرف الدولار مقابل الجنيه **X_7 = عدد السكان بالمليون نسمة**

X_8 = قيمة الدعم الحقيقي للزيوت النباتية مليون جنيه

القيمة بين القوسين أسفل معامل الانحدار للمتغيرات المستقلة تعبر عن قيمة *t* المقدرة.

(**) معنوية عند مستوي ١% **(*) معنوية عند مستوي ٥%** **(–) غير معنوية**

جدول (٤) معامل تضخم التباين والدليل الشرطي والجذور الكامنة وحدود التسامح للمتغيرات المستقلة قبل إزالة الازدواج الخطي

Variables	*VIF*	*Condition Index*	*Eigenvalues*	*Tolerance*
X_1	٢,٩٣٣	٤,٥٥	٠,٣٩٨	٠,٣١
X_2	٢,٩٥٥	٥,٨٣٥	٠,٢٤٢	٠,٣٥
X_3	١٠,٣٦٤	١٠,٠٩٨	٠,٠٨١	٠,٠٨
X_4	٣٥,٩٦٣	١٥,٩٦٥	٠,٠٣٢	٠,٠٢
X_5	٨,١٦٧	٢٥,٠٢٣	٠,٠١٣	٠,١٣
X_6	٨,٤٣٤	٤٨,٢٦٩	٠,٠٠٤	٠,١٢
X_7	١٠٥,٧٨٢	٦٣,٣٥٣	٠,٠٠٢	٠,٠٠٨
X_8	٢١,٥٧٩	٣٣٠,٣١٥	٠,٠٠٠٠٧	٠,٠٤

المصدر: نتائج التحليل الإحصائي ببرنامج *SPSS*.

توضح المعادلة السابقة أن زيادة كمية الواردات من الزيوت النباتية بنحو ألف طن سوف يؤدي إلي زيادة الاستهلاك بنحو ٠,٨٤ ألف طن، أما عند زيادة الدخل الفردي الحقيقي بنحو جنيه/ السنة سوف يؤدي إلي زيادة استهلاك الزيوت النباتية بنحو ٠,١٥ ألف طن، كما يتبين أن انخفاض متوسط سعر التجزئة الحقيقي بنحو جنيه/ الطن سوف يؤدي إلي انخفاض الكمية المستهلكة من الزيوت النباتية بنحو ٠,٢٠٦ ألف طن سنوياً.

تشير المعادلة السابقة إلي أن المتغيرات التفسيرية موضع الدراسة تفسر نحو ٩٨% من إجمالي التغيرات في كمية استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠– ٢٠١٥)، كما يتبين معنوية معاملات الانحدار لكل من كمية واردات الزيوت النباتية والدخل الفردي الحقيقي ومتوسط سعر التجزئة الحقيقي في حين لم تثبت معنوية معاملات الانحدار لبقية المتغيرات التفسيرية بالمعادلة، هذا ولم تتفق إشارات كل من المتغيرات X_3 و X_6 و X_7 و X_8 مع المنطق الاقتصادي.

عند إجراء تحليل الانحدار الخطي المتعدد المرحلي(١٢) (*Stepwise*) للمتغيرات الاقتصادية موضع الدراسة تبين أن المعادلة كانت على النحو التالي:

$$\hat{Y}_I = \underset{(2.81)^{*}}{1805.363} + \underset{(11.866)^{**}}{0.729}\,X_2 + \underset{(4.994)^{**}}{0.157}\,X_4 - \underset{(-2.48)^{**}}{0.080}\,X_5 - \underset{(-3.21)^{**}}{30.520}\,X_7$$

$$R = 0.98 \qquad R^2 = 0.97 \qquad F = (81.50)^{**}$$

تشير المعادلة السابقة إلى أهم العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠- ٢٠١٥) ومنها يتبين أن زيادة كمية واردات الزيوت بألف طن سنوياً سوف يؤدي إلي زيادة الكمية المستهلكة من الزيوت النباتية بنحو ٠,٧٢٩ ألف طن، وأن زيادة الدخل الفردي الحقيقي بجنيه واحد سوف يؤدي إلى زيادة الكمية المستهلكة من الزيوت النباتية بنحو٠,١٥٧ ألف طن، وأن زيادة متوسط سعر التجزئة الحقيقي بجنيه واحد سنوياً سوف يؤدي إلى انخفاض الكمية المستهلكة من الزيوت النباتية بنحو٠,٠٨٠ ألف طن، وأن زيادة عدد السكان بنحو مليون نسمة سوف يؤدي إلى نقص الكمية المستهلكة من الزيوت النباتية بنحو٣٠,٥٢٠ ألف طن ويلاحظ أن إشارة هذا المتغير لا تتفق مع المنطق الاقتصادي.

يتبين من القيمة (t) (١١) المقدرة التي بين القوسين أسفل معاملات الانحدار معنوية ثوابت المعادلة كما تبين قيمة (F) المقدرة معنوية المعادلة كما يتبين معنوية المعادلة عند المستوي الاحتمالي ١%، أن المتغيرات السابقة تفسر نحو٩٧% من إجمالي العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر خلال فترة الدراسة، وتوضح قيمة معامل تضخم التباين (VIF) المقدرة أنها بلغت نحو ١,٣٤ للمتغير (X_2) ونحو ١٥,١١ للمتغير (X_4) ونحو ١,٤٩ للمتغير (X_5) ونحو ١٦,١٠ للمتغير (X_7) وهو ما يشير إلي وجود الازدواج الخطي للمتغيرين (X_4) و (X_7) حيث كانت قيمة (VIF) أعلي من ١٠ بينما كانت قيمة (VIF) للمتغيرين (X_2) و(X_5) أقل من ١٠ القيمة ١٠ أي أن إجراء تحليل الانحدار الخطي المتعدد المرحلي لم يتمكن من إزالة ازدواج الخطي الموجود بين المتغيرات موضع الدراسة.

تحليل الانحدار المتعدد باستخدام طريقة انحدار ريدج *RR* (*Ridge Regression*):

إذا كانت العلاقة الخطية بين المتغير المستقل والمتغير التابع تأخذ الصورة $\hat{Y} = Xb + e$ ..... المعادلة (١) حيث Y هي متجه يستجيب لعدد n من القيم (وهي المتغير التابع) أما المتغير X هو عبارة عن $n \times p$ (المتغيرات التفسيرية) مع الرتبة p ومعامل الانحدار b هو متجه للمعامل p للمتغيرات التفسيرية وقيمة e هي الخطأ العشوائي وبافتراض أنه يساوي صفر وتباينه ثابت، أي أن $E(e) = 0$ و $Var(e) = \delta^2 I_n$ ، ففي ظل افتراضات طريقة (OLS) فإن تقدير b في المعادلة السابقة تساوي $\hat{b} = (X^T X)^{-1} X^T Y$ ..... المعادلة (٢)، وفي ظل وجود الازدواج الخطي فإن قيمة الثابت k سوف تضاف لعناصر المصفوفة القطرية في مصفوفة $X^T X$ في المعادلة (٢) وذلك لتقليل التبعية في المتغيرات التفسيرية وهو ما يعرف بطريقة *Ridge Regression* لتأخذ الصورة التالية $\hat{b}_R = (X^T X + KI_n)^{-1} X^T Y$ ..... المعادلة (٣).[١٦]

نتائج تطبيق طريقة انحدار ريدج *RR* (*Ridge Regression*):

نظراً لوجود مشكلة الازدواج الخطي بين متغيرات دالة استهلاك الزيوت تم استخدام طريقة (*Ridge Regression*) ولكون هذه الطريقة تعتمد على تحديد قيمة K وبما أنه هناك أكثر من طريقة لاختيار قيمة K فإن قيمتها تقع بين الصفر والواحد الصحيح ويوضح جدول (٥) قيم K تم اختيارها وتم تحديد معامل التحديد R^2 ومعامل تضخم البيانات VIF و $Sigma$ ومنها يتم اختيار قيمة K التي تعطي قيمة صغيرة نسبياً VIF أقل من ١٠ بالتالي تم حساب المعادلة بعد إزالة مشكلة الازدواج الخطي كالتالي:

$$\hat{Y}_i = 147.033 + \underset{(0.392)^{-}}{0.42}\,X_1 + \underset{(4.805)^{**}}{0.567}\,X_2 - \underset{(-0.758)^{-}}{0.05}\,X_3 + \underset{(1.714)^{-}}{0.036}\,X_4 - \underset{(-1.400)^{-}}{0.09}\,X_5 + \underset{(1.332)^{-}}{55.35}\,X_6 + \underset{(0.501)^{-}}{2.01}\,X_7 - \underset{(-4.386)^{**}}{0.025}\,X_8$$

$$R = 0.95 \quad R^2 = 0.90 \quad F = (8.28)^{**} \quad CV = 0.17 \quad K = 0.1$$

$$Ave\ VIF = 1.25$$

أوضحت النتائج أنه عندما تكون $K = 0.1$ فإن المعادلة سوف تكون معنوية عند المستوي الاحتمالي ١% حيث بلغت قيمة F نحو (٨,٢٨)** وأن معامل الاختلاف سوف يكون منخفض حيث بلغت قيمته $CV = 0.17$ في حين يبلغ متوسط قيمة معامل تضخم التباين للمتغيرات موضع الدراسة نحو $Ave\ VIF = 1.25$ ، هذا وتفسر المتغيرات موضع الدراسة نحو ٩٠% من إجمالي العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠– ٢٠١٥) كما تبين معنوية معاملات انحدار المتغير (X_2) والمتغير (X_8) فقط في حين لم يتبين معنوية بقية المتغيرات موضع الدراسة في النموذج ومن ثم يتبين أن زيادة واردات الزيوت النباتية بنحو ألف طن سوف يؤدي إلي زيادة استهلاكها سنوياً في مصر بنحو ٠,٥٦٧ ألف طن، كما أن زيادة قيمة الدعم الحقيقي للزيوت النباتية سوف يؤدي إلي زيادة الاستهلاك سنوياً بنحو ٠,٠٢٥ ألف طن وذلك خلال فترة الدراسة.

يلاحظ أن زيادة الإنتاج المحلي من الزيوت النباتية سوف يؤدي لزيادة الاستهلاك وذلك لانخفاض أسعارها محلياً عن الزيوت المستوردة، بينما يزيد الاستهلاك نتيجة زيادة الكمية المستوردة من الزيوت وذلك لزيادة المعروض منها، حيث يعتمد السوق المصري بصفة رئيسية علي توفير الزيوت اللازمة للمستهلك بالاستيراد من الخارج، وعند زيادة سعر صرف الدولار مقابل الجنيه المصري ستظل الكمية المستهلكة من الزيوت النباتية كما هي إن لم تنخفض، وذلك لأهمية الزيوت النباتية وتنوع استخدامها لدي المستهلك في مصر فضلاً عن ارتفاع أسعار السلع البديلة لها مثل المسلي الحيواني.

معاملات انحدار *Ridge* القياسية:

يوضح جدول (٥) معاملات انحدار *Ridge* القياسية للمتغيرات المستقلة المتحصل عليها من تطبيق طريقة (RR) عند تحديد قيم مختلفة من K والتي تراوحت بين (٠,١ – ٠,٠٠٠٠٠١) ومنه يتبين أن المتغيرين (X1) و (X_7) أخذا قيم سالبة عند تكرار التحليل باختلاف قيمة K إلي أن بلغت قيمة (K= 0.1) تغيرت عندها إشارة المتغيرين إلي موجبة، أما المتغيرين (X_2) و (X_6) فقد احتفظا بإشارتهما الموجبة مع تكرار قيمة K وبالنسبة للمتغيرين (X_5) و (X_8) فقد احتفظا بإشارتهما السالبة مع تكرار قيمة K وأخيراً فإن أشارة المتغيرين (X_3) و (X_4) فقد تغيرت من موجبة إلي سالبة عندما بلغت قيمة K=0.01.

معامل تضخم التباين (*VIF*) عند مستويات مختلفة من *K* :

تبين النتائج الواردة في جدول (٦) معامل تضخم التباين (VIF) المقدر لكل من المتغيرات المستقلة موضع الدراسة عند مستويات مختلفة من K أنه عندما بلغت قيمة K= 0.000001 فإن قيمة (VIF) كانت أعلي من ١٠ لكل من المتغيرات (X_3)، (X_4) و(X_7) بينما كانت أقل من ١٠ لبقية المتغيرات تحت الدراسة، واستمر الوضع علي ما هو عليه عند تغير قيمة K إلي أن بلغت نحو K= 0.001 حيث انخفضت قيمة (VIF) للمتغير (X_3) فقط لأقل من ١٠ ويفيد العمود الأخير بالجدول بانخفاض متوسط قيم (VIF) للمتغيرات التفسيرية تدريجياً من ٢٤,٤٩ عندما بلغت قيمة K= 0.000001 إلي نحو ٩,٣٠ عندما بلغت قيمة K= 0.007 في حين عندما بلغت قيمة K= 0.1 كانت قيمة (VIF) لجميع المتغيرات المستقلة للدراسة أقل من ١٠ وبذلك يبدو أن أفضل قيمة هي القيمة التي تم تحديدها والتي توفر أدنى قيم لمعامل تضخم التباين (VIF).

يوضح جدول (٦) أيضاً معامل التحديد (R^2) وهو يشير إلي أنه قدر بنحو ٩٩% تقريباً عندما كانت قيمة K=0.000001 وأخذ في الانخفاض تدريجياً إلي أن بلغ نحو ٩٠% عندما أفضل قيمة مقدرة لقيمة K وهي تساوي ٠,١ ، أي أن قيمة K تتناسب تناسباً عكسياً مع قيمة R^2 وهو ما يشير إلي أن قيمة R^2 أخذت قيمة أكبر من قيمتها الحقيقية في حالة وجود الازدواج الخطي، وبالمثل فإن العلاقة عكسية بين قيمة K وقيمة

(*VIF*) وفي المقابل فإنه بزيادة قيمة *K* تتزايد قيم *Sigma* والتي تعبر عن الجذر ألتربيعي لمتوسط مربعات الخطأ أي أن العلاقة طردية بينهما حيث يتبين أنه عندما بلغت قيمة *K*= 0.1 كانت قيمة *Sigma* تساوي ١٦٣,٧٧.

جدول (٥) معاملات انحدار *Ridge* القياسية للمتغيرات المستقلة عند مستويات مختلفة من قيم *K*

K	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
٠,٠٠٠٠٠١	٠,٠٢٥-	٠,٨٦٠	٠,١٩٩	٠,٩٧٥	٠,٤٢٥-	٠,٢٢٩	٠,٨٤٦-	٠,٣٥٧-
٠,٠٠٠٠١	٠,٠٢٥-	٠,٨٦٠	٠,١٩٩	٠,٩٦٠	٠,٤٢٥-	٠,٢٢٩	٠,٨٤٥-	٠,٣٥٧-
٠,٠٠٠١	٠,٠٢٥-	٠,٨٥٨	٠,١٩٦	٠,٩٠٦	٠,٤٢٥-	٠,٢٢٨	٠,٨٣٢-	٠,٣٥٩-
٠,٠٠١	٠,٠٢٥-	٠,٨٦٠	٠,١٦٦	٨٥٠.	٠,٤٢٣-	٠,٢١٩	٠,٧٢٥-	٠,٣٧٢-
٠,٠٠٧	٠,٠٢٩-	٠,٨٤٩	٠,٠٦٠	٠,٦٧٨	٠,٣٩٥-	٠,١٩٨	٠,٣٤٦-	٠,٣٧٨-
٠,٠١	٠,٠٣٣-	٠,٧٧٠	٠,٠٣٣	٠,١٦١	٠,٣٧٩-	٠,١٩٦	٠,٢٨١-	٠,٣٦٤-
٠,١	٠,٠٤٥	٠,٥٨٣	٠,٠٩٦-	٠,٠٩٦-	٠,١٨٩-	٠,١٩٩	٠,٠٤٦	٠,٠٦٣-

المصدر: حسبت من نتائج تحليل جدول (١) بواسطة برنامج *NCSS*.

جدول (٦) معامل تضخم التباين"(*VIF*) عند مستويات مختلفة من *K*

K	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	R^2	*Sigma*	*Ave VIF*
٠,٠٠٠٠٠١	٢,٩٣	٢,٩٥	١٠,٣٣	٣٥,٩٤	٨,١٧	٨,٤١	١٠٨,٥١	٢١,٦٠	٠,٩٨٨	٦٠,٦٧	٢٤,٤٩
٠,٠٠٠٠١	٢,٩٣٣	٢,٩٥	١٠,٣١	٣٥,٨٦	٨,١٧	٨,٤١	١٠٥,٢٤	٢١,٥٧	٠,٩٨٦	٦٠,٧٣	٢٤,٤٣
٠,٠٠٠١	٢,٩٢	٢,٩٣	١٠,٣١	٣٥,١٩	٨,١٣	٨,٣٦	١٠٢,٥٣	٢١,٣١	٠,٩٨٦	٦١,٣١	٢٣,٤٩
٠,٠٠١	٢,٨١	٢,٧٦	٨,٦١	٢٩,٥٧	٧,٧٣	٧,٩١	٨٠,٤٥	١٩,١٣	٠,٩٨٣	٦٦,٤٣	١٩,٨٧
٠,٠٠٧	٢,٤٣	٢,٢١	٤,٦٢	١٤,٦٠	٦,٠٢	٦,١٩	٢٦,١٢	١٢,١٤	٠,٩٧٠	٨٧,٢٥	٩,٣٠
٠,٠١	٢,٣٢	٢,٠٧	٣,٩١	١١,٨١	٥,٤٣	٥,٦٧	٢٢,٨٦	١١,٥٠	٠,٩٦٨	٩٣,٩٨	٧,٤٤
٠,١	١,٢٨	١,٠٨	١,١٥	١,٤٨	١,٣١	١,٦٤	٠,٦٢	١,٤٦	٠,٩٠٠	١٦٣,٧٧	١,٢٥

المصدر: حسبت من نتائج تحليل جدول (١) بواسطة برنامج *NCSS*.

في ضوء ما تقدم يتبين فعالية طريقة انحدار *Ridge* في معالجة مشكلة الازدواج الخطي في حالة تحليل الانحدار الخطي المتعدد وهذا يتطلب تفعيل طرق اكتشاف المشكلة من جانب الباحثين الإحصائيين والاقتصاديين وتحديد أسبابها بدقة، وتوصي الدراسة باستخدام طريقة انحدار *Ridge* عند معالجة الازدواج الخطي باعتبارها طريقة فعالة ومفيدة وموثوق فيها مع بيان عيوبها.

الملخص

يجب أن يتوفر في نموذج الانحدار المتعدد عدة افتراضات فإذا سقط فرض أو أكثر من هذه الافتراضات يمكن أن تظهر بعض المشاكل التي لها بعض الآثار غير المرغوبة، والتي ينشأ عنها تقديرات غير دقيقة لمعالم المعادلة ولعل أهم هذه المشاكل ما تناولته هذه الدراسة وهي مشكلة الازدواج الخطي والتي تعني وجود علاقة خطية قوية بين المتغيرات التفسيرية المستخدمة في نموذج الانحدار المتعدد، وتوجد عدة طرق لاكتشاف ومعالجة هذه المشكلة من بينها طريقة انحدار *Ridge* والتي تعطي تقديرات متحيزة نوعاً ما لمعلمات المعادلة ولكن مع توفير تباين أصغر علي الرغم من أهميتها كوسيلة فعالة في تخفيف آثار الازدواج الخطي، فضلاً عن مقارنة نتائجها مع نتائج طريقة (*OLS*) وهذه المشكلة نادراً ما يتم بحثها عند تحليل المتغيرات الاقتصادية والتي يترتب عليها تقديرات خاطئة، وقد كشفت الدراسة عن وجود الازدواج الخطي في النموذج الإحصائي للمتغيرات المؤثرة علي استهلاك الزيوت النباتية في مصر خلال الفترة (٢٠٠٠-٢٠١٥)، حيث تؤثر المتغيرات الاقتصادية موضع الدراسة بنحو ٩٨% في ظل وجود مشكلة الازدواج الخطي ونحو ٩٠% بعد معالجة مشكلة الازدواج الخطي بطريقة (*RR*) وذلك من إجمالي العوامل المؤثرة علي استهلاك الزيوت النباتية في مصر، وتوصي الدراسة بتفعيل طرق اكتشاف المشكلة من جانب الباحثين

الإحصائيين والاقتصاديين وتحديد أسبابها بدقة، والتوسع في استخدام طريقة انحدار *Ridge* عند معالجة الازدواج الخطي باعتبارها طريقة فعالة ومفيدة وموثوق فيها مع بيان عيوبها.

المراجع:

١- أموري هادي كاظم (٢٠٠٥)، مقدمة في الاقتصاد القياسي، الطبعة الأولي، مطبعة جامعة الموصل، العراق.

٢- مزاحم محمد يحيي، محمود حمدون عبد الله، (٢٠٠٧)، تشخيص التعدد الخطي واستخدام انجراف الحرف في اختيار متغيرات دالة الاستثمار الزراعي في العراق في الفترة (١٩٨٠- ٢٠٠٠)، مجلة تكريت للعلوم الإدارية والاقتصادية، المجلد (٣)، العدد (٨)، جامعة تكريت كلية الإدارة والاقتصاد.

٣- الجهاز المركزي للتعبئة العامة والإحصاء، المركز القومي للمعلومات، أعداد متفرقة.

٤- الجهاز المركزي للتعبئة العامة والإحصاء، الكتاب السنوي، أعداد متفرقة.

٥- الجهاز المركزي للتعبئة العامة والإحصاء، النشرة السنوية للتجارة الخارجية، أعداد متفرقة.

٦- الجهاز المركزي للتعبئة العامة والإحصاء، النشرة السنوية لأسعار المواد والمنتجات الغذائية والخدمات، أعداد متفرقة.

7- Abdalla M. EL-Habil, Khaled I.A. Almghari (2011). "Remedy of Multicollinearity using Ridge regression", Journal of Al Azhar University-Gaza (Natural Sciences), 13 : 119-134.

8- Berk, Kenneth N(1977).Tolerance and condition in Regression computations, Journal of American statistical ssocations,Vol.72,P(863-866).

9- Brown,C.E (1998). Applied Multivariate , statistics in Geohydrology and related sciences , Springer – verlag . Berlin Heidelberg , chapter 6 , multiple regression . pp. 62-66.

10- Draper, N.R. and smith , H.(1981). Applied Regression analysis , New York , P.413.

11- Freund R.J, Wilson W.J (1998). Regression analysis Statistical modeling of aresponse variable, Acadimic press, San Diego, California, USA.

12- Gomez, K.A.,A. A. Gomez (1984). Statistical procedures for agricultural research, 2nd ed, John Wiley and Sons, New York, U S A.

13- Issam A.W. Mohamed (2011). Effect of Multicollinearity on the Estimation of Macroeconomic Variables: Using Data from Sudan, Al-Neelain University, 2011 :34- 35.

14- Mahavidyalaya Halkarni Y C, Tal-Chandgad, Kolhapur, Maharashtra (2014). On Comparison of Some Ridge Parameters in Ridge Regression Ashok V. Dorugade, Sri Lankan Journal of Applied Statistics, Vol (15-1),31- 46, India.

15- Vedide REZAN Uslu, Erol Egrioglu, Eren Bas, (2014). Finding optimal value for the shrinkage parameter in Ridge Regression via particle Swarm optimization, American Journal of Intelligent system, Turkey,4(4):142:147.

16- Shariff N.S.Md. and N.A.Ferdaos(2017). An application of robust ridge regression model in the presence of outliers to real data problem، Journal of physices, Conf. Ser. Fac. Of Sci. and Tech. Bandar Baru Nilai ,Malaysia : 890-896.

Increase The Efficiency Of Prediction Equation For Multicollinearity In Vegetable Oils

Sahar Abd El-Salam Ibraheem
Cent. Lab. for Design & Stat. Analysis Res., ARC.

summary

Egypt's lack of vegetable oil production is of paramount importance and is an important target because it is an important food for humans and a commodity of strategic food commodities, The nutritional importance of vegetable oils is due to the fact that it contains important fat-soluble vitamins and is a major source of energy, The problem of vegetable oils in Egypt is limited to the limited quantities of vegetable oil produced for the required quantities, Local production is unable to cover the needs of local consumption and thus the food gap widened.

In the multiple regression model there should be several assumptions. If one or more of these assumptions fall, some of the problems that have some undesirable effects may arise, resulting in inaccurate estimates of the equation. Perhaps the most important of these problems is multicollinearity problem which means that there is a strong linear relationship between the explanatory variables used in the multiple regression model. There are several ways to detect and address this problem, including the Ridge regression method, which gives somewhat biased estimates of equation parameters but with a smaller variance in addition to comparing its results with those of OLS method. The study revealed the existence of multicollinearity in the statistical model of the variables affecting the consumption of vegetable oils in Egypt during the period of the study. (2000 - 2015), where the economic variables affected about 98% in the presence of the problem of multicollinearity and about 90% of the total factors affecting the consumption of vegetable oils in Egypt after the treatment of the problem using (RR) method (RR). The study recommends activation methods of identifying multicollinearity problem by economists and statisticians and accurately determine their causes, and expand the use of Ridge regression method when processing multicollinearity as an effective, useful and reliable way with a statement in which its flaws.