استخدام طريقة المربعات الصغرى الجزئية للتخلص من تعدد العلاقة الخطية

د.صفاء يونس الصفاوي سيف الدين ضياء الدين صالح مؤيد شاكر
أستاذ مساعد مدرس مساعد مدرس مساعد
كلية علوم الحاسبات والرياضيات / قسم الإحصاء والمعلوماتية

الملخص:

في هذا البحث تم توضيح تقدير معلمات الانحدار باستخدام طريقة المربعات الصغرى الجزئية ومقارنتها بطريقة المربعات الصغرى الاعتيادية من حيث قابليتها على التخلص من مشكلة التعدد الخطي بين المتغيرات التنبؤية ،تتميز الطريقة الأولى بإمكانية إيجاد تحليل الانحدار لعدة متغيرات استجابة مع عدة متغيرات تنبؤية في إن واحد،وتم تطبيق كلتا الطريقتين على بيانات إنتاج مادة الاسمنت حيث وصفت المواد الداخلة في صناعة مادة الاسمنت بالمتغيرات التنبؤية إما بالنسبة إلى متغيرات الاستجابة فتمثلت بمادة الاسمنت والكلنكر والغبار والنفايات الصلبة.

Use Partial Least Squares Method To Remove Multicollinearity

Abstract:

In the clarification of this research was estimating the regression parameters using least squares and partial least squares compared with the normal in terms of its ability to of Remove the problem of Multicollinearity between the predictive variables, the first method is the possibility of a regression analysis of several response variables with a number of predictive variables at the same time, was the application of both methods on the data production of cement as described in the manufacture Material of cement, either predictive variables for the response variables by cement and clinker, dust and solid waste.

مقدمة :

ظهرت الحاجة إلى استخدام طريقة المربعات الصغرى الجزئية((PLS) Partial Least Squares) في العديد من الدراسات منها الكيمياء والعلوم الاجتماعية ويعد الباحث (Wold ; 1966)اول من وجد هذه الطريقة وأصبحت شائعة الاستخدام في الطب وخاصة في المعالجات السريرية حيث قلة عدد المشاهدات (عدد المرضى) مع كثرة عدد المتغيرات (X'_S) المتمثلة بأعراض المريض مع عدد المتغيرات المعتمدة(Y'_S) والمتمثلة بالمستوى الصحي للمريض

مع تحسين حالته الصحية ونتائج أخرى.كما طورت هذه الطريقة من قبل (Friedman ; 1993) .

إن الهدف من أسلوب انحدار المربعات الصغرى الجزئية هو التنبؤ بمتغيرات الاستجابة (Y'_S) من المتغيرات التنبؤية (X'_S) ومن ثم وصف هذا التنبؤ من خلال علاقة متمثلة بالنماذج الخطية التي ستكونها.

عندما يكون Y متجه و X مصفوفة ذات رتبة كاملة فان عملية تكوين نموذج خطي للتنبؤ بـ Y من خلال X ممكنة ويمكن تطبيق انحدار متعدد متغيرات بشكل اعتيادي ، ولكن عندما يكون عدد المتغيرات التنبؤية كبير ويفوق عدد المشاهدات فانه في هذه الحالة ستكون المصفوفة X شاذة Singular ولا يمكن تطبيق الانحدار لمتعدد المتغيرات الاعتيادي وهذا أيضا بدوره سيظهر مشكلة تعدد العلاقة الخطية بين المتغيرات التنبؤية.

هناك حالتين في استخدام طريقة المربعات الصغرى الجزئية ، الأولى والتي يرمز لها بـ (PLS1) تتمثل بتطبيق هذه الطريقة للتنبؤ بقيم متغير معتمد واحد Y مقابل K من المتغيرات التنبؤية X وهذه حالة خاصة ، والثانية هي الحالة العامة يرمز لها بـ (PLS2) تتمثل بتطبيق هذه الطريقة في حالة (j) من المتغيرات المعتمدة ($y_{ij} \quad ,i=1,2,...,n, \quad j=2,3,...,n$) مع (K) من المتغيرات التنبؤية ($x_{ik} \quad ,i=1,2,...,n, \quad k=2,3,...,n$) وهذه الطريقة تتطلب تكرار للوصول إلى الحل النهائي. (Harald & luis ; 1986).

الجانب النظري:

تعتمد طريقة انحدار المربعات الصغرى الجزئية على مصفوفة التباين المشترك((cov (X,Y) ، إذ أن هذه الطريقة تسمح بتحديد العوامل والتي هي عبارة عن تراكيب خطية للمتغيرات التنبؤية (X) ، وتعرف هذه العوامل بالمتغيرات الصماء (Latent Variables) والتي بدورها تكون أفضل نموذج للمتغيرات المعتمدة (Y'_S)،تعد طريقة المربعات الصغرى الجزئية مشابهة إلى كل من الطرائق(Canonical Correlation) و(Principal Component) و(Discriminate Analysis) إلا أن هذه الطريقة تفرض شروط عند تطبيقها منها(Dante ; 2006):

1-العوامل أو المكونات (Extracted Factors) لكل من المتغيرات Y والمتغيرات X تكون مستخلصة من المصفوفات ($Y'Y$) و($X'X$) فقط وليس من المصفوفات التي جاءت من حاصل ضرب المتغيرات X في Y أي ليس من ($X'Y$) او غيرها.

2-عدد المعادلات التنبؤية يجب أن لا يتجاوز الحد الأدنى لعدد المتغيرات المعتمدة اوالتنبؤية.
بالنسبة لطريقة المربعات الصغرى الجزئية فهي طريقة موسعة للانحدار الخطي المتعدد ولكن بدون الشرطين السابقين ، حيث إن المعادلات التنبؤية والتي مثلت من خلال العوامل التي تم الحصول عليها من المصفوفات ($Y'X'XY$) للحصول على نموذج الانحدار للتنبؤ بقيم Y.

التركيب أو البناء الداخلي لطريقة المربعات الصغرى الجزئية :

إن طريقة المربعات الصغرى الجزئية تسمى بالمتغيرات الصماء او الذاتية احياناً وذلك لأنها تكون مشابها بتركيبتها للمتجهات الذاتية في التحليل الذاتي ويمكن وصف هذه التركيبة بالشكل التالي :

ليكن لدينا مدور مصفوفة المتغيرات التنبؤية ($X's$) مضروبة بمتجه عشوائي قبل U أي أن:

$$X'U = \begin{bmatrix} \sum X_{i1}U_i \\ \sum X_{i2}U_i \\ \vdots \\ \sum X_{ip}U_i \end{bmatrix} \quad \ldots. (1)$$

ومن خلالها يمكن الحصول على مصفوفة الـ β والتي توضح نوع وشكل العلاقة بين كل متغير معتمد مع كل متغير تنبؤي أي أن :

	Y_1	Y_2	Y_3	$\ldots$	Y_j
X_1	b_{11}	b_{12}	b_{13}	$\ldots$	b_{1j}
X_2	b_{21}	b_{22}	b_{23}	$\ldots$	b_{2j}
X_3	b_{31}	b_{32}	b_{33}	$\ldots$	b_{3j}
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
X_K	b_{K1}	b_{K2}	b_{K3}	$\ldots$	b_{Kj}

إن طريقة المربعات الصغرى الجزئية عبارة عن تركيبة خطية للمربعات الصغرى لمصفوفة الارتباط والتباين المشترك بين المتغيرات التنبؤية والمتغيرات المعتمدة ، والجزء الذي تعتمد عليه طريقة المربعات الصغرى الجزئية في مصفوفة الارتباط والتباين المشترك هو جزء التقاطع Cross Block أي الارتباطات بين المتغيرات التنبؤية والمتغيرات المعتمدة ، كما تقدم هذه الطريقة عوامل Factor Scores على شكل مجموعات خطية بين المتغيرات التنبؤية الأصلية ، بذلك سوف لا يكون هناك ارتباطات بين عوامل المتغيرات المستخدمة في نموذج الانحدار التنبؤي .

وعليه يمكن إجراء تحليل للمصفوفتين X ، Y من خلال العلاقــة (Saikat & Jun ; 2008)

$$X=TP'+K \qquad ...(2)$$
$$Y=UC'+R \qquad ...(3)$$

حيث ان:

T = مصفوفة العوامل المستخلصة من المصفوفة X.

P = هو متجه محمل للمصفوفة X.

U = مصفوفة العوامل المستخلصة من المصفوفة Y.

C = متجه محمل للمصفوفة Y.

وان عملية استخلاص عوامل المصفوفة X تتمثل بإيجاد تركيبة خطية من أعمدة المصفوفة X ، أي إيجاد متجه ابتدائي يتم ضربه بالمصفوفة X لإيجاد التراكيب الخطيــة t أي ان :

$$t=XW$$

حيث ان W هو متجه لقيم عشوائية أو هو أول متجه مميز المقابـل لأول قيمـة مميـزة للمصفوفة ($X'YY'X$) ، بنفـس الطريقة فإن التركيبة الخطيـة للمـصفوفة Y والمتمثلـة بالمتجه U كما يلي :

$$U=YC$$

حيث أن C هو أول متجه مميز مقابل إلى أول جذر مميـز للمـصفوفة ($Y'XX'Y$) مـع العـلم أن ($X'Y$) تمثل مصفوفة التباين المشترك بين X و Y .

إن هذه الطريقة تبحث عن مجموعة من المكونات والتي تدعى بالمتجهات الـصماء (Latent Vectors) وان هذه المتجهات توضح قدر الإمكان التباين المشترك بين X و Y .

ويمكن تحليل المتغيرات المستقلة من خلال المعادلة التالية :

$$X=TP'$$

حيث T هي عبارة عن مجموعة خطية من المتغيرات التنبؤية ولكن علـى شـكل عوامـل متعامدة ، أي أن كل عمود يحوي على جميع المتغيرات المستقلة الموجودة في X ، ولكـن على شكل مجموعة خطية من الأوزان ، أما P فهو متجه محمل والمقصود به عبارة عـن مجموعة خطية بين العوامل المتعامدة t والمصفوفة الأصلية للمتغيرات المستقلة أي ان :

$$P=X't$$

حيث ان t هو عمود من أعمدة المصفوفة T

$$T'T=I$$

$P'P \neq I$

وبعد إيجاد أول متجه ذاتي ، يتم طرحه من كل من X و Y ويتم تكــرار أو إعــادة هــذا الإجراء حتى تصبح X مصفوفة صفرية .

خوارزمية طريقة انحدار المربعات الصغرى الجزئية :

يمكن توضيح آلية عمل طريقة المربعات الصغرى الجزئية من حيث حساب معلمات الانحدار ومصفوفة المتغيرات الصماء T والأعمدة المحملة P ، وكخطوة أولية يــتم تحويــل مصفوفة المتغيرات التنبؤية X ومصفوفة المتغيرات المعتمدة Y إلى الــصيغة القيــاسية بحيث تصبح E و F على التوالي ، وخطوات هذه الخوارزمية كما يلي :

1- اخذ قيم ابتدائية عشوائية للمتجه U.

2- إيجاد المتجه W من خلال العلاقة:

$$W=E'U$$

3- إيجاد المتجه t وهو احــد أعمــدة المــصفوفة T للمتغيــرات المستخلــصة مــن المصفوفة X أي أن:

$$t_{old} = EW$$

4-قسمة المتجه t على طول المتجه t أي أن :

$$t_{new} = \frac{t_{old}}{\|t\|}$$

حيث أن

$$\|t\| = \sqrt{t_1^2 + t_2^2 + t_3^2 + \ldots\ldots + t_n^2}$$

5-إيجاد المتجه الموزون C للمصفوفة Y :

$$C=F't$$

6-إيجاد المتجه U وهو احد أعمدة المصفوفة U للمتغيرات المستخلصة للمصفوفة Y

$$U_{old} = FC$$

7-قسمة المتجه U على طول المتجه U وكما يلي :

$$U_{new} = \frac{U_{old}}{\|U\|}$$

حيث أن

$$\|U\| = \sqrt{U_1^2 + U_2^2 + U_3^2 + \dots\dots + U_n^2}$$

8-مقارنة قيم المتجه U في الخطوة (7) مع قيم المتجه t في الخطوة (4) فإذا كانت القيم متقاربة يتم حساب قيم b وفي حالة عدم التقارب يتم حساب متجه المحمل P للمصفوفة X من خلال المعادلة التالية:

$$P = E't$$

بعد ذلك يتم طرح تأثير المتجه t (Partial Out) من كل من المصفوفتين E و F وكما يلي:

$$E_1 = E - tP'$$
$$F_1 = F - b\,tC'$$

ومن المعادلتين من الخطوة (5) و (8) نجد :

$$\begin{aligned} E_1 &= E - t(E't)' \\ &= E - tt'E \end{aligned}$$

$$\begin{aligned} F_1 &= F - b\,t(F't)' \\ &= F - b\,tt'F \end{aligned}$$

بعد ذلك تعاد الخوارزمية من جديد بنفس الخطوات ولكن باستخدام المصفوفتين الجديدتين E_1 و F_1 مع العلم أن في كل عملية تكرار يتم مقارنة قيم المتجه (t) للدورة الأخيرة مع قيم المتجه (t) للدورة السابقة ، وتستمر عملية التكرار إلى أن يتم الحصول على جميع أعمدة المصفوفة T (والتي عددها مساو إلى رتبة المصفوفة X) ومن ثم يتم إيجاد مصفوفة المعلمات b وكما يلي :

$$b = E'U(T'EE'U)^{-1}T'F$$

الجانب العملي :

البيانات المستخدمة في هذا البحث تم الحصول عليها من معمل اسمنت كركوك وللفترة (2002 - 2006) حيث تمثلت العملية الإنتاجية من أربعة عناصر أساسية كمواد أولية (مدخلات) والتي اعتبرت كمتغيرات تنبؤية لإنتاج مادة الاسمنت بالإضافة إلى أربعة مخرجات مع الاسمنت والتي تعتبر متغيرات معتمدة ، بذلك سيكون هناك أربع متغيرات تنبؤية (مدخلات) وأربع متغيرات معتمدة (مخرجات) ، وكما موضحة بالجدول الأتي :

جدول (1) المتغيرات التنبؤية والمتغيرات المعتمدة لعملية إنتاج الاسمنت

المتغيرات المعتمدة	المتغيرات التنبؤية
Y_1 : كمية إنتاج الاسمنت / طن	X_1 : الطاقة الكهربائية /
Y_2 : إنتاج الكلنكر / طن	X_2 : النفط / طن
Y_3 : النفايات الصلبة / طن	X_3 : الحجر / طن
Y_4 : كميات الغبار المنبعثة / م^3	X_4 : الطين / طن

1- أسلوب المربعات الصغرى الاعتيادية (OLS):

تم تحليل البيانات أعلاه باستخدام أسلوب المربعات الصغرى الاعتيادية بانحدار كل متغير معتمد على حدا مع بقية المتغيرات التنبؤية أي انه تم تحليل انحدار $(y_1/x_1,x_2,x_3,x_4)$ و $(y_2/x_1,x_2,x_3,x_4)$ و $(y_3/x_1,x_2,x_3,x_4)$ و $(y_4/x_1,x_2,x_3,x_4)$ وكانت نتائج التحليل بالشكل التالي :

1-تحليل انحدار $(y_1/x_1, x_2,x_3,x_4)$:

$$Y_1 = -0.000736 + 1.82\,X_1 - 0.72\,X_2 - 0.108\,X_3 \qquad \ldots..(4)$$

جدول (2) تحليل التباين لإنتاج الاسمنت

Source	D.F	S.S	M.S	F	P
Regression	3	5.5079	1.836	9180	0.007
Residual	1	0.0002	0.0002		
Total	4	5.508			
R^2 = 100 %	R^2(adj) = 100 %				

جدول (3) قيم معلمات الانحدار b والــ VIF والمختبر الإحصائي T

Predictor	D.F	Coef.	SE Coef.	T	P	VIF
Constant	--	–0.000736	0.00796	– 0.09	0.94	---
X_1	1	1.824	2.022	0.90	0.53	145523.8
X_2	1	– 0.72	1.858	– 0.39	0.765	123504.4
X_3	1	– 1083	0.171	– 0.63	0.641	963.4

2-تحليل انحدار $(y_2/x_1,x_2,x_3,x_4)$:

$$Y_2 = 0.0000 + 0.0000\ X_1 - 0.0000\ X_2 + 1.00\ X_3 \qquad .....(5)$$

جدول (4) تحليل التباين لإنتاج الكلنكر

Source	D.F	S.S	M.S	F	P
Regression	3	5.0993	1.6998	*	*
Residual	1	0.0000	0.0000		
Total	4	5.0993			
R^2 = 100 %	R^2(adj) = 100 %				

جدول (5) قيما معلمات الانحدار b والـــ VIF والمختبر الإحصائي T

Predictor	D.F	Coef.	SE Coef.	T	P	VIF
Constant	--	0.0000	0.0000	*	*	---
X_1	1	0.0000	0.0000	*	*	145523.8
X_2	1	– 0.0000	0.0000	*	*	123504.4
X_3	1	1.000	0.0000	*	*	963.4

3-تحليل انحدار $(y_3/x_1,x_2,x_3,x_4)$:

$$Y_3 = -\ 0.00000414 + 0.00385\ X_1 - 0.00356X_2 + 1.00\ X_3 \qquad .....(6)$$

جدول (6) تحليل التباين لكميات النفايات الصلبة

Source	D.F	S.S	M.S	F	P
Regression	3	5.0994	1.6998	169980	0.000
Residual	1	0.0000	0.00001		
Total	4	5.0994			
R^2 = 100 %	R^2(adj) = 100 %				

جدول (7) قيم معلمات الانحدار b والـــ VIF والمختبر الإحصائي T

Predictor	D.F	Coef.	SE Coef.	T	P	VIF
Constant	--	-0.00000414	0.00002868	-0.14	0.909	---
X_1	1	0.003846	0.007281	0.53	0.691	145523.8
X_2	1	– 0.00356	0.006691	-0.14	0.689	123504.4
X_3	1	0.999708	0.000616	1623.58	0.000	963.4

4-تحليل انحدار $(y_4/x_1,x_2,x_3,x_4)$:

$$Y_4 = -0.431 + 152\ X_1 + 141X_2 + 11.00\ X_3 \qquad .....(7)$$

جدول (8) تحليل التباين لكميات الغبار المنبعثة

Source	D.F	S.S	M.S	F	P
Regression	3	1.424	0.4747	1.66	0.505
Residual	1	0.2833	0.2853		
Total	4	1.7093			
R^2 = 100 %	R^2(adj) = 100 %				

جدول (9) قيم معلمات الانحدار b والـــ VIF والمختبر الإحصائي T

Predictor	D.F	Coef.	SE Coef.	T	P	VIF
Constant	--	-0.4306	0.3420	-1.26	0.427	---
X_1	1	-151.78	86.83	-1.75	0.331	145523.8
X_2	1	141.13	79.80	1.77	0.328	123504.4
X_3	1	11.011	7.343	1.50	0.374	963.4

من خلال ملاحظة تحليل الانحدار في الجدول رقم (3) ، (5) ، (7) ، (9) تم حذف المتغيــر الرابع (X_4) والسبب في ذلك يعود إلى قوة التداخل الخطي لهذا المتغير مع بقيــة المتغيــرات التنبؤية مما يعطي ضعف للمتغيرات المدروسة علما أن المتغير الرابع المتمثل بــالطين هــو مادة أولية مهمة في إنتاج الاسمنت ولا يمكن إنتاج الاسمنت بدون هذه المادة .

كما أظهرت النتائج إن إشارات المعلمات X_2 و X_3 للنموذج الأول سالبة وهي منافية للعلاقــة المنطقية الصحيحة بين كل من المتغيرين الحجر والطين مع المتغير المعتمد لإنتاج الاســمنت والتي هي علاقة طردية، ونفس الاختلاف ظهر في النموذج الثاني بين المتغيــر X_2 والتغيــر المعتمد Y_2 إذ ظهرت العلاقة عكسية وهذا مناف للعلاقة الصحيحة بين المتغيــرين ، وكــذلك للنموذج (6) حيث ظهرت العلاقة عكسية بين المتغير الثاني X_2 (النفط) والمتغير المعتمد Y_3 (النفايات الصلبة) ، وكذلك الحال في النموذج (7) إذ ظهرت العلاقة ايضاً عكسية بين المتغير الأول X_1 (الطاقة الكهربائية) والمتغير المعتمد Y_4 (كميات الغبار المنبعثة) .

ويعزى وجود العلاقات الغير منطقية الى وجود مشكلة تعدد علاقة خطيـة بـين المتغيـرات التنبؤية والذي أدى بدوره إلى تشويه قيم معلمات الانحدار وابتعادها عـن التفـسير العلمـي الصحيح.

وللتأكيد على وجود تداخل خطي بين المتغيرات التنبؤية فقد تم اسـتخدام بعـض الأسـاليب للكشف عنها منها :

1-1 – استخدام عامل تضخم التباين variance Inflation Factors (VIF)

من ملاحظة قيم عوامل تضخم التباين نجد أنها مرتفعة جداً ولقد بـين (Marquardt,1970) على وجود تعدد العلاقة الخطية بين المتغيرات التنبؤية في حالة كون قـيم عوامـل تـضخم التباين VIF اكبر من (4) أو (10) ومن خلال التحليل السابق يلاحـظ أن هـذه القـيم جـداً مرتفعة وهذا يدل على وجود تداخل خطي كبير بين المتغيرات التنبؤية .

1-2 – إيجاد محدد المصفوفة $|x'x|$:

اقترح هذا المعيار من قبل (Mason & Webster ; 1975) إذ ينص على انه إذا كـان $|x'x|=0$ فهذا يدل على وجود تداخل تام بين المتغيرات التنبؤية أما إذا كان $|x'x|=1$ فهـذا يدل على استقلالية المتغيرات التنبؤية فيما بينها .

ومن خلال القيم الذاتية للمصفوفة $(x'x)$ تم حساب محدد المـصفوفة $(x'x)$ والـذي كـان يساوي :

$$|x'x| = \prod_{j=1}^{p} \lambda_j = 0.000000039$$

وهذه القيمة صغيرة جداً وقريبة من الصفر وهذا دليل آخر على وجود تداخل خطي كبير بين المتغيرات التنبؤية .

1-3 – استخدام عناصر مصفوفة الارتباط الواقعة خارج القطر :

اقترح هذا الأسلوب من قبل (Gunst & Mason ; 1980) إذ نلاحظ من خـلال مـصفوفة الارتباط $(x'x)$ والموضحة أدناه :

$$(x'x)=\begin{vmatrix} 1.00 & 0.999 & 0.98 & 0.984 \\ 0.999 & 1.00 & 0.982 & 0.982 \\ 0.984 & 0.982 & 1.00 & 1.00 \\ 0.984 & 0.982 & 1.00 & 1.00 \end{vmatrix}$$

نلاحظ أن هناك ارتباط شبه تام بين المتغيرات (X_1,X_2) وبين (X_1,X_3) وبين (X_1,X_4) كذلك بين (X_2,X_3) و (X_2,X_4) وكذلك بين (X_3,X_4) مما يدل على وجود تداخل خطي كبير بين هذه المتغيرات والذي يؤدي إلى عدم دقة نتائج التحليل .

كذلك نلاحظ كدليل آخر على التداخل الخطي بين المتغيرات التنبؤية هو قيمة R^2 او معامل التحديد حيث ظهرت كبيرة جداً مع انه تم حذف المتغير الرابع من النموذج وهذا مخالف للمنطق إذ أن المتغير الرابع هو مادة أساسية في إنتاج الاسمنت .

ولذلك ومن اجل رفع التداخل الخطي ومن ثم الحصول على نتائج منطقية من دون حذف أي متغير تنبؤي مهم فقد تم اللجوء إلى أسلوب المربعات الصغرى الجزئية (PLS).

2 – أسلوب المربعات الصغرى الجزئية (PLS) :

كما تم توضيح أسلوب المربعات الصغرى الجزئية في الجانب النظري وما تتميز به من حيث إمكانية تحليل البيانات لمصفوفة من المتغيرات المعتمدة مع مصفوفة من المتغيرات التنبؤية في آنٍ واحد فقد تم إجراء هذا التحليل باستخدام هذا الأسلوب وكانت النتائج كم يلي :

جدول (10) معلمات انحدار المربعات الصغرى الجزئية

	Y_1	Y_2	Y_3	Y_4
Constant	0.016967	- 0.013546	- 0.013538	- 0.806517
X_1	0.250112	0.240652	0.240652	0.021208
X_2	0.249182	0.239757	0.239758	0.021129
X_3	0.259901	0.250071	0.250071	0.022038
X_4	0.259901	0.250071	0.250071	0.022038

جدول (11) تحليل التباين لإنتاج الاسمنت

Source	D.F	S.S	M.S	F	P
Regression	1	5.46298	5.46297	363.79	0.000
Error	3	0.04505	0.01502		
Total	4	5.50803			

جدول (12) تحليل التباين لإنتاج الكلنكر

Source	D.F	S.S	M.S	F	P
Regression	1	5.05755	5.05757	363.33	0.000
Error	3	0.04177	0.01392		
Total	4	5.09933			

جدول (13) تحليل التباين لكميات النفايات الصلبة

Source	D.F	S.S	M.S	F	P
Regression	1	5.05758	5.05759	363.07	0.000
Error	3	0.04178	0.01393		
Total	4	5.09936			

جدول (14) تحليل التباين لكميات الغبار المنبعثة

Source	D.F	S.S	M.S	F	P
Regression	1	0.03928	0.039279	0.07	0.808
Error	3	1.67004	0.556679		
Total	4	1.70931			

يلاحظ من نتائج تحليل انحدار المربعات الصغرى الجزئية عودة المتغير التنبؤي الرابع (X_4) والمتمثل بمادة الطين إلى التحليل مما يشير إلى إزالة التداخل الخطي بين المتغيرات التنبؤية ومما يؤكد ذلك ظهور العلاقات الطردية (المنطقية) بين المتغيرات التنبؤية والمتغيرات المعتمدة والتي ظهرت واضحة في الجدول (10) .

الاستنتاجات :

1 – تميزت طريقة المربعات الصغرى الجزئية (PLS) عن طريقة المربعات الصغرى الاعتيادية (OLS) بإمكانيتها في تحليل مصفوفة من المتغيرات المعتمدة مع مصفوفة من المتغيرات التنبؤية في آنٍ واحد .

2 – أظهرت طريقة المربعات الصغرى الاعتيادية (OLS) ضعفاً بالتحليل نتيجة صغر عدد المشاهدات على الرغم من إمكانية هذه الطريقة من تحليل أي بيانات بشرط أن يكون عدد

المشاهدات اكبر من عدد المتغيرات التنبؤية بمقدار واحد ، أما طريقة المربعات الصغرى الجزئية (PLS) تميزت بإمكانيتها بتحليل بيانات حتى وان كان عدد مشاهداتها اقل من عدد متغيراتها التنبؤية وهذا ما جاء في الجانب النظري وأكده الجانب العملي من البحث .

3 – نتيجة للتداخل الخطي الكبير بين المتغيرات التنبؤية عجزت طريقة المربعات الصغرى الاعتيادية (OLS) عن إجراء تحليل الانحدار إلا بعد حذف المتغير التنبؤي الرابع (X_4) والمتمثل بمادة الطين والتي هي مادة أساسية في عملية إنتاج الاسمنت ، وبذلك كانت نتائج تحليل هذه الطريقة غير منطقية ، وكذلك العلاقات العكسية بين (الحجر والاسمنت) و(والنفط والاسمنت) في حين أعادت طريقة المربعات الصغرى الجزئية (PLS) المتغير الرابع (X_4) إلى التحليل وظهور العلاقات الطردية بين المتغيرات والتي تمثل علاقات منطقية مما يدل إلى دقة ومنطقية النتائج المستحصل عليها باستخدام هذه الطريقة مقارنة بطريقة المربعات الصغرى الاعتيادية (OLS) .

المصادر:

1-Dante M. Pirouz ,(2006), "An Overview of Partial Least Squares " ,The Paul Merage School of Business University of California , Irvine, pp(1-15).
2-Frank , I. E. and Friedman ,J. H. (1993)."A Statistical View of Chemometrcs Regression Tools". Technometrics ,35 ,pp(109-148).
3-Gunst ,R. F. and Mason ,R. L. (1980),"Regression Analysis and its Applications", Marcel Dekker ,Inc. New York, U.S.A.
4-Harald , M. and Luis , I. (1986)," Partial Least Squares Regression on Design Variables As an Alternative to Analysis of Variance, Analytica Chimica Act, 191, pp(133-148).
5-Marquardt , D. W. (1970) ,"Generalized Inverse , Ridge Regression ,Biased Linear Estimation and Nonlinear Estimation" ,Technometrics ,Vol. 12 ,pp(591-612).
6-Mason ,R. L. ,Gunst ,R. F. and Webster , I. T.,(1975),"Regression Analysis and Problems of Multicollinearity ",Comm. In statistics .VOL.(4) No.(3).pp(277-292).
7-Saikat Maitra and Jun Yan ,(2008),"Principle Component Analysis and Partial Least Squares: Two Dimension Reduction Techniques for Regression" , casualty Actuarial Society , pp(79-90).
8-Wold , H.(1966)."La Regression PLS. Paris : Technip Iterative Least Squares" .In P.R. Kishnaiaah (Ed). Multivariate Analysis ,New York: Academic Press .pp(391-420).