بحوث مستقبلية (١) ٢٠٠٠
ص ص (١١٩-١٣٥)

مقارنة بعض طرائق الانحدار الخطي المتعدد بالاعتماد على القيم والمتجهات الذاتية لمصفوفة الارتباط

الدكتور موفق محمد توفيق القصاب(*)
سالم مؤيد شكر(**)

الملخص

تم ايجاد مقدرات المربعات الصغرى والتحليل الذاتي والحرف لمعلمات الانحدار بدلالة القيم والمتجهات الذاتية لمصفوفة الارتباطات البسيطة بين المتغير المعتمد والمتغيرات التنبؤية فضلاً عن تقديم علاقة رياضية عامة تربط هذه المقدرات حيث يتغير شكل المعادلة بتغير الوزن الذي يعتمد على نوع المقدر، وتمت دراسة هذه المقدرات من خلال خواصها التي تم ايجادها بدون قيود وذلك باستخدام تجزئة المصفوفات.

المقدمة

يعتبر أسلوب انحدار الحرف من أكثر الأساليب استخداماً وشيوعاً حيث اقترح هذا الأسلوب من قبل Hoerl & Kennard(1970-a) ، وطبق من قبل الكثير من الباحثين أمثال Marquardt & Snee (1975), Gunst & Mason(1977)، القصاب وهشام (١٩٩٤)، القصاب وسعيد (١٩٩٧).

لقد أطلق اسم انحدار الحرف (Ridge Regression) من قبل العالم Hoerl وذلك لكونه مشابهاً لطريقة رياضية مستخدمة في تحليل الحرف لوصف وتمثيل خواص الدرجة الثانية للاستجابة لمعاملات تابعة لمتغيرات كثيرة.

ان انحدار الحرف هو أسلوب لتعديل المربعات الصغرى عند وجود تداخل خطي بين المتغيرات التنبؤية فضلاً عن أنها من الطرائق المميزة اذ تعتمد على إضافة الثابت (k)

---

* استاذ- قسم الاحصاء- كلية الادارة والاقتصاد/ جامعة الموصل.
** قسم الاحصاء - كلية الادارة والاقتصاد/ جامعة الموصل.
قبل للنشر في ١٩٩٩/٩/١٤

الى عناصر القطر الرئيس للمصفوفة X X (X مصفوفة المتغيرات التنبؤية) قبـــل أخـــذ المعكوس لها، يتم اختيار (k) بعدة أساليب وبذلك يعرف انحدار الحرف على انه دالة لعامل الحرف k (Shrinkage factor) او لمعلمة الحرف، وان اختيار قيمة مناسبة لـــ k ســـوف يؤدي الى اعطاء تقديرات أكثر دقة من تقديرات المربعات الصغرى، اذ ان اضافـــة هـــذا الثابت سيؤدي الى تقليل عناصر القطر الرئيسي للمصفوفة $(X X)^{-1}$ وبهذه الطريقـــة يتـــم تقليل قيم المقدرات والتباينات المتضخمة، كذلك نحصل على مجموع مربعات الخطأ الكليـــة (TSE) أقـــل ممـــا عليـــــه للمربعــــات الصغـــــرى، وقـــد أكـــد كـــل مـــن Hoerl & Kennard(1970-a) عدم ايفاء المربعات الصغرى بالغرض عند وجود تداخـــل خطي بين المتغيرات، حيث استخدما مقياس الحد الأدنى لوسط البعد بين متجه المعلمـــــات الحقيقية β ومقدرة $\hat{\beta}$ الذي يأخذ الشكل الآتي $\sigma^{2}/$ lmin (σ^2 يمثل تباين الخطأ العشـــوائي و lmin انحدار أصغر الجذور المميزة للمصفوفة (X X). ان هذا المقدار فسوف يكـــون كبيرا في حالة وجود تداخل خطي بين المتغيرات أي سوف تكون المسافة كبيرة بيـــن $\underline{\beta}$ و $\hat{\underline{\beta}}$ وهذا لايفي بالغرض. كما اقترحا مقياس الطول لمتجه المعاملات من خـــلال lmin اذ قارناه بالواحد الصحيح فوجدا ان1> lmin عندما تكون المتغيرات متعددة العلاقة الخطيـــة مما يدل على ان طول متجه المعاملات طويل بالمعدل. أما أسلوب التحليل الذاتي فقد أقترح وطور من قبل كل من Webster & Mason (1974), Hawkins (1973) علـــى التوالـــي بديلا لطريقة المربعات الصغرى نتيجة لظهور مشكلة التداخل الخطـــي بيـــن المتغـــيرات التنبؤية، حيث يؤدي هذا التداخل الى ظهور تشوهات وارباك فـــي مقـــدرات المربعـــات الصغرى.

كما يمكن وصف أسلوب التحليل الذاتي (شاكر ١٩٩٩) على أنه تعديل او تحسين لأسلوب المكونات الرئيسة من حيث امتلاك او عدم امتلاك المتجهات الذاتية لقيم تنبؤيـــة، ويتم ذلك من خلال اضافة متجه مشاهدات المتغير المعتمد Y الـــى مصفوفـــة مشـــاهدات المتغيرات المستقلة X اذ تتكون لدينا المصفوفة A=(Y:X) ذات بعد ((nx(p+1)) حيـــث ان (n: حجم الصف، p: عدد المتغيرات المستقلة) وبمتغيرات قياسية كمـــا ان المصفوفـــة (A A) هي مصفوفة الارتباط لكل من المتغيرات المستقلة والمتغير المعتمد.

ان القيم الذاتية للمصفوفة $(A\ \bar{A})$ فسوف تكون $(\lambda_0>\lambda_1>..>\lambda_p)$ بمتجهات ذاتية مقابلة لها $(\underline{\gamma}_0,\underline{\gamma}_1,\ldots\ldots\ldots,\underline{\gamma}_p)$ حيث ان $(\gamma_j'=\gamma_{oj},\gamma_{1j},\ldots\ldots,\gamma_{pj})$ يرمز للمتجه الذي يحوي على عناصر المتجه $(\underline{\gamma}_j)$ ما عدا العنصر الأول $(\underline{\gamma_{0j}})$ بـ $(\underline{\gamma}^0{}_j)$ أي ان :

$$\underline{\gamma}_j^{0'}=(\gamma_{1j}\quad \gamma_{2j}\ldots\ldots\gamma_{pj})$$

عندئذ فأن:

$$\underline{\gamma}_j'=(\gamma_{oj}\downarrow\underline{\gamma}_j^{0'})$$

الغرض من إضافة المتغير المعتمد Y الى مصفوفة المتغيرات المستقلة X هو إظهار تأثير المتغير المعتمد على التداخلات الخطية في المصفوفة A لذلك فأن تأثير العنصر (γ_{oj}) على Y ليس من المعتاد ان يكون صغيرا، من هنا يبرز دور المتغير المستقل في قيمته التنبؤية. فاذا كانت قيمة ذاتية λ_j وقيمة العنصر الأول من المتجه الذاتي المقابل للقيمة الذاتية (γ_{oj}) صغيرتين كلتيهما $(\gamma_{oj}\approx o,\lambda_j\approx o)$ ، فهذا يعني ان المتجه الذاتي يعطي تعدد علاقة خطية غير تنبؤية للمتغير المعتمد، كذلك نستطيع تحديد أي المتغيرات التي تعاني من تداخل خطي وذلك من خلال معرفة عناصر المتجه الذاتي التي تكون قيمتها كبيرة حيث يمكن توضيح ذلك من خلال المعادلة الآتية:

$$\gamma^{\circ}{}_jY+\gamma_{1j}X_1+\cdots+\gamma_{pj}X_p\approx 0$$

اذ نلاحظ من المعادلة اعلاه انه في حالة كبر أي من عناصر المتجه الذاتي $(\underline{\gamma}_j^{\circ})$ فأن هذا دليل على وجود تعدد علاقة خطية ما بين ذلك المتغير والمتغيرات الأخرى، فعلى سبيل المثال اذا كانت قيمة العنصر (γ_{pj}) كبيرة نوعا ما فأن هذا دليل على ان المتغير X_p يمتلك تعدد علاقة خطية مع بقية المتغيرات.

يتضمن المبحث الثاني ايجاد مقدرات معلمات الانحدار اعتمادا على طريقة المربعات الصغرى بدلالة القيم والمتجهات الذاتية للمصفوفة $(A\ \bar{A})$ وكذلك مقدرات الحرف فضلا عن دراسة العلاقة بين هذه الطرائق من خلال المقدرات وخواصها فضلا عن ربط مقدرات الطرائق الثلاث بعلاقة واحدة من خلال وزن تتغير قيمته حسب كل طريقة من هذه الطرائق.

أما المبحث الثالث فأنه يتضمن الجانب التطبيقي حيث تضمن تطبيق الجانب النظري لبيانات تم الحصول عليها من معمل السكر والخميرة في الموصل للمدة (١٩٦٠-١٩٧٥) .

(٢) مقدرات معلمات الانحدار

سوف يتم ايجاد معلمات الانحدار بطرائق المربعات الصغرى بدلالة القيم والمتجهات الذاتية للمصفوف ($A \bar{A}$) ومقدرات الحرف وايجاد العلاقة بين هذه الطرائق من خلال المقدرات وخواصها فضلا عن ربط المقدرات المذكورة بعلاقة واحدة (للمزيد من المعلومات وتفاصيل طريقة الاشتقاق أنظر شاكر (١٩٩٨)).

(٢-١) اسلوب التحليل الذاتي:

يمكن تقدير معلمات الانحدار بالاعتماد على اسلوب القيم والمتجهات الذاتية من خلال استخدام المعادلات التنبؤية حيث اتبع هذا الأسلوب Webster, et al (1974) وبالنظر لعدم وجود التوزيع الاحتمالي لمقدرات القيم الذاتية وذلك لكونها دالة في القيم والمتجهات الذاتية للمصفوفة ($A \bar{A}$) فقد تم ايجاد خواص المقدرات من دون قيود من خلال استخدام اسلوب تجزئة المصفوفات، حيث ان المعادلة التنبؤية سوف تكون بالصيغة:

$$\hat{\underline{y}} = y\,\underline{I} - \eta\, X \sum_{j=0}^{p} aj\, \underline{\gamma}^{\circ}_{j}$$

وهي نفس المعادلة التنبؤية لطريقة المربعات الصغرى، حيث أن

$$\hat{\beta} o = \overline{Y^{*}}, \qquad \hat{\underline{\beta}} = \underline{\ }^{-}\eta \sum_{j=0}^{p} aj \ \underline{\gamma}^{\circ}_{j}$$

كما أن

$$\eta^2 = \sum_{i=1}^{n} (\gamma_i - \bar{y})$$

$\bar{Y}^*$ هو المتوسط للمتغير المعتمد للمشاهدات الأصلية

وبذلك فأن مقدرات المربعات الصغرى بدلالة القيم والمتجهات الذاتية للمصفوفة $(A\ \bar{A})$ هي:

$$\hat{\underline{\beta}}_{LS} = -\eta \frac{\sum_{j=0}^{p} \gamma_{oj}\ \gamma_j^{o}\ \lambda_j^{-1}}{\sum_{j=0}^{p} \gamma^2{}_{oj}\ \lambda_j^{-1}} \qquad \text{.....................................}(2.2)$$

أما مجموع مربعات البواقي فسوف يأخذ الشكل

$$SSE(X1, X2, \ldots, Xp) = \eta^2 (\sum_{j=0}^{p} (\lambda^{*-1}{}_j\ )^{-1} \qquad \text{..........}(2.3)$$

حيث ان

$$\lambda_j^{*}\ {}_j = \lambda_j /\ {}^2{}_{oj}$$

واذا ما أردنا الحصول على مقدرات القيم الذاتية فأننا سوف نحذف القيم والمتجهات الذاتية غير التنبوية من المعادلة (2.1) فعلى سبيل المثال اذا كان لدينا -p ($\bar{K}$+1) من مقدرات العلاقة الخطية غير التنبوية في المتغير المعتمد فأن $\gamma_{oj} = 0$ لجميع قيم $j = \bar{K}+1,\ \bar{K}+2..,p$

وان مقدرات القيم الذاتية سوف تأخذ الصيغة

$$\hat{\beta}_{EV} = -\eta \frac{\sum_{j=0}^{k} \gamma_{oj}\ \gamma^{o}{}_j / \lambda_j}{\sum_{j=0}^{k} \gamma^2{}_{oj} / \lambda_j} \qquad \ldots\ldots\ (2.4)$$

ومجموع مربعات خطأ

$$SSE(X1,X2.,...X_p)=\eta^2(\sum_{j=0}^{k'}\lambda_j^{*-1} \quad ....(2.5)$$

غير ان المشكلة الرئيسة هي كيفية اختيار القيم والمتجهات الذاتية التي سوف تحذف ولمعالجة هذه المشكلة فقد اقترح كل من Webster, et al (1974) القيمة التي يمكن ان يعول عليها في حذف القيم الذاتية هي λ_j<0.05 ، أما المتجه المقابل لهذه القيم فيحكم عليه تنبؤي او غير تنبؤي (لايحذف او يحذف) من خلال القيمة الأولى لذلك المتجه فاذا كانت ($\gamma_{0j} \prec 0.1$) فهذا يعني ان المتجه الذاتي ($\underline{\gamma}_j$) غير تنبؤي وبعكسه فيعتبر تنبؤياً.

كما اقترح كل من Fisher & Mason (1981) القيم التي يعول عليها $|\gamma_{0j}| \prec 0.1$ λ_j<0.1, فضلا عن امكانية تحديد المتغيرات المتداخلة خطيا من خلال القيم المقابلة للمتغيرات التنبؤية الموجودة في المتجه المقابل للقيم الذاتية الصغيرة، حيث تعتبر هذه المتغيرات متداخلة خطيا اذا كانت القيم المقابلة لها أكبر او مساوية لـ 0.4 في المتجه الذاتي المقابل للقيمة الذاتية الصغيرة.

سوف نعول على القيمة λ_j<0.05 في ترشيح القيم الذاتية الصغيرة للحذف والقيمة ($|\gamma_{0j}| \prec 0.1$) لحذف المتجهات الذاتية من حيث كونها تنبؤية او غير تنبؤية، والقيم المقابلة للمتغيرات التي تكون أكبر من 0.4على انها متعددة العلاقة الخطية.

(٢-٣) اسلوب الحرف:

يمكن اضافة الثابت k الى عناصر قطر المصفوفة ($A'A$) بدلا من المصفوفة X X وبذلك نستطيع ان نبين ان أسلوب الحرف يمكن صياغته على أساس القيم والمتجهات الذاتية لمصفوفة الارتباط للمتغير المعتمد والمتغيرات المستقلة ($A'A$) كما يمكن ان نقدر معلمات الحرف بدلالة القيم والمتجهات الذاتية للمصفوفة ($A'A$) فان هذه المصفوفة تأخذ الشكل:

$$A'A = \begin{pmatrix} \underline{Y}'\underline{Y} & \underline{Y}X \\ X'\underline{Y} & X'X \end{pmatrix}$$

وباضافة الثابت k الى قطرها الرئيس نحصل على

$$A'A + KI_{P+1} = \begin{pmatrix} \underline{Y}'\underline{Y}+K & \underline{Y}'X \\ X'\underline{Y} & X'X+KI_p \end{pmatrix}$$

ان المتجهات الذاتية لا تتأثر باضافة الثابت k الى عناصر قطر المصفوفة $A'A$ وانما فقط القيم الذاتية هي التي تتأثر لذا فأن مقدرات الحرف بدلالة القيم والمتجهات الذاتية سوف تكون بالصيغة (شاكر (١٩٩٨)).

$$\hat{\beta}_R = -\eta \frac{\sum_{j=0}^{p} \gamma_{oj}\, \gamma^{o}{}_{j} /(\lambda_j + k)}{\sum_{j=0}^{p} \gamma^2{}_{oj} /(\lambda_j + k)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.6)$$

وبمجموع مربعات خطأ

$$SSE\,(X1, X2., \ldots X_p) = \eta^2 \left(\sum_{j=0}^{p} \frac{\gamma^2{}_{oj}}{(\lambda_j + k)}\right)^{-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.7)$$

يمكن ربط (2.2) و(2.4) و(2.6) بمعادلة عامة واحدة هي:

$$\hat{\beta} = -\eta \frac{\sum_{j=0}^{p} W_j \gamma_{oj}\, \gamma^{o}{}_{j} / \lambda_j}{\sum_{j=0}^{p} W_j\, \gamma^2{}_{oj} / \lambda_j} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.8)$$

اذ نحصل على مقدرات المربعات الصغرى في حالة $W_j=1$ لجميع قيم j ونحصل على مقدرات القيم الذاتية من خلال حذف المتجهات الذاتية التي تعرف تعدد العلاقة الخطية غير التنبؤية وذلك بجعل $W_j=0$ للمتجهات المحذوفة و $W_j=1$ لبقية المتجهات كما نحصل على مقدرات الحرف بجعل $W_j = \frac{\lambda_j}{\lambda_j + k}$.

ان انحدار الحرف سوف يعطي أوزانا قليلة للمتجهات التي تشير الى وجود تعدد علاقة خطية في المصفوفة A A وذلك باضافة الثابت k الى λ_j التي سوف تضخم قيمة المقام ومن ثم تقلل أوزان تلك المتجهات المقابلة للقيم الذاتية الصغيرة التي تشير الى تعدد العلاقة الخطية.

(٣) خواص مقدرات معلمات الانحدار

يمكن ايضا توضيح العلاقة بين انحدار الحرف والتحليل الذاتي من خلال خواص المقدرات اذ يمكن توضيح ذلك من خلال التوقع الرياضي والتباين.

(٣-١) التوقع الرياضي:

لايجاد التحيز في مقدرات القيم الذاتية نفرض وجود تداخل خطي بين المتغيرات التنبؤية وهذا التداخل غير تنبؤي، بذلك سوف تكون لدينا مقدرات القيم الذاتية متمثلة بالمعادلة (2.4) اذ أنها عبارة عن مقدرات المربعات الصغرى محذوفا منها المتجهات والقيم الذاتية غير الممتلكة لقيم تنبؤية، لذلك يمكن كتابة المعادلة (2.4) بالصيغة:

$$\hat{\underline{\ }}_{EV} = \hat{\underline{\ }}_{LS} - \left(-\eta \frac{\sum_{j=K+1}^{p} \gamma_{oj}\, \underline{\gamma_j}^{\circ} / \lambda_j}{\sum_{j=K'+1}^{p} \gamma^2_{oj} / \lambda_j}\right)$$

$$= -\eta \frac{\sum_{j=0}^{k} \gamma_{oj}\, \underline{\gamma}^{\circ}_j / \lambda_j}{\sum_{j=0}^{k} \gamma^2_{oj} / \lambda_j}$$

وبأخذ التوقع الرياضي لطرفي المعادلة نحصل على:

$$E(\hat{\underline{\beta}}_{EV}) = \underline{\beta} + \eta \frac{\sum_{j=K+1}^{p} \gamma_{oj}\, \underline{\gamma}^{\circ}_j / \lambda_j}{\sum_{j=K'+1}^{p} \gamma^2_{oj} / \lambda_j}$$

لذا فأن مقدار التحيز في مقدرات القيم الذاتية هو :

$$b(\hat{\underline{\beta}}_{EV}) = \eta \frac{\sum_{j=K'+1}^{p} \gamma_{oj}\, \underline{\gamma}^{o}{}_{j} / \lambda_{j}}{\sum_{j=K'+1}^{p} \gamma^{2}{}_{oj} / \lambda_{j}} \quad ....(3.1)$$

ولإيجاد التحيز في مقدرات الحرف نحصل بعد أخذ التوقع الرياضي لمعادلة مقدرات الحرف (شاكر ١٩٩٨) على:

$$E(\hat{\beta}_{iR}) = \beta_i - K\left(\sum_{j=0}^{p} (\lambda_j + k)^{-1} \gamma_{ij}^{2} - \frac{\sum_{j=0}^{p} \gamma_{0j} \gamma_{ij} / \lambda_j + k)^2}{\sum_{j=0}^{p} \gamma_{0j}^{2} / (\lambda_j + k)}\right) \beta_i$$

أي أن مقدار التحيز للمعلمة $\hat{\underline{\beta}}_i$ سوف يكون

$$b(\hat{\beta}_{iR}) = -K\left(\sum_{j=0}^{p} (\lambda_j + k)^{-1} \gamma^{2}{}_{ij} - \frac{\sum_{j=0}^{p} \gamma_{0j} \gamma_{ij} / \lambda_j + k)^2}{\sum_{j=0}^{p} \gamma_{0j}^{2} / (\lambda_j + k)}\right) \beta_i \quad .....(3.2)$$

(٣-٣) التباين

إن تباين مقدرات المربعات الصغرى للمعلمة β_I بدلالة القيم والمتجهات الذاتية هو:

$$V(\hat{\beta}_{iLS}) = \sigma^2 \left(\sum_{j=0}^{p} \lambda_j^{-1} \gamma_{ij}^{2} - \frac{\left(\sum_{j=0}^{p} \lambda_j^{-1} \gamma_{0j} \gamma_{ij} \right)^2}{\sum_{j=0}^{p} \lambda_j^{-1} \gamma_{0j}^{2}} \right) \cdots (3.3)$$

وبجعل $\gamma_{0j} = 0$ لحالة $\gamma_{0j} \approx 0$ و $\lambda_j \approx 0$ لقيم $j = k' + 1, k' + 2, \ldots, p$

فأن تباين مقدرات القيم الذاتية هو:

$$V\left(\hat{\beta}_{iEV}\right)=\sigma^{2}\left(\sum_{j=0}^{p}\lambda_{j}^{-1}\gamma_{ij}^{2}-\frac{\left(\sum_{j=0}^{k'}\lambda_{j}^{-1}\gamma_{0j}\gamma_{ij}\right)^{2}}{\sum_{j=0}^{k}\lambda_{j}^{-1}\gamma_{0j}^{2}}\right)\quad(3.4)$$

كما ان تباين مقدرات الحرف $\hat{\beta}_{iR}$ بدلالة القيم والمتجهات الذاتية (شاكر ١٩٩٨) هو:-

$$V\left(\hat{\beta}_{LR}\right)=\left(\begin{array}{l}\sum_{j=0}^{p}(\lambda_{j}+K)^{-1}\gamma_{lj}^{2}-\frac{\left(\sum_{j=0}^{p}\gamma_{0j}\gamma_{lj}/(\lambda_{j}+K)\right)^{2}}{\sum_{j=0}^{p}\gamma_{0j}^{2}/(\lambda_{j}+K)}\\ -K\left(\sum_{i=1}^{p}\sum_{j=0}^{p}\gamma_{lj}\gamma_{ij}/(\lambda_{j}+K)\right)^{2}-2\sum_{i=1}^{p}\left(\sum_{j=0}^{p}\gamma_{lj}\gamma_{oj}/(\lambda_{j}+K)\right)\\ \left(\sum_{j=0}^{p}\gamma_{ij}\gamma_{oj}/(\lambda_{j}+K)\right)\left(\sum_{j=0}^{p}\gamma_{ij}\gamma_{lj}/(\lambda_{j}+K)\right)\Big/\left(\sum_{j=0}^{p}\gamma_{oj}^{2}/(\lambda_{j}+K)\right)\\ +\frac{\sum_{i=1}^{p}\left(\left(\sum_{j=0}^{p}\gamma_{lj}\gamma_{oj}/(\lambda_{j}+K)\right)\left(\sum_{j=0}^{p}\gamma_{ij}\gamma_{oj}/(\lambda_{j}+K)\right)\right)^{2}}{\left[\sum_{j=0}^{p}\gamma_{oj}^{2}/(\lambda_{j}+K)\right]^{2}}\end{array}\right)\sigma^{2}$$

..................(3.5)

(٤) الجانب التطبيقي:

تم تطبيق الجانب النظري على بيانات معمل السكر والخميرة في الموصل للمـــدة (١٩٦٠-١٩٧٥) حيث كانت المتغيرات المستقلة والمتغير المعتمد كما يلي :

Y^*: كمية السكر المنتج (طن)

X_1^*: عدد أيام العمل خلال الموسم

X_2^*: مجموع زمن التوقف في الانتاج (بالايام)

X_3^*: عدد أيام العمل الفعلية في الموسم

X_4^*: اجمالي كمية البنجر المورد الى المعمل (طن)

X_5^*: اجمالي كمية البنجر المصنع(طن)

X_6^*: اجمالي كمية الأتربة والاحراش(طن)

X_7^*: معدل انتاجية الدونم (طن/دونم)

X_8^*: توقفات بسبب عدم وجود بنجر (ساعة)

X_9^*: عطلات ميكانيكية (ساعة)

تم استخدام النموذج الخطي الآتي:

$$\underline{Y}^*=X^*\underline{\beta}+\underline{\in}$$

اذ ان Y^* متجه مشاهدات المتغير المعتمد الأصلية ذو بعد (٣١×١) و X^* مصفوفة مشاهدات المتغيرات الأصلية ذات بعد (٣١×١٠) وتحوي على متجه الواحدات في عمودها الأول، β متجه المعلمات ذا بعد (١٠×١) و $\underline{\in}$ متجه المتغيرات العشوائية ذا بعد (٣١×١) ، ثم تحويل المتغيرات المستقلة والمتغير المعتمد الى متغيرات قياسية وبذلك أصبح النموذج وبحسب المتغيرات التي لدينا كـ ما يأتي:-

$$Y_i=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+..+\beta_9X_{i9}+\epsilon_i \qquad i=1,2,..,31$$

(٤-١) وصف المتغيرات

تم حساب المصفوفة X X التي تمثل معاملات الارتباطات البسيطة بين المتغيرات المستقلة حيث كانت المصفوفة بالشكل الآتي:

XX'=

	X1	X2	X3	X4	X5	X6	X7	X8	X9
X1	1.0000								
X2	.3247	1.0000							
X3	.77716	-.33248	1.0000						
X4	.85669	.39964	.58528	1.0000					
X5	.85229	.41078	.57267	.99939	1.0000				
X6	.43564	.11770	.36128	.38923	.38555	1.0000			
X7	.4734	.06986	.40980	.58717	.58390	.30864	1.0000		
X8	-.76995	-.44885	-.47875	-.68305	-.67419	-.25590	-.27209	1.0000	
X9	.20459	.86247	-.37626	.36273	.37182	-.00156	.06581	-.50736	1.000

لوحظ من مصفوفة الارتباط المذكورة وجود ارتباطات منطقيـــة واخـــرى غـــير منطقية بين المتغيرات غير انها في كل الأحوال مؤثرة وتسبب ارباك التقديرات. لقـــد تـــم اختبار معنوية هذه الارتباطات والجدول الآتي يوضح القيم المعنوية فقط علما بأن قيمـــة t الجدولية عند مستوى معنوية 0.05 وبدرجات حرية (29) هي 2.045.

الجدول (١)

اختبار لمعنوية الارتباطات بين المتغيرات

Variables	t-cal	Variables	t-cal
(X_1,X_3)	6.65	(X_3,X_7)	2.41
(X_1,X_4)	8.943	(X_3,X_8)	2.93
(X_1,X_5)	8.774	(X_3,X_9)	2.18
(X_1,X_6)	2.6	(X_4,X_5)	154.1
(X_1,X_7)	2.89	(X_4,X_6)	2.27
(X_1,X_8)	-6.497	(X_4,X_7)	3.9
(X_2,X_4)	2.34	(X_4,X_8)	5.0
(X_2,X_5)	2.42	(X_4,X_9)	2.09
(X_2,X_8)	2.7	(X_5,X_6)	2.25
(X_2,X_9)	9.176	(X_5,X_7)	3.873
(X_3,X_4)	3.887	(X_5,X_8)	4.9
(X_3,X_5)	3.761	(X_5,X_9)	2.15
(X_3,X_6)	2.086	(X_8,X_9)	3.17

كما تم حساب $(X X)^{-1}$ حيث ان عناصر القطر الرئيس لهذه المصفوفة تعـــرف بعوامل تضخم التباين حيث كانت المصفوفة بالشكل الآتي:

	X1	X2	X3	X4	X5	X6	X7	X8	X9
	85.55	-51.25	-86.44	76.91	- 68.06	-.07	-1.61	-1.39	-5.75
		35.93	54.53	-25.19	23.27	-.59	1.3	2.25	1.55
			97.45	-60.1	55.04	-.45	1.6	7.15	12.18
				1346	- 1331.5	-3.5	-8.2	36.11	11.08
$(X'X)^{-1}$					1323.8	3.41	6.89	-37.15	-13.82
						1.31	-.15	-0.2	.27
							1.73	-.48	-.14
								6.33	4.99
									9.08

يلاحظ من المصفوفة المذكورة وجود تضخم كبير في عناصر القطر وبخاصة تلك التي تقابل المتغيرات X_1، X_2، X_3، X_4، X_5، وهذا يوضح وجود تعدد علاقة خطية بيـــن التغيرات التنبؤية التي تؤدي الى تضخم في تباينات المعلمات المقدرة.

وتم أيضا احتساب القيم والمتجهات الذاتية للمصفوفة (X X) (شاكر ١٩٩٨) حيث نلاحظ ان القيم الذاتية l_{10},l_9 أقل من 0.05 وهذا دليل على وجود علاقة خطية في البيانات.

كما تم حساب القيم والمتجهات الذاتية للمصفوفة A A (شاكر ١٩٩٨) وكانت قيم $\lambda_9, \lambda_8, \lambda_7$ تمتلك قيما أصغر من 0.05 وهذا يعني وجود تعدد علاقة خطية في البيانات غير ان المتجهين الذاتيين $\gamma 8,\gamma 9$ يمتلكان القيم $|\gamma 08| = .05716, |\gamma 09| = .0258$ وهي أقل من 0.1 وهذا يدل على ان قيمة تعدد العلاقة الخطية التي تعرفها القيمتان الذاتيتان λ_9, λ_8 هما غير تنبؤيتين أما القيمة $|\gamma 071| = .6262$ فانها أكبر من 0.1 وهذا يعني ان القيمة الذاتية λ_7 تعرف تعدد علاقة خطية تنبؤية.

(٤-٣) تقدير المعلمات

تم الحصول على المعلمات المقدرة بطريقة المربعات الصغرى حيث كانت قيم هذه المعلمات كما معرفة في النموذج الآتي:

$$Y_i=0.21551\ X_{i1}+.00136\ X_{i2}+.85917\ X_{i3}+2.9457\ X_{i4}-2.9839\ X_{i5}-.06877\ X_{i6}-.01029\ X_{i7}+.0898\ X_{i8}+.00713\ X_{i9} \quad (4.1)$$

كما تم الحصول على مقدرات المعلمات بطريقة القيم الذاتية بعد حذف المتجهات الذاتية غير التنبؤية وكانت قيم هذه المعلمات كما موضحة في النموذج الآتي:

$$Y_i=0.518817X_{i1}-235423X_{i2}+.4941206X_{i3}+.194121Xi4-.007135X_{i5}-.005839X_{i6}+.007662X_{i7}+.02793\ X_{i8}-.07242\ X_{i9} \quad (4.2)$$

لوحظ ان اشارات بعض المقدرات لهذه الطريقة قد تغيرت عما كانت عليه في النموذج (4.1) فمثلا اشارة β_{EV} (2) قد تغيرت من الاشارة الموجبة الى السالبة علما بأن الاشارة السالبة هي الأكثر منطقية، حيث ان مجموع زمن التوقف عن الانتاج يتناسب عكسيا مع انتاج السكر والكلام نفسه ينطبق على المقدرات β_{EV} (5) و β_{EV} (7) و β_{EV} (8) و β_{EV} (9).

كذلك تم الحصول على مقدرات للمعلمات بطريقة الحرف (الأسلوب التكراري) وكانت قيم هذه المعلمات كما موضحة في النموذج الآتي:

$$Y_i=0.195261X_{i1}-.13644X_{i2}+.289741 \quad X_{i3}+.082725X_{i4}+.07634X_{i5}-.017069X_{i6}+.0355451X_{i7}-.10411X_{i8}-.149894X_{i9} \qquad (4.3)$$

لوحظ ان اشارات المعلمات المقدرة جاءت منسجمة مع أسلوب القيم الذاتية.

(٤-٣) مقارنة المقدرات من خلال خواصها

لغرض مقارنة المقدرات التي تم الحصول عليها بالطرائق الثلاث فقد تم احتساب عوامل تضخم التباين لمقدرات الطرائق الثلاث وكما في الجدول (٢)

الجدول (٢)

VIF لمقدرات المربعات الصغرى والقيم الذاتية والحرف

estimates	VIF(LS)	VIF(EV)	VIF(R)
β_1	85.55	3.26141	0.0962
β_2	35.93	3.3753	.15397
β_3	97.47	1.9986	.10629
β_4	1346.03	.9988	.0867
β_5	1323.83	1.362203	.0898
β_6	1.31	1.29592	.26616
β_7	1.73	1.6676	.24986
β_8	6.33	3.77991	.18674
β_9	9.08	7.97265	.14373

نلاحظ من الجدول المذكور ان تباينات مقدرات المربعات الصغرى كبيرة ومتضخمة وخصوصا في معاملات المتغيرات $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9$ وهذا بسبب التداخل الخطي الموجود بين المتغيرات مما لا يمكننا الاعتماد على مقدرات المربعات الصغرى.

أما عوامل تضخم التباين لمقدرات القيم الذاتية فهو أقل مما هو عليه لحالة المربعات الصغرى حيث ان مقدر القيم الذاتية قد خفضها بشكل كبير، كذلك فقد ظهر لدينا ان مقدر الحرف هو الأفضل اذ أنه يمتلك قيم تباينات المعلمات المقدرة أقل مما هو عليه لمقدرات القيم الذاتية.

(٤-٤) الاستنتاجات والتوصيات

١. تعد طريقة ايجاد القيم والمتجهات الذاتية للمصفوفة $(A' A)$ من الطرائق المميزة في الكشف عن تعدد العلاقات الخطية حيث ان القيم الصغيرة $\lambda_j \approx 0$ تحدد العلاقة الخطية والعنصر الأول من المتجه الذاتي المقابل للقيمة الذاتية $|\approx 0$ | تحدد العلاقة الخطية غير التنبؤية. كما ان فحص عناصر القطر للمصفوفة $(X' X)$ تعد طريقة جيدة اذ توضح لنا المتغيرات التي تعاني من تضخم في تباينات معاملاتها. أما مقياس معاملات الارتباطات الزوجية فلا يمكن الاعتماد عليه.

٢. يمكن ربط مقدرات المربعات الصغرى مع مقدرات القيم الذاتية والحرف وذلك من خلال الوزن W_j اذ تتغير قيمته بحسب المقدرات، حيث تكون $W_j=1$ في حالة مقدرات المربعات الصغرى و $W_j=0$ للمتجهات الذاتية المحذوفة للحصول على مقدرات القيم الذاتية، وتكون $W_j=\lambda_j 1(\lambda_j +K)$ للحصول على مقدرات الحرف.

٣. ان التوزيع والخواص النظرية للتحليل الذاتي غير معروفة الى حد الآن لذلك لجأنا الى ايجاد خواص مقدرات القيم الذاتية من دون قيود وذلك من خلال استخدام تجزئة المصفوفات.

التوصيات:

في حالة وجود تعدد علاقة خطية قوية بين المتغيرات فأننا نوصي بعدم استخدام أسلوب الحرف اذ أنه سوف يعطي تقديرات ليست دقيقة بشكل كامل وكما لاحظنا في الجانب التطبيقي أنه بقي أثر تعدد العلاقة الخطية ظاهرا حتى بعد حذف المتغيرات المتعددة العلاقة الخطية على انفراد. واستخدام أسلوب التحليل الذاتي في حالة وجود تعدد علاقة خطية قوية يعطي تقديرات مناسبة وتنبؤا دقيقا.

المصادر

- صالح مؤيد شاكر، " العلاقات بين انحدار الحرف والتحليل الذاتي لمصفوفة الارتباط المتضخمة"، أطروحة ماجستير مقدمة إلى مجلس كلية الإدارة والاقتصاد (غير منشورة) ١٩٩٨.
- موفق محمد القصاب وهيفاء عبد الجواد سعيد، " استخدام معيار متوسط مربعات الخطأ لمقارنة مقدرات المربعات الصغرى والطرائق المتحيزة في الانحدار"، مجلة تنمية الرافدين، (١٩٩٨).
- موفق محمد توفيق القصاب ومحمد علي بسام هشام، "التقدير بطريقتي الحرف والمربعات الصغرى"، منشور في مجلة وقائع المؤتمر العلمي السابع للجمعية العراقية للعلوم الإحصائية (١٩٩٥).
- A.E Hoerl, & R.W. Kennard, (1970-a), " Ridge Regression : Biased Estimation for Non or thogonal Problems" , Technometrics, Vol.12, PP(55-67).
- D.M. Hawkins, (1973), "On the Investigation of Alternative Regression by Princilal Component Analysis", Applied Statistics, Vol. 22, pp(275-286).
- D.W. Marquardt, & R.D. Snee, (1975), "Ridge Regression in Practice", The American Statistician, Vol. 29, No.1, PP(3-20).
- J.C Fisher, & R.L. Maeson (1981), The Analysis of Multicollinear Data in Criminology", Acadimic press Inc.
- J.T. Webster, R.F Gunts. & R.L. Mason, (1974), "Latent Root Regression Analysis", Technometrics, Vol.16, No.4, pp(513-533).
- R.F Gunts & R.L. Mason, (1977), "Biased Estimation in Regression : An Evaluation Using Mean Squar Error", JASA, Vol.72, No.359, PP(616-628).