Compared Some Estimators Ordinary Ridge Regression And Bayesian Ridge Regression With Practical Application

مقارنة بعض مقدرات انحدار الحرف الاعتيادية وانحدار الحرف البيزية مع تطبيق عملي

أ.م.د. عماد حازم عبودي/ كلية الادارة والاقتصاد / جامعة بغداد
الباحث / احمد جبار حمود

OPEN ACCESS
P - ISSN 2518 - 5764
E - ISSN 2227 - 703X

Received:27/11/2018
Accepted :18/12/2018

المستخلص

تظهر مشكلة التعدد الخطي دائماً عندما يرتبط اثنان او اكثر من المتغيرات المستقلة مع بعضها البعض. وبذلك يكون فيها خرق لاحد الفروض الاساسية لطريقة المربعات الصغرى الاعتيادية مما يؤدي الى تقديرات متحيزة. وقد اقترحت عدة طرق لمعالجة هذه المشكلة منها طريقة المركبات الرئيسة لمعالجة مشكلة التعدد الخطي التام وطريقة انحدار الحرف لمعالجة مشكلة التعدد الخطي شبه التام . في هذا البحث ستتم المقارنة بين طريقة انحدار الحرف المتحيزة وطريقة انحدار الحرف غير المتحيزة وكذلك طريقة انحدار الحرف البيزية باستعمال توزيع كاما بالإضافة الى طريقة المربعات الصغرى الاعتيادية. استخدمنا اسلوب المحاكاة للمفاضلة بين الطرائق الثلاثة التي تم الاعتماد عليها وهي طريقة انحدار الحرف الاعتيادية المتحيزة وطريقة انحدار الحرف الاعتيادية غير المتحيزة بالإضافة الى طريقة انحدار الحرف البيزية واستخدم معيار متوسط مربعات الخطأ للمقارنة بين الطرائق وكانت طريقة انحدار الحرف الاعتيادية المتحيزة اعطت نتائج جيدة باستخدام احجام عينات مختلفة .

المصطلحات الرئيسة للبحث/ انحدار الحرف الاعتيادية المتحيزة، انحدار الحرف الاعتيادية غير المتحيزة، انحدار الحرف البيزية.

بحث مستل من رسالة ماجستير

1-1 **المقدمـــة** :

يعتبر الانحدار الخطي المتعدد واحداً من النماذج الخطية التي تستعمل بكثرة في تحليل بيانات العديد من البحوث في المجالات الطبية ، والصحية ، والاقتصادية ، والادارية، والاجتماعية ، والعلوم التطبيقية الاخرى . ان تفسير معادلة الانحدار الخطي المتعدد يعتمد ضمنياً على الفرضية التي تقول ان المتغيرات التفسيرية لا تكون مرتبطة مع بعضها ارتباطا قويا ، فمعامل الانحدار عادة يفسر على انه مقياس للتغير في المتغير عندما يزداد المتغير التفسيري المناظر بمقدار وحدة واحدة وان كل المتغيرات التفسيرية الاخرى تعد ثابته ، وهذا التفسير يمكن ان لا يكون صحيحاً اذا وجدت علاقة خطية قوية بين المتغيرات التفسيرية ، هذه المشكلة تدعى بمشكلة التعدد الخطي $Maulticollinearity$ بحيث يصعب فصل اثر كل متغير على المتغير المعتمد في الواقع التطبيقي ، وبذلك يتعذر تقدير معالم النموذج عندما تكون هنالك علاقة خطية تامة ما بين اثنين او اكثر من المتغيرات التفسيرية (المستقلة) ويرجع السبب في ذلك الى استحالة ايجاد معكوس مصفوفة المعلومات $X'X$() .

مشكلة التعدد الخطي تكون على نوعين ، الاول يسمى بالتعدد الخطي التام (Maulticollinearity perfect) في هذه الحالة يكون محدد مصفوفة المعلومات مساوياً الى الصفر ، وابرز طرق المعالجة هو اسلوب المركبات الرئيسية (principle components)، اما النوع الثاني يسمى التعدد الخطي شبه التام (semi Maulticollinearity) عندما تكون قيمة محدد مصفوفة المعلومات صغيرة جداً وعندها تكون المعالم المقدرة ذات تباين كبير جداً، ومن ابرز طرق المعالجة في مثل هذه الحالة هو اسلوب انحدار الحرف Ridge Regression().

أن التوسعة التي تم اضافتها من خلال بحثنا تتضمن تشخيص مشكلة التعدد الخطي شبه التام ومعالجتها باستعمال ثلاث طرائق و كالاتي الطريقة الاولى هي طريقة انحدار الحرف الاعتيادية المتحيزة Ordinary Ridge Regression اما الطريقة الثانية التي تم استعمالها هي طريقة انحدار الحرف الاعتيادية غير المتحيزة *Unb*iased Ordinary Ridge Regression فضلاً عن طريقة انحدار الحرف البيزية Bayesian Ridge regression

2-1 **الانحدار الخطي المتعدد multiple linear regression** :

يعتبر الانحدار الخطي من أوسع الطرق الاحصائية استخداماً في دراسة وتحليل العلاقة السببية بين المتغيرات المدروسة ولكافة العلوم وخصوصاً الاقتصادية منها (5:pp21) يستخدم أنموذج الانحدار الخطي البسيط لغرض تحديد العلاقة بين متغير الاستجابة (response variable) وعلاقته بمتغير توضيحي واحد (explained variable) ، ولكن هنالك دراسات تتطلب وضع متغير الاستجابة بوصفه دالة لأكثر من متغير توضيحي واحد، مثال ذلك تحليل دالة الطلب على سلعة استهلاكية معينة والتي تعتمد على متغيري الدخل و السعر وكذلك اسعار السلع البديلة وكذلك الحال عند دراسة دالة الانتاج بوصف دالة تعتمد على العمل وراس المال ، مثل هذه الدراسات يمكن استعمال الأنموذج الخطي العام فضلا عن النموذج الخطي البسيط (6:pp92). يمكن كتابة أنموذج الانحدار الخطي العام (13:pp71-72) كالاتي :

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \mathcal{E}_i \\ &= \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \mathcal{E}_i \ , \ i = 1,2,3,\ldots\ldots,n \end{aligned} \quad (1)$$

وباستخدام اسلوب المصفوفات يمكن كتابة النموذج كالاتي:

$$\underline{Y} = X\underline{\beta} + \underline{\mathcal{E}} \quad (2)$$

حيث ان:

$$\underline{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \underline{\mathcal{E}} = \begin{bmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \\ \vdots \\ \mathcal{E}_p \end{bmatrix}$$

حيث ان:

$\underline{Y}$: يمثل متجه مشاهدات متغير الاستجابة من الدرجة ($n \times 1$) .

X : يمثل مصفوفة مشاهدات المتغيرات التوضيحية من الدرجة ($((n \times p)+1)$).

$\underline{\beta}$: يمثل متجه المعلمات المجهولة من الدرجة ($(p+1) \times 1$).

p: عدد المتغيرات التوضيحية .

n : عدد المشاهدات.

$\underline{\varepsilon}$: يمثل متجه الأخطاء العشوائية من الدرجة ($n \times 1$).

3-1 مشكلة التعدد الخطي Multicollinearity problem

أحدى الافتراضات الخاصة لأنموذج الانحدار الخطي أن ليس هنالك علاقة خطية بين المتغيرات التوضيحية فاذا كانت هنالك علاقة بين هذه المتغيرات سوف تحصل مشكلة التعدد الخطي(12:pp68) ولكن والفرضية الخاصة بهذه المشكلة تعتمد على نوع الاعتماد الخطي بينها وهذا يعني هل أن العلاقة هي خطية تامة او شبه تامة بين هذه المتغيرات المستقلة وبوجود شرط أن عدد المعالم المطلوب تقديرها يجب أن يكون اقل من حجم العينة المدروسة اي ان(3:pp181):

$$\text{rank}(X) = (p+1) < n \qquad (3)$$

$$r_{x_i x_j} \neq 1$$

فاذا كانت العلاقة الخطية تامة بين المتغيرات التوضيحية ففي هذه الحالة لا يمكن ان نفصل اثر اي متغيرين عن بعضهما في نموذج الانحدار الخطي العام غالبا ما توجد بعض الارتباطات بين المتغيرات بشكل مباشر او غير مباشر كما(5:pp149).

4-1 انواع التعدد الخطي Multicollinearity Types

1-4-1 التعدد الخطي التام Perfect Multicollineaity

في هذا النوع من التعدد الخطي تكون مصفوفة المعلومات ($X'X$) برتبة غير كاملة(1:pp88). و يكون محدد مصفوفة المعلومات او مصفوفة فيشر (Fisher Matrix) مساوياً للصفر0 =$|X'X|$ وبذلك يصبح من المستحيل ايجاد تقدير لمعلمات أنموذج الانحدار نتيجة لعدم امكانية حساب معكوس مصفوفة المعلومات ($X'X$) وابرز طرق المعالجة في مثل هذه الحالة استخدام اسلوب المركبات الرئيسية (Principle Components)(3:pp182).

2-4-1 التعدد الخطي شبه التام Semi Perfect Multicollinearity

في هذا النوع تكون بعض المتغيرات التوضيحية دالة في التركيبة نفسها لمتغيرات اخرى مع قيم عشوائية كاملة(1:pp88)، عندها يكون محدد مصفوفة المعلومات صغير جداً يقترب من الصفر $|X'X| \approx 0$ وعندها تكون المعالم المقدرة ذات تباين كبير جداً وبذلك يمكن ان نستنتج بان بعض المتغيرات التوضيحية غير مهمة وابرز طرق المعالجة في مثل هذه هو اسلوب انحدار الحرف (Ridge Regression) (3:pp184).

وهنالك مجموعة من العواقب تسببها مشكلة التعدد الخطي منها عدم توفر خاصية $BLUE$ في مقدرات المربعات الصغرى الاعتيادية وهذه تؤدي الى تقديرات غير دقيقة للمعالم فضلاً عن فترات الثقة سوف تكون واسعة جداً مما يؤدي الى قبول فرضية العدم بشكل سهل جداً وربما القرار الصائب يكون رفض ناهيك ان اختبار t للمتغيرات المستقلة سوف يكون غير مؤكد من ناحية قبول معنوية تأثير هذا المتغير في النموذج او عدم تأثيره وبالتالي يكون غير معنوي احصائياً وانعكاس ذلك على كل من معامل التحديد R^2 وايضاً جدول تحليل التباين.

5-1 الكشف عن التعدد الخطي Detecting of Maulticollinearity

"لا يوجد اختبار لفرضية احصائية للدلالة على وجود تعدد العلاقة الخطية لأنها لا معلمة لها ، وتتبع طرائق احصائية مختلفة على احتمالية وجود مثل هذه العلاقة ، ان فشل طريقة ما ليست دليلا على عدم توافرها"(1:pp88).

هنالك عدة طرق للكشف عن التعدد الخطي نذكر منها:

1- حساب مصفوفة الارتباط (correlation matrix) للمتغيرات التوضيحية ، فاذا كانت القيمة عالية للارتباط بين اي متغيرين قد يشير الى ان المتغيرات تعاني من مشكلة التعدد الخطي .هذه الطريقة سهلة ، لكنها لا يمكن ان تنتج تقدير واضح لمعدل (درجة) التعدد الخطي.(10:pp586)

.كما وضع (Klein,) قاعدة لاختبار وجود مشكلة التعدد الخطي انه اذا كانت قيمة معامل الارتباط البسيط لأي متغيرين اكبر من قيمة معامل التحديد دل ذلك على وجود هذه المشكلة (4:pp224).

2- الدليل الشرطي ((Condition Index(CI) :

يستعمل الدليل الشرطي (CI) كأختبار لتشخيص مشكلة التعدد الخطي وهو بدوره يعتمد على الجذور المميزة (eigen values) والتي توضح كمية الاختلافات الكلية (Total variation) بين المتغيرات ، وعندما تكون الجذور المميزة مساوية للصفر فان ذلك يدل على وجود التعدد الخطي التام (Perfect Multiconllinearity) ، اما اذا كانت الجذور المميزة تقترب من الصفر فهذا يدل على وجود تعدد خطي عالي ، واذا كانت تلك الجذور مساوية الى الواحد فأنها تعبر عن عدم وجود المشكلة(2:pp173-174).

و الصيغة الرياضية للدليل الشرطي (CI) تكتب بالشكل التالي:

$$CI = \sqrt{\frac{\text{Maximum eigenvalue}}{\text{Minimum eigenvalue}}} = \sqrt{cn} \tag{4}$$

حيث ان cn يمثل العدد الشرطي (Condition number) والذي يعد ايضاً مؤشرا للكشف عن مشكلة التعدد الخطي وصيغته الرياضية تكون بالشكل التالي:

$$cn = \frac{\text{Maximum eigenvalue}}{\text{Minimum eigenvalue}} \tag{5}$$

اذا كانت قيمة cn بين 100 و 1000 دل ذلك على وجود مشكلة التعدد الخطي معتدلة الى قوية ، اما اذا كانت قيمة cn تتجاوز 1000 فهنالك مشكلة تعدد خطي شديدة. اما بالنسبة للدليل الشرطي CI فاذا كانت قيمته ما بين 10 و 30 دل ذلك على وجود مشكلة التعدد الخطي معتدلة الى قوية ، اما اذا كانت قيمته تتجاوز 30 فهنالك مشكلة تعدد خطي شديدة(13:pp298).

6-1 طريقة انحدار الحرف الاعتيادية (المتحيزة) Biased Ordinary Ridge Regression Method(RR)

اقترح كل من Hoerl and Kennard عام **1970** طريقة انحدار الحرف الاعتيادية (Ordinary Ridge Regression) لمعالجة مشكلة التعدد الخطي شبه التام ، بإضافة كمية صغيرة موجبة تدعى معلمة الحرف او معلمة التحيز(k) الى العناصر القطرية لمصفوفة المعلومات $X'X$ للحصول على تقديرات اكثر دقة ، والصيغة الرياضية لهذه الطريقة هي:

$$\hat{\beta}_{RR} = (X'X + kI_n)^{-1}\, X'Y \qquad , k \geq 0 \tag{6}$$

حيث ان:

I_n: مصفوفة الوحدة (Identity Matrix)

عندما تكون قيمة معلمة التحيز تساوي صفر $K = 0$ فأننا نحصل على تقديرات المربعات الصغرى الاعتيادية (OLS) **(9:pp57).**

و تتميز هذه الطريقة بمجموعة من الخصائص**(8:pp3)** .

1- مقدرات انحدار الحرف الاعتيادية تكون متحيزة ويمكن اثبات ذلك كالتالي:

$$\hat{\beta}_{RR} = (X'X + KI_n)^{-1}\, X'Y$$
$$= [(X'X + KI_n)^{-1}X'\left(X\underline{\beta} + \underline{\varepsilon}\right)]$$
$$= [(X'X + KI_n)^{-1}X'X\underline{\beta} + (X'X + KI_n)^{-1}X'\underline{\varepsilon}]$$

وبأخذ التوقع للطرفين نحصل على

$$E\left(\hat{\beta}_{RR}\right) = E[(X'X + KI_n)^{-1}X'X\underline{\beta} + (X'X + KI_n)^{-1}X'\underline{\varepsilon}]$$

$$\because E\left(\underline{\varepsilon}\right) = 0$$

$$\therefore E\left(\hat{\beta}_{RR}\right) = (X'X + KI_n)^{-1}X'X\underline{\beta} \tag{7}$$

نفرض ان $Z = (X'X + KI_n)^{-1}X'X$ والتي تمثل مقدار التحيز

$$E\left(\hat{\beta}_{RR}\right) = Z\underline{\beta}$$

2- تباين مقدرات انحدار الحرف الاعتيادية المتحيزة يكون:

$$Var\left(\hat{\beta}_{RR}\right) = Var[(X'X + KI_n)^{-1}\, X'Y]$$
$$= (X'X + kI_n)^{-1}X'Var(Y)X'(X'X + kI_n)^{-1} \tag{8}$$

3- متوسط مربعات الخطأ يحسب وفق الصيغة الاتية**(7:pp24):**

$$Mse\left(\hat{\beta}_{RR}\right) = tr\left[Var\left(\hat{\beta}_{RR}\right)\right] + \left[Bias\left(\hat{\beta}_{RR}\right)\right]^2$$
$$= \sigma^2\, tr[(X'X + KI_n)^{-1}X'X + (X'X + KI_n)^{-1}] + \underline{\beta}'(Z - I)(Z - I)\underline{\beta}$$
$$= \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2\beta'(X'X + KI_n)^{-2}\beta \tag{9}$$

ويمكن كتابته بصيغة اخرى كالتالي**(14:pp711)**

$$Mse(\hat{\beta}_{RR}) = \sigma^2 S_k^{-1} S S_k^{-1} + k^2 S_k^{-1} \underline{\beta}\, \underline{\beta}' S_k^{-1} \qquad (10)$$

حيث ان:

$S_k = X'X + kI_p$, $S = X'X$

ولغرض الحصول على قيمة واحدة لمتوسط مربعات الخطأ ناخذ اثر المصفوفة (trace matrix) لمصفوفة متوسط مربعات الخطأ.

7-1 طريقة انحدار الحرف الاعتيادية غير المتحيزة Unbiased Ordinary Ridge Regression Method (URR)

اقترح Swindle في عام 1976 تقنية لدمج المعلومات السابقة (prior information) مع انحدار الحرف وبذلك ستكون الصيغة المطورة لطريقة انحدار الحرف الاعتيادية كالتالي:

$$\hat{\beta}(KI, J) = (X'X + KI_p)^{-1} + (X'Y + K\underline{J}) \qquad , k \geq 0 \qquad (11)$$

حيث نلاحظ ان من الصيغة اعلاه انه ادخل متجه $\underline{J}$ والذي يمثل المعلومات السابقة (prior information) مع معلمة التحيز k ، فاذا كانت قيمة معلمة التحيز تساوي صفر $K = 0$ فأننا نحصل على تقديرات المربعات الصغرى الاعتيادية (OLS) (15:pp1072)

متجه المعلومات السابقة J يتوزع توزيعاً طبيعياً بمتوسط $\hat{\beta}$ وتباين $\frac{\sigma^2}{k} I_p$ اي ان($J \sim N(\underline{\hat{\beta}}, \frac{\sigma^2}{k} I_p)$ و $k > 0$ ، اما بالنسبة الى معلمة التحيز فتعطى من خلال الصيغة الاتية :

$$\hat{k}_{cjh} = \begin{cases} \dfrac{p\hat{\sigma}^2}{(\hat{\beta}_{LS} - J)'(\hat{\beta}_{LS} - J) - \hat{\sigma}^2 tr(X'X)^{-1}} & \text{if } (\hat{\beta}_{LS} - J)'(\hat{\beta}_{LS} - J) > \hat{\sigma}^2 tr(X'X)^{-1} \\ \dfrac{p\hat{\sigma}^2}{(\hat{\beta}_{LS} - J)'(\hat{\beta}_{LS} - J)} & , otherwise \end{cases}$$

حيث ان $\hat{\sigma}^2 = \frac{(Y - X\hat{\beta}_{LS})'(Y - X\hat{\beta}_{LS})}{(n-p)}$ وهي تقدير غير متحيز لتباين المجتمع σ^2. مع الاشارة الى ان $\hat{k}_{CJH}$ تعميم لـ $\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}_{LS}'\hat{\beta}_{LS}}$. (11:pp100)

و تتميز هذه الطريقة بمجموعة من الخصائص (14:pp710-711).

1- تكون هذه الطريقة غير متحيزة ويمكن اثبات ذلك كما يلي:

$\hat{\beta}(KI, J)$ يمكن كتابتها بالشكل التالي:

$$\underline{\hat{\beta}}(KI, J) = S_k^{-1} S \underline{\hat{\beta}}_{LS} + k S_k^{-1} \underline{J}$$

حيث ان $S_k = X'X + kI_p$, $S = X'X$ ويكون $\hat{\beta}(KI, J)$ تقدير غير متحيز في حالة استقلالية المتجه J عن متجه مقدرات $\hat{\beta}_{LS}$

في حالة توفر معلومات اولية عشوائية تعطى كالتالي:

$$h = HR + u$$

حيث ان h متجه عشوائي ذات رتبة$(l \times 1)$ ، H مصفوفة التصميم ذات رتبة $(l \times p)$ ، u متجه الخطأ العشوائي يتوزع طبيعياً بمتوسط (δ) ومصفوفة تباين مشترك $(\sigma^2 R)$ تكون σ^2 غير معلومة و R معلومة.

التقدير في النموذج $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ يخضع الى المعادلة اعلاه نحصل على

$$\breve{\beta} = (X'X + H'R^{-1}H)^{-1}(X'Y + H'R^{-1}h) \quad (12)$$

وبذلك فان التوقع والتباين المشترك لـ $(\breve{\beta})$ يكون:

$$E(\breve{\beta}) = \beta + (X'X + H'R^{-1}H)^{-1}H'R^{-1}\delta$$

$$Var(\breve{\beta}) = \sigma^2(X'X + H'R^{-1}H)^{-1}$$

عندما تكون $h = J$, $H = I$, $R = \frac{1}{k}I$ و $\delta = 0$ فان $E(\breve{\beta}) = \beta$ تساوي مقدر انحدار الحرف غير المتحيز.

2- متوسط مربعات الخطأ لمقدرات انحدار الحرف الاعتيادية غير المتحيزة يحسب من خلال الصيغة:

$$Mse\hat{\beta}(KI, J) = \sigma^2 S_k^{-1} S S_k^{-1} + k\sigma^2 S_k^{-2} \quad (13)$$

وللحصول على قيمة واحدة لمتوسط مربعات الخطأ نأخذ اثر المصفوفة (trace matrix) للمصفوفة في المعادلة (13).

8-1 انحدار الحرف البيزية Bayesian ridge regression (BRR)

في سياق نظرية بيز تعامل معلمة التحيز k كمعلمة غير معروفة و التي تمتلك توزيع اولي $p(k)$ وتكون مستقلة عن σ و β . لدينا(17:pp6-7):-

$$\underline{\beta}|\sigma, k \sim N_p\left(0, \frac{1}{k} I_p\right) \ , p(\sigma|k) \alpha\ \sigma^{-1} \quad (14)$$

والتوزيع الاولي لـ k يتناسب مع $p(k)$ ، ليس من الضروري ان يكون هذا التوزيع الاولي تاماً (proper)

لذلك التوزيع المشترك اللاحق يكون على النحو التالي:-

$$p\left(\underline{\beta}, \sigma, k \middle| y, X\right) \alpha\ \sigma^{-(n+1)} k^{p/2} p(k) \exp\left[-\frac{\left(y - X\underline{\beta}\right)'\left(y - X\underline{\beta}\right) + k\sigma^2 \underline{\beta}'\underline{\beta}}{2\sigma^2}\right] \quad (15)$$

وبإكمال المربع $Q = \left(y - X\underline{\beta}\right)'\left(y - X\underline{\beta}\right) + k\sigma^2 \underline{\beta}'\underline{\beta}$ نحصل على

$$Q = \left(\underline{\beta} - b_k\right)'\left(X'X + k\sigma^2 I_p\right)\left(\underline{\beta} - b_k\right) + y'M_k y$$

حيث ان : $M_k = I_p - XV_kX'$, $\quad V_k = \left(X'X + k\sigma^2 I_p\right)^{-1}$

لذلك ، التوزيع اللاحق المشترك يكون:

$$p\left(\underline{\beta},\sigma,k\middle|y,X\right)\alpha\sigma^{-(n+1)}k^{p/2}p(k)\exp\left[-\frac{\left(\underline{\beta}-b_k\right)'\left(X'X+k\sigma^2I_p\right)\left(\underline{\beta}-b_k\right)+y'M_ky}{2\sigma^2}\right]$$

حيث ان :

$$b_k=(X'X+k\sigma^2I_p)^{-1}X'y$$

من المعادلة السابقة يمكننا استخراج التوزيعات اللاحقة الشرطية كالتالي:

$$\underline{\beta}|\sigma,k,y,X\sim N_p\left(b_k,\sigma^2\left(X'X+k\sigma^2I_p\right)^{-1}\right),$$

$$\frac{\left(y-X\underline{\beta}\right)'\left(y-X\underline{\beta}\right)}{\sigma^2}\left|\underline{\beta},k,y,X\sim\chi^2_{(n)}\right)$$

حيث $\chi^2_{(n)}$ يدل على توزيع $(chi-square)$ مربع كاي بدرجة حرية (n) ، للحصول على التوزيع الشرطي اللاحق لمعلمة التحيز يمكننا اشتقاق المعادلة لذلك :

$$p\left(k\middle|\underline{\beta},\sigma,y,X\right)\alpha\, k^{p/2}p(k)\exp\left(-\frac{k\underline{\beta}'\underline{\beta}}{2}\right) \qquad (16)$$

التوزيع الاولي المرافق الشرطي لمعلمة التحيز k هو بوضوح توزيع $Gamma(a/2,b/2)$ التي تكون دالة الكثافة الاحتمالية له تاخذ الشكل الاتي:

$$p(k)=\frac{(b/2)^{a/2}}{\Gamma(a/b)}k^{(a/2)-1}\exp\left(-\frac{b}{2}k\right) \qquad (17)$$

حيث $a\geq 0,b\geq 0$ تمثل المعلمات الفوقية $(hyperparameters)$. لذلك نحصل على :

$$p\left(k\middle|\underline{\beta},\sigma,y,X\right)\alpha\, k^{((p+a)/2)-1}\exp\left(-\frac{b+\underline{\beta}'\underline{\beta}}{2}k\right)$$

لذلك ، التوزيع الشرطي اللاحق لمعلمة التحيز يكون:

$$\left(k\middle|\underline{\beta},\sigma,y,X\right)\sim Gamma\left(\frac{n+a}{2},\frac{b+\underline{\beta}'\underline{\beta}}{2}\right) \qquad (18)$$

1-2 **الجانب التجريبي**:

يمكن وصف تجربة المحاكاة الخاصة بالبحث من خلال توليد بيانات لمتغيرات عشوائية من توزيعات معينة مثل التوزيع المنتظم للمتغيرات التوضيحية والتوزيع الطبيعي لمتغير الخطأ العشوائي وبحجوم عينات $(15,25,50,75,100,150)$ مشاهدة بالاعتماد على نموذج الانحدار الخطي العام .

$$y_i=\beta_0+\beta_1x_{1i}+\beta_2x_{2i}+\beta_3x_{3i}+\beta_4x_{4i}+\mathcal{E}_i$$

وان القيم الافتراضية للمعلمات[1]

$(\beta_0=0.2,\beta_1=1.2,\beta_2=0.8,\beta_3=-4.2,\beta_4=2.34)$ وقيم المعلمات لتوزيع كاما فقد اخذ اربع حالات كالاتي:

[1] تم اختيار قيم محددة للمعلمات في الجانب التجريبي بالاعتماد على الخبرة بالاطلاع على نماذج في دراسات سابقة.

جدول (1) يبين القيم الافتراضية[2] لمعلمة التوزيع الموظف في التقدير

الحالات / المعالم	1	2	3	4
a	0.2	0.04	0.02	0.004
b	2	0.4	0.2	0.04

اما بالنسبة الى المتجه J فقد ولد من خلال التوزيع الطبيعي $\underline{J}\sim N(\underline{\hat{\beta}},\frac{\sigma^2}{K}I_P)$ حيث ان $\underline{\hat{\beta}}$ متجه المعالم الافتراضية اما قيمة K فقد اخذت قيم معالم التحيز $K_{MI1},K_{HKB},K_{LW},K_{MI2}$.

جدول رقم (2) يوضح قيم متوسط مربعات الخطأ عندما تكون قيمة التباين $\sigma^2 = 0.3$

حجوم العينات	متوسط مربعات الخطأ	طرق التقدير	$K_{LW} = 0.4084$	$K_{HKB} = 0.0080$	$K_{MI1} = 0.4344$	$K_{MI2} = 1.2968$
$n = 15$	MSE	OLS	1.1045	1.1045	1.1045	1.1319
		RR	1.0303	1.0245	5.6392	8.6589
		URR	1.1003	1.1003	0.7876	0.6752
		BRR	25.2332	25.2332	25.2332	25.2335
$n = 25$	MSE	OLS	0.5500	0.5500	0.55	0.5269
		RR	0.5412	0.5345	1.32553	1.4845
		URR	0.5533	0.5533	0.5767	0.5936
		BRR	25.2149	25.2149	25.2149	25.2142
$n = 50$	MSE	OLS	0.2284	0.2284	0.2284	0.2339
		RR	0.2276	0.2253	0.2560	0.2306
		URR	0.2284	0.2282	0.3163	0.3080
		BRR	25.1679	25.1679	25.1679	25.1675
$n = 75$	MSE	OLS	0.1513	0.1513	0.1513	0.1520
		RR	0.1511	0.1502	0.1527	0.1502
		URR	0.1516	0.1516	0.2513	0.1833
		BRR	25.1207	25.1207	25.1207	25.1203
$n = 100$	MSE	OLS	0.1155	0.1155	0.1155	0.1126
		RR	0.1154	0.1146	0.1145	0.1120
		URR	0.1156	0.1156	0.1480	0.1136
		BRR	25.0736	25.0736	25.0736	25.0742
$n = 150$	MSE	OLS	0.0765	0.0765	0.0765	0.0710
		RR	0.0765	0.0762	0.0760	0.0708
		URR	0.0765	0.0765	0.1184	0.0742
		BRR	24.9803	24.9803	24.9803	24.9803

من خلال الجدول (2) يمكن ان نلاحظ ان :

[2] تم اخذ هذه القيم الافتراضية بالاعتماد على المصدر 17

عند حجم عينة $n = 15$ نلاحظ ان طريقة RR اعطت اقل Mse باستخدام معلمتي التحيز k_{LW}) ، k_{HKB}) ، وعند نفس حجم العينة فان طريقة URR اعطت اقل Mse باستخدام معلمتي التحيز الآخرتين k_{MI1}) ، k_{MI2}).

اما بالنسبة الى احجام العينات $(25,50,75,100)$ فان طريقة RR كانت هي الأفضل من بين بقية الطرق لمعظم معالم التحيز حيث اعطت اقل Mse .

اما بالنسبة لحجم العينة 150 في التجارب التي اخذت فان طريقة RR هي الافضل بشكل مطلق لتبين افضلية الطريقة باستخدام معالم التحيز الاربعة.

عندما كانت $\sigma^2 = 0.3$ فان التحليل اعلاه يشير الى ان طريقة RR هي الافضل من بين باقي الطرق حيث امتلكت اقل Mse .

جدول رقم (3) يوضح قيم متوسط مربعات الخطأ عندما تكون قيمة التباين $\sigma^2 = 0.8$

حجوم العينات	متوسط مربعات الخطأ	طرق التقدير	$K_{LW} = 4.6195$	$K_{HKB} = 0.1235$	$K_{MI1} = 1.0983$	$K_{MI2} = 4.6195$
$n = 15$	MSE	OLS	7.0842	7.0842	7.0842	7.0357
		RR	5.7762	16.0901	16.0901	16.2375
		URR	8.2529	2.2439	2.2439	2.2447
		BRR	25.2386	25.2386	25.2386	25.2384
$n = 25$	MSE	OLS	3.5527	3.5527	3.5527	3.5499
		RR	3.3467	8.4847	8.4847	5.9472
		URR	4.4682	1.7341	1.7341	2.0565
		BRR	25.2240	25.2240	25.2240	25.2240
$n = 50$	MSE	OLS	1.5084	1.5850	1.5850	1.5854
		RR	1.4893	2.1939	2.1939	1.4093
		URR	1.9839	1.2296	1.2296	1.4677
		BRR	25.1865	25.1858	25.1858	25.1858
$n = 75$	MSE	OLS	0.9793	0.9793	0.9793	1.0087
		RR	0.9740	1.0392	1.0392	0.9480
		URR	1.5732	0.8775	0.8775	0.9833
		BRR	25.1491	25.1491	25.1491	25.1491

$n = 100$	MSE	OLS	0.7081	0.7081	0.7081	0.7081
		RR	0.7058	0.6951	0.6951	0.6882
		URR	0.8060	0.6824	0.6824	0.7167
		BRR	25.1120	25.1120	25.1120	25.1120
$n = 150$	MSE	OLS	0.4943	0.4943	0.4943	0.4943
		RR	0.4936	0.4775	0.4775	0.4901
		URR	0.5868	0.4892	0.4892	0.5013
		BRR	25.0410	25.0410	25.0410	25.0410

من خلال الجدول (3) يمكن ان نلاحظ ان:

عند حجم عينة $n = 15$ فان طريقة URR اعطت اقل Mse عند استخدام معالم التحيز k_{MI1}) ، k_{MI2} ، k_{HKB}) لكن اختلفت عنها عند استخدام معلمة التحيز k_{LW}) (حيث كانت طريقة RR هي الافضل لأنها امتلكت اقل Mse .اما بالنسبة لأحجام العينات $(25,50,75,100)$ فان طريقة URR امتلكت اقل Mse باختلاف معالم التحيز لمعظم التجارب وعند حجم عينة 150 فان طريقة RR هي الافضل من خلال معيار Mse باختلاف معالم التحيز.

جدول رقم (4) يوضح قيم متوسط مربعات الخطأ عندما تكون قيمة التباين $\sigma^2 = 1.5$

حجوم العينات	متوسط مربعات الخطأ	طرق التقدير	$K_{LW} = 0.6019$	$K_{HKB} = 0.2376$	$K_{MI1} = 3.2431$	$K_{MI2} = 13.2106$
$n = 15$	MSE	OLS	24.4293	24.4293	24.4293	24.4293
		RR	17.9009	20.9270	20.9270	20.9270
		URR	22.7612	6.1836	6.1836	6.1836
		BRR	25.2412	25.2412	25.2412	25.2412
$n = 25$	MSE	OLS	12.3152	12.3152	12.3152	12.3152
		RR	11.1639	15.8724	15.8724	15.8724
		URR	13.0462	3.7022	3.7022	3.7022
		BRR	25.2289	25.2289	25.2289	25.2289
$n = 50$	MSE	OLS	5.5083	5.5083	5.5083	5.5083
		RR	5.3958	7.3042	7.3042	7.3042
		URR	6.1226	2.5121	2.5121	2.5121
		BRR	25.1963	25.1963	25.1963	25.1963

n = 75	MSE	OLS	3.5027	3.7189	3.5027	3.5027
		RR	3.4702	3.7189	3.7189	3.7189
		URR	3.7582	2.0528	2.0528	2.0528
		BRR	25.1658	25.1658	25.1658	25.1658
n = 100	MSE	OLS	2.4656	2.4656	2.4656	2.4656
		RR	2.4531	2.3876	2.3876	2.3876
		URR	2.5066	1.8057	1.8057	1.8057
		BRR	25.1361	25.1361	25.1361	25.1361
n = 150	MSE	OLS	1.7127	1.7127	1.7127	1.7127
		RR	1.7088	1.5501	1.5501	1.5501
		URR	1.7446	1.4834	1.4834	1.4834
		BRR	25.0825	25.0825	25.0825	25.0825

من خلال الجدول (5) يمكن ان نلاحظ ان:

نلاحظ من الجدول اعلاه ان كل الطرق التي استخدمت في تجارب المحاكاة الستة فان طريقة RR كانت هي الافضل عند استخدام معلمة التحيز k_{LW}() وبشكل واضح.

اما عند استخدام الانواع الثلاثة الاولى لمعالم التحيز فان طريقة URR كانت الافضل من خلال معيار Mse.

جدول رقم (6) يوضح قيم متوسط مربعات الخطأ عندما تكون قيمة التباين $\sigma^2 = 6$

حجوم العينات	متوسط مربعات الخطأ	طرق التقدير	$K_{LW} = 13.1995$	$K_{HKB} = 0.9250$	$K_{MI1} = 34.0594$	$K_{MI2} = 71.6780$
$n = 15$	MSE	OLS	390.7326	399.4694	399.4694	390.7326
		RR	266.3429	104.8778	24.4472	24.4157
		URR	221.8387	115.5429	93.6642	91.8593
		BRR	25.2430	25.2429	25.2429	25.2430
$n = 25$	MSE	OLS	196.1868	196.1727	196.1727	196.1868
		RR	172.5317	57.7452	23.3784	23.3639
		URR	152.6030	57.8612	46.9447	46.8705
		BRR	25.2319	25.2318	25.2318	25.2319

$n = 50$	MSE	OLS	87.6791	87.6940	87.6940	87.6791
		RR	85.2252	31.2908	20.6629	20.6643
		URR	81.9933	26.1607	21.0002	20.9964
		BRR	25.2030	25.2030	25.2030	25.2030
$n = 75$	MSE	OLS	54.0268	55.7316	55.7316	54.0268
		RR	53.2829	22.4524	18.0223	18.2297
		URR	53.2829	16.0259	13.1294	12.7224
		BRR	25.1772	25.1770	25.1770	25.1772
$n = 100$	MSE	OLS	39.3323	39.1507	39.1507	39.3323
		RR	39.0760	18.4308	15.8587	15.8639
		URR	39.2922	11.9387	9.8278	9.9130
		BRR	25.1531	25.1532	25.1532	25.1531
$n = 150$	MSE	OLS	27.1924	27.1924	27.1924	27.1924
		RR	27.1113	14.7067	12.6810	12.6810
		URR	27.2295	8.5731	7.5129	7.5129
		BRR	25.1149	25.1149	25.1149	25.1149

من خلال الجدول (6) يمكن ان نلاحظ ان:

ان طريقة BRR اعطت اقل Mse عند استخدام معلمة التحيز (k_{LW}) ولجميع احجام العينات الستة.

اما عند استخدام معالم التحيز الاخرى (k_{MI1} ، k_{MI2} ، k_{HKB}) و لأحجام العينات $(25,50)$ فان طريقتي BRR و URR كانتا هما الافضل.

وعند تكبير حجم العينة الى $(75,100,150)$ فان طريقة URR كانت هي الافضل من الطريقتين الاخريتين عندما كان حجم العينة صغير.

من خلال الجداول الاربعة اعلاه لتجارب المحاكاة باختلاف احجام العينات واختلاف التباينات تبين ان طريقة RR هي الافضل عندما يكون حجم العينة كبير وتباين صغير وتبين انها افضل طريقة من خلال معيار متوسط مربعات الخطأ التي اعطته باستخدام معالم التحيز الاربعة.

ويمكن ملاحظة انه عند تقليل التباين فان Mse ايضاً ينخفض عند استخدام طريقة اندار الحرف الاعتيادية المتحيزة وبشكل طردي ولجميع احجام العينات.

3-1 **وصف المرض**[16]

السُكَّري أو داء السكري (Diabetes Mellitus) هو ارتفاع في تركيز سكر الدم الناجم عن نقص هرمون الأنسولين أو انخفاض حساسية الأنسجة للأنسولين. وهو مرض خطير ويمكن ان يؤدي الى الوفاة، ويمكن للمريض ان يتخذ تدابير علاجية لغرض السيطرة على هذا المرض ويكون على نوعين الأول : يتميز النمط الأول (I) من السكري بخسارة الخلايا بيتا المنتجة للأنسولين في خلايا لانغرهانس بالبنكرياس مما يؤدي إلى نقص الأنسولين اما النوع الثاني يتميز باختلافه عن النوع الأول من حيث وجود مقاومة مضادة لمفعول الأنسولين بالإضافة إلى قلة إفراز الأنسولين ولا تستجيب مستقبلات الأنسولين الموجودة في الأغلفة الخلوية لمختلف أنسجة الجسم بصورة صحيحة للأنسولين, وان من اهم اعراض الاصابة بهذا المرض التعب الشديد والاحساس الشديد بالعطش وزيادة عدد مرات التبول والفقدان المستمر للوزن بالإضافة الى انه عند جرح المريض لا يشفى الجرح بسهولة.

اما بالنسبة للجانب التطبيقي لبيانات حقيقية لمجموعة من المرضى المصابين بداء السكري ، إذ تم الحصول على البيانات من البطاقات الخاصة بالمرضى في مركز السكري والغدد الصم التخصصي في محافظة ذي قار لعام 2017، تم الحصول على بيانات مجموعة من المرضى حيث بلغ عددهم (112) مريض وكان نموذج الانحدار الخطي المتعدد يتضمن :

المتغير المعتمد Y : السكر Diabetes، والمتغيرات المستقلة هي :

عمر المريض X_1 :(age)

الكولسترول X_2 :(S. cholesterol)

الدهون الثلاثية X_3 :(S.Triglyceride)

يوريا الدم X_4: (Blood Urea)

الكرياتين X_5:(S.creatinine)

معدل ضغط الدم X_6: (Avarage Blood pressure)

مؤشر كتلة الجسم X_7 : (Body Mass Index) ويحسب من خلال قسمة وزن المريض على مربع طول المريض بالمتر.

3-2 **اختبار وجود مشكلة التعدد الخطي**:

يمكن ملاحظة وجود مشكلة التعدد الخطي من استخراج مصفوفة الارتباطات للمتغيرات التوضيحية ، حيث يلاحظ قوه هذه الارتباطات مما يعني وجود هذه المشكلة وكالاتي:

جدول(7) قيم مصفوفة الارتباط بين المتغيرات المستقلة

	X1	X2	X3	X4	X5	X6	X7
X1	1.0000						
X2	0.1185	1.0000					
X3	0.1712	0.8322	1.0000				
X4	0.9308	0.1120	0.1358	1.0000			
X5	0.3941	0.1588	0.1732	0.4449	1.0000		
X6	0.1326	-0.0294	-0.1009	0.1136	0.0811	1.0000	
X7	0.7840	0.0600	0.0900	0.7990	0.3034	0.0548	1.0000

يلاحظ من مصفوفة الارتباطات للمتغيرات التوضيحية بان هنالك علاقة قوية بين المتغيرات (X_1,X_4) اذ بلغ معامل الارتباط بينهما (0.9308) ، وهي علاقة طردية بين المتغيرين ، وفي المرتبة الثانية كانت العلاقة قوية بين المتغيرين (X_2,X_3) اذ بلغ معامل الارتباط بينهما (0.8322) وهي علاقة طردية قوية بينهما ، وفي المرتبة الثالثة العلاقة بين المتغيرين (X_4,X_7) حيث بلغ معامل الارتباط بينهما (0.7990) وهي علاقة طردية قوية الى حد ما بينهما ، وكان معامل الارتباط بين المتغيرين (X_1,X_7) قد بلغ (0.7840) وهذا يدل على ان العلاقة طردية الى حد ما ، ذلك يدل على وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية .

اما بالنسبة لقاعدة كلاين فان قيمة معامل الارتباط البسيط بين المتغيرات (X_1,X_4) ، (X_2,X_3) ، (X_4,X_7)، (X_1,X_7)، ولجميع هذه المتغيرات اكبر من معامل التحديد R^2 والتي بلغت قيمته (0.6577) وهذا يدل على وجود مشكلة التعدد الخطي.

الدليل الشرطي :

بلغت قيمة الدليل الشرطي $CI = 2856.4593$ وهي اكبر من 30 ، وكذلك بالنسبة للعدد الشرطي بلغت قيمته 8159359.6851 وهي اكبر من 1000 ، و المؤشران يدلان على وجود مشكلة التعدد الخطي.

ان طريقة انحدار الحرف الاعتيادية المتحيز RR من خلال تجارب المحاكاة والتي اعطتنا اقل Mse ففي الجانب التطبيقي سنطبق هذه الطريقة .

جدول(8) قيم معلمات الانحدار بطريقة انحدار الحرف الاعتيادية المتحيزة

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$
$\hat{\beta}_{(KHB)}$	0.1570	0.9263	0.0969	0.3543	−15.9389	0.3029	−1.1291
$\hat{\beta}_{(LW)}$	0.0347	0.9298	0.1129	0.7068	−55.6828	0.4550	−1.0912
$\hat{\beta}_{(MI1)}$	0.1494	0.9265	0.0979	0.3769	−18.4339	0.3125	−1.1271
$\hat{\beta}_{(MI2)}$	0.0349	0.9298	0.1129	0.7069	−55.6129	0.4547	−1.0912

4-1 **الاستنتاجات**:

1- من خلال التحليل للبيانات الحقيقية تبين ان متغير الكولسترول له تأثير فعال على مرض السكر ويكاد يكون سبب رئيسي للمرض.

2- ان متغيري يوريا الدم ومعدل ضغط الدم يؤثران على مرض السكر بصورة اقل من تأثير الكولسترول عليه.

3- عند انخفاض التباين فان تجارب المحاكاة اعطت افضل طريقة بزيادة احجام العينات عند استخدام معيار Mse للمفاضلة بين الطرق.

4- من خلال تجارب المحاكاة ظهرت طريقة انحدار الحرف الاعتيادية المتحيزة RR افضل طريقة من بين الطرق الاخرى.

2-4 **التوصيات**:

1- من خلال الجانب التطبيقي نوصي مراجعة الطبيب لتقليل الكولسترول ويوريا الدم.

2- استخدام طريقة انحدار الحرف الاعتيادية المتحيزة RR لكونها الافضل من خلال تجارب المحاكاة.

5- **المصادر** :

1-دبدوب ، مروان عبد العزيز و النعيمي ، اسوان محمد طيب 2006م ، " طرائق مقترحة في انحدار الحرف " المجلة العراقية للعلوم الاحصائية جامعة الموصل ،المجلد 10،العدد 6 ، ص [85-106] .

2-يحيى، مزاحم محمد و عبدالله ،محمود حمدون،2007م،"تشخيص التعدد الخطي واستخدام انحدار الحرف في اختيار متغيرات دالة الاستثمار الزراعي في العراق للفترة 1980-2000"،مجلة تكريت للعلوم الادارية والاقتصادية،المجلد3،العدد8، 171-187.

3-كاظم ، أموري هادي و مسلم ، باسم شليبة 2002م ، " القياس الاقتصادي المتقدم النظرية والتطبيق"، مكتبة دنيا الامل ، بغداد.

4-محمد ، رواء صالح 2010م ،"استخدام انحدار الحرف (Ridge) لدراسة اثر بعض العوامل على المؤشر العام لسوق الاوراق المالية"، مجلة القادسية للعلوم الادارية والاقتصادية ، المجلد 12، العدد 4،ص 220-240.

5-العزاوي ، دجلة ابراهيم مهدي و الشمري، نذير عباس ابراهيم، 2011م ، "الاقتصاد القياسي وتطبيقاته" مكتبة الجزيرة للطباعة و النشر، بغداد ص19.

6-الصفاوي، صفاء يونس و طه ،عمار حازم 2005، " بعض طرائق المقدرات التقليدية ومقدر بيز لمعلمات نموذج الانحدار الخطي العام(دراسة مقارنة مع تطبيق في مجال طبي)، تنمية الرافدين 80(27) ص (91-103).

7-Al-Hassan, Y.M., Performance of a New Ridge Regression Estimator.Journal of The Association of Arab Universities for Basic and Applied Sciences, 2010. 9 :pp.(23 – 26)

8-Al-sadoum, Muhannad Faiz,2005," Empirical Bayes and Bayes for Ridge regression ",Quarterly Specialized Refereed Journal ,vol.(7) ,no.(1).

9-Arthur , E. Hoerl , Robert W. Kennard (1970) ,''Ridge Regression Biased Estimation for Nonorthogonal Problems'' ; Technometrics Vol.12 , No.1 , pp.55-67.

10-El-Dereny,M. & Rashwan,N.I. , (2011) , “ Solving Multicollinearity Problem Using Ridge Regression Models “ Int.J.Contemp.Math.Scienes, Vol.6,No.12,pp(585-600).

11-Feras Sh. M. B. and Sharad D. G., (2009) “ Ridge Regression Estimator Combining Unbiased And Ordinary Ridge Regression Methods of Estimation” , Surveys in Mathematics and its Applications Volume 4,PP(99 – 109).

12-Gujarati, D. , (2011) ,”Econometrics By Examples” ,Palgrave Macmillan .

13-Montgomery, D.C. et al ,(2012) ," Introduction to Linear Regression Analysis", Published by John Wiley & Sons, Inc., Hoboken, New Jersey Published simultaneously in Canada:71-72.

14-Ozkale,m.revan and Kaciranlar,selahattin , Comparisons of the Unbiased Ridge Estimation to the Other Estimations, Communications in Statistics—Theory and Methods, 36: 707–723, 2007.

15-Swindel, B.F. (1976).'' Good ridge estimators based on prior information''. Communications in Statistics-Theory and Methods, 5 1065-1075.

16-http://ar.wikiepedia.org 17-https://ssrn.com/abstract=2387582

Compared Some Estimators Ordinary Ridge Regression And Bayesian Ridge Regression With Practical Application

Abstract:

Maulticollinearity is a problem that always occurs when two or more predictor variables are correlated with each other. consist of the breach of one basic assumptions of the ordinary least squares method with biased estimates results, There are several methods which are proposed to handle this problem including the principle components **method To address a problem** Maulticollinearity perfect **and** Ridge Regression **method To address a problem** semi Maulticollinearity**, In this research a comparisons are employed between the biased** Ridge Regression **method and unbiased** Ridge Regression **method with Bayesian** Ridge Regression **using Gamma distribution method addition to Ordinary Least Square method, We will use the simulation to compare these methods using the mean squares error criteria. The method of biased** Ridge Regression **gave good results by using sizes different samples.**

Key words: **biased** Ridge Regression, **unbiased** Ridge Regression, **Bayesian** Ridge Regression.