A Simulation Study Of Ridge regression Method With Autocorrelated Errors

Hussein Yousif Abdallah*

Abstract

The objective of this paper is to discuss the case of the general linear model that suffered from both problems autocorrelation (AR(1) and AR(2)) and multicollinearity, we used simulation technique to create data containing these problems at the same time , and we used the two stages least squares procedure (2SLS) to deal with problem of autocorrelation , and ridge regression (RR) to deal with the problem of multicollinearity of the data that had originally treated in consideration of autocorrelation. Moreover we used the evaluation methods as a base for the process of evaluation and comparison.

Throughout the simulation experiment results domain we concluded that dealing with autocorrelation from data that suffered from multicollinearity, multicollinearity increases when the error term follows first or second order autoregressive scheme. Whereas, multicollinearity decreases if the model has a few explanatory variables .Among the types of ridge regression method , if we take the *MSE* as a criteria of comparison we find that ordinary ridge regression is the best when the sample size is too large, otherwise , generalized ridge regression is the best one.

* Assistance professor - Faculty of Sciences & Technology- Shendi University

ملخص

تتناول الدراسة النموذج الخطي العام الذي يعاني من مشكلة الارتباط الذاتي من الدرجة الأولى والدرجة الثانية بالاضافة لمشكلة التداخل الخطي المتعدد، وقد تم استخدام أسلوب المحاكاة في توليد بيانات تعاني من المشكلتين في وقت واحد ، وتم استخدام طريقة المربعات الصغرى ذات المرحلتين في معالجة الارتباط الذاتي بدرجتيه الاولى والثانية ، كما أُستخدم اسلوب Ridg Regression في معالجة التداخل الخطي المتعد من البيانات بعد ازالة مشكلة الارتباط الذاتي منها علاوة على ذلك أُستخدمت بعض المقاييس الاحصائية في عملية التقييم والمقارنة.

ومن خلال نتائج تجربة المحاكاة تم التوصل الى أن ازالة الارتباط الذاتي من الدرجة الاولى أو الثانية من البيانات التى تعاني من التداخل الخطي المتعدد يؤدي ذلك الى زيادة خطورة التداخل الخطي المتعدد وتقل خطورة المشكلة اذا كان عدد المتغيرات التوضيحية قليل.أما من بين انواع أساليب انحدار الحرف أنه اذا تم استخدام MSE كأساس في عملية المقارنة نجد أن أسلوب الحرف العمومي أفضل من أسلوب الحرف الاعتيادي اذا كان حجم العينة صغيراً بينما نجد أن الأخير هو الافضل اذا كان حجم العينة كبيرا .

Introduction

The ordinary least squares method is considered as one of the most important ways of estimating the parameters of the general linear model because of its ease and simplicity and rationality of the results that obtained when the specific assumptions are achieved regarding the general linear model about error term and explanatory variables which are supposed to be orthogonal.

Yet if these assumptions are not verified , the ordinary least squares method will give undesirable results , and there appears the problem of inaccurate estimation , one of which is associated with the autocorrelation of errors which occurs when the value of the error term in any particular period is correlated with its own preceding value or values [$E(U_t\ U_{t-s}) = 0 \quad s \neq 0$] multicollinearity is another significant problem , this occurs when the explanatory variables are correlated with each other.

Suppose there is a linear relation between dependent variable Y_i and explanatory variables $X_1, X_2, \ldots, X_p$ and error term U_i , we can write this relation as follows [Draper & Smith1981:23] :

$$Y_i = B_0 + B_1 X_{1i} + B_2 X_{2i} + \ldots + B_p X_{pi} + U_i \quad \ldots\ldots (1)$$

$$i = 1,2,\ldots\ldots,n$$

Where:

Y_i : is the $\bar{i}\,th$ observation of response variable .

X_{ji} : is the $\bar{i}\,th$ observation of explanatory variable j .

$B_0, B_1, B_2, \ldots, B_p$: are the parameters or regression coefficients .

U_i : is a random error term or disturbance term .

In matrix form the general linear model GLM (1) is :

$$\underline{Y} = X\underline{B} + \underline{U} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Where :

$\underline{Y}$: $(n\text{x}1)$ vector observations of the dependent variable .

X : $(n\text{x}(p+1))$ matrix of explanatory variables .

$\underline{B}$: $((p+1)\text{x}1)$ vector of regression coefficients.

$\underline{U}$: $(n\text{x}1)$ vector of errors .

The estimation of B using OLS is as follows [Younis & Others ,2002:156]

$$\underline{b} = (X'X)^{-1}X'Y \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Autocorrelation Problem

This problem occurs when the assumption of the classical linear model about the independence of the disturbances from observation to others ($E(U_jU_s)=0 \quad \forall \quad s \neq j$) is not verified ,therefore the errors in one time period are correlated with their own values in other periods[Ronald ,2002] .

The model with first–order autoregressive process AR(1) has the form [John Nester & others ,1985:448] :

$$U_t = rU_{t-1} + V_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Where

r is autocorrelation parameter(coefficient) ($|r.(1 >|$

V_t is a random disturbance , where

$$V_t \sim (0, \sigma_V^2)$$
$$E(V_t V_{t-s}) = \sigma_V^2 \quad \forall \quad s = 0$$
$$= 0 \quad \forall \quad s \neq 0$$

The model with second – order autoregressive process AR(2) has form

$$U_t = r_1 U_{t-1} + r_2 U_{t-2} + V_t \ldots\ldots\ldots\ldots\ldots\ldots \text{(5)}$$

To estimate the Autocorrelation Coefficient we used Durbin Two – Step Method ,This method proposed by Durbin (1960) [Durbin,1970].The steps in the estimation procedure are as follows :

For AR(1) :- Consider the transformation model :

$$Y_t = B_0(1- r) + rY_{t-1} + B_1X_{1t} - B_1rX_{1t-1} + \ldots\ldots + B_pX_{pt} - B_prX_{pt-1} + V_t$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\text{(6)}$$

Let

$$a_0 = B_0(1- r) \;,\quad a_1 = B_1 \;,\; a_2 = B_1r \;\ldots\ldots\ldots\ldots\;,\; a_p = B_pr$$

Therefore, we can rewrite equation (6) as

$$Y_t = a_0 + rY_{t-1} + a_1X_{1t} + a_2X_{1t-1} + \ldots\ldots + a_prX_{pt-1} + V_t \;\ldots\ldots\ldots\text{(7)}$$

Estimate the regression equation (7) by OLS and obtain estimated coefficient of the lagged variable Y_{t-1} ($\hat{r}$)

For AR(2): Consider the transformation model

$$(Y_t - r_1Y_{t-1} - r_2Y_{t-2}) = B_0(1- r_1 - r_2) + B_1(X_{1t} - r_1X_{1t-1} - r_2X_{1t-2}) + \ldots\ldots + B_p$$

$$(X_{pt} - r_1X_{pt-1} - r_2X_{pt-2}) + W_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(8)}$$

Where $W_t = r_1U_{t-1} + r_2U_{t-2}$

We rewritten equation (8) as follows

$$Y_t = B_0(1- r_1 - r_2) - r_1Y_{t-1} + r_2Y_{t-2} + B_1X_{1t} - B_1r_1X_{1t-1} - B_1r_2X_{1t-2} \ldots\ldots +$$

$$B_pX_{pt} - B_pr_1X_{pt-1} - B_pr_2X_{pt-2} + W_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

Estimate the regression equations (9) by OLS and obtain estimated coefficients of the lagged variables Y_{t-1} , Y_{t-2} (r_1,r_2).

In order to deal with this problem we used The Two Stages least squares procedure (2SLS) ,we can summarize this method as follows [Kadiyala,1968]:

Pre multiply equation (2) by T we obtain:

$$\underline{T}\,\underline{Y} = \underline{T}X\underline{B} + \underline{T}\,\underline{U} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

where :

$$T = \begin{pmatrix} -\rho & 1 & 0 & - & - & 0 & 0 & 0 \\ 0 & -\rho & 1 & 1 & 1 & 0 & 0 & 0 \\ - & - & - & - & - & 0 & 0 & 0 \\ - & & & & & - & - & - \\ - & & & & & - & - & - \\ 0 & 0 & 0 & & & 0 & -\rho & 1 \end{pmatrix} \qquad \text{If } U_t \text{ follows } AR$$

And

$$T = \begin{pmatrix} \sigma_u / \sigma_V & 0 & 0 & . & . & 0 & 0 & 0 \\ \frac{-\rho_1}{1-\rho_1}\sqrt{1-\rho_2^2} & \sqrt{1-\rho_2^2} & 0 & . & . & 0 & 0 & 0 \\ -\rho_2 & -\rho_1 & 1 & 0 & . & 0 & 0 & 0 \\ 0 & -\rho_2 & -\rho_1 & 1 & . & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & -\rho_2 & -\rho_1 & 1 & 0 \\ 0 & 0 & 0 & . & . & -\rho_2 & -\rho_1 & 1 \end{pmatrix} \qquad \text{If } U_t \text{ follows } AR(2)$$

multicollinearity Problem :

Originally multicollinearity meant the existence of a (perfect) or exact linear relation among some or all explanatory variables of regression model . For the p variable regression involving explanatory variables $X_1, X_2, \ldots\ldots, X_P$, an exact linear relationship is said to exist if the following condition is satisfied

[Gujarati ,1995:323]:

$$C_1X_1 + C_2X_2 + \ldots\ldots\ldots\ldots + C_PX_P = 0$$

Where:

$C_1, C_2, \ldots., C_P$ are constants thus that not all of them are zero simultaneously.

Multicollinearity is also the name we give to the problem of nearly perfect linear relationships among explanatory variables , this is the more common problem , and it said to exist if the following condition is satisfied :

$$C_1X_1 + C_2X_2 + \ldots\ldots\ldots\ldots + C_PX_P + Vi = 0$$

Where : V_i is stochastic error term.

Note that the perfect collinearity does not usually happen except in the case of the dummy variable trap

Use of Eigenvalues and Eigenvectors:

Suppose we consider the $X'X$ matrix (correlation form) we know that there exist an orthogonal matrix [Myres,1986:86] .

$$V = [V_1, V_2, \ldots\ldots\ldots\ldots, V_P]$$

Such that : $V`(X`X)V = \quad diag(\lambda_1, \lambda_2, \ldots\ldots, \lambda_P)$ …………….(11)

The λ_i is the *i-th* eigenvalue of the correlation matrix , the column of V are normalized eigenvectors associated with eigenvalues of $(X'X)$.

For our purpose here, we need to denote the i^{th} element of the vector V_i by V_{ij} . Now if multicollinearity is present at least one λ_i @ 0 thus , we write , for at least one value of j , $V`_j(X`X)V_j$ @ 0 .Which implies that for at least one eigen-vector V_j ,

Thus the number of small eigenvalues of the correlation matrix relate to the number of multicollinearities according to the definition in (11) and the "weights" C_i are the individual elements in the associated eigenvectors.

There are several methods that have been proposed to remedy multicollin-earity problem by modification the method of OLS to allow biased estimators of regression coefficients , these methods are ridge regression , principal components regression and latent roots regression in this study we used ridge regression.

Ridge Regression:

This method first suggested by Hoerl in 1962 , it discussed at length by Ho-erl and Kenard in (1970) [Hoerl & Kennard ,1970] .The ridge regression estima-tors are obtain by introducing into the least squares estimator b a biasing constant

$C^{3}0$ therefore, yields the ridge regression coefficients[Ghufran ,1990]:

$$\underline{b}_R = (X'X + \underline{C}I)^{-1} X'\underline{Y} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (12)$$

When all variables are transformed to the correlation form, the ridge regression coefficients become:

$$\underline{b}_R = (r_{xx} + \underline{C}I)^{-1} r_{xy} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (13)$$

Where:

r_{xx}: $(p'p)$ matrix containing the pairwise coefficients of simple correlation between the explanatory variables.

r_{xy}: $(p'1)$ vector containing the coefficients of simple correlation between the dependent variable and the explanatory variables.

The form (13) of b_R called standardized ridge regression coefficients.

The constant C reflects the amount of bias in the estimators , when $C = 0$ (13) reduces to *OLS* regression coefficients in standardized form, when $C > 0$ the ridge regression coefficients are biased but tend to be more stable than *OLS* estimators (in application the interesting value of C usually lie in the range (0,1)) [Draper & Smith ,1980:313] .The ridge regression estimator has two forms [Kadiyala,1980] :

When the constant C takes a sequential values ($C=CI$),ridge regression estimator called **ordinary ridge regression estimator** as in (12)

When C takes an estimated values ($C_i=diag(c)$) , ridge regression estimator called **generalized ridge regression estimator** , and written as

$$b_R = (X'X + diag(C))^{-1} X'Y \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(14)$$

We can write the ridge regression estimators of equation (13) using the eigen values and eigen vectors as follows ($C=CI$):

$$b_R = V(\lambda + CI)^{-1} V' r_{XY} = \sum_{i=1}^{p} (\lambda_i + C)^{-1} V_j V_j' \; r_{XY}$$

And when ($C_i = diag(C)$):

$$b_R = \sum_{i=1}^{p} (\lambda_i + C_i)^{-1} V_j V_j' \; r_{XY} \quad .$$

The Mean square error <u>when C=CI</u> is given as:

$$MSE\ (b_R) = C^2 \sum_{i=1}^{p} \frac{a_i^2}{(\lambda_i + C)^2} + \sigma_u^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + C)^2} \quad \ldots\ldots\ldots\ \ldots..(15) \quad .$$

Where

$$a_i^2 = \sum_{i=1}^{p} (V_i' b_{ols})^2$$

and <u>when C_i=diag(c)</u> is given as:

$$MSE\,(b_R) = C^2 \sum_{i=1}^{p} \frac{1}{(\lambda_i + C)^2} \sum_{j=1}^{p} (C_j B_j V_{ij})^2 \quad .$$
$$+ \sigma_u^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + C)^2} \quad \ldots\ldots\ldots\ldots\ldots(16)$$

Note that the first term of *MSE* in equations (15) and (16) presents the $Bias^2 (b_R)$ which is increasing function of C , and the second one is $Var(b_R)$ which is decreasing function of C , this means we accepted some bias in order to decrease the variance [El Any,1987] .

Choice of biasing C:

There are many procedures suggested to choose the optimum value of C , in this study we used Hoerl , Kenard & Baldwin technique as one of the simulation techniques when C is a constant ,this technique takes the following formula [Hoerl & others,1975]

$$C_h = pS^2/b`_{ols}b_{ols}$$

Where
p: number of the estimated parameters

S^2: mean square errors of OLS estimators

b_{ols} : vector of OLS estimators

And we used the following formula for $C_i=diag(C)$ [Dwived,1980]**:**

$$C_i = S^2/b_i^2{}_{ols} \quad i = 1,2,\ldots\ldots,p$$

Where: b_{iols} : the i-th estimator of OLS.

Simulation Experiment :

We used computer simulation to generate data contain multicollinearity and autocorrelation both .Delphi Language (version 6.0) used to construct statistical package designed by researcher.

Step1: Generate a random number U which follows uniform distribution $U\sim U(0,1)$.

Step2: Generate a random variable V_t which follows standard normal distribution

$V_t \sim N(0,1)$, we used Box – Mullar , Polar and Inverse transform methods

Step3: Compute the disturbance term U_t which follows :-

AR(1) $\quad U_t = \rho U_{t-1} + V_t \qquad -1 \leq \rho \leq 1$

AR(2) $\quad U_t = \rho_1 U_{t-1} + \rho_2 U_{t-2} + V_t \qquad -1 \leq \rho_1, \rho_2 \leq 1$

Step4: Generate an explanatory variables :

Generate X_1:

We used a random number generation to generate X_1, ($X_1 \sim U(0,1)$) b-

Compute $X_2, X_3, \ldots\ldots\ldots\ldots\ldots, X_p$

Two methods used to compute : $X_2, X_3, \ldots\ldots\ldots\ldots\ldots, X_p$:

MethodI [Ghufran ,1990]:

$$X_2 = g_1 + g_2 X_1 + D$$
$$X_3 = X_2 + g_3 X_2$$
$$.$$
$$.$$
$$X_p = X_{p-1} + g_p X_{p-1}$$

Where

$g_1, g_2, \ldots\ldots\ldots\ldots, g_p$ are arbitrary .

D : (n´1) vector follows uniform distribution ($U(0,1)$) .

MethodII [Yue Fang and Sergio,2003]

$$X_j = D_{j-1} + X_{j-1} \qquad j = 2,3,\ldots\ldots\ldots,p$$

Step5: Compute dependent variable Y_t :

$$Y_i = B_0 + B_1 X_{1t} + B_2 X_{2t} + \ldots\ldots + B_p X_{pt} + U_t \qquad t = 1,2,\ldots\ldots,n$$

Where: $B_0, B_1, B_2, \ldots\ldots, B_p$: are arbitrary.

Step6: *Standardize* Y_t *and* X_s

Step7: Applying *OLS* .

Dealing with Autocorrelation:

We used the Two Stages least squares procedure (2SLS) to deal with the problem of autocorrelation as follows :

Transform the original generated data

Standardize the transformed data .

Applying OLS .

Dealing with Multicollinearity:

After we eliminated the problem of autocorrelation ,we used the ridge regression method to deal with the problem of multicollinearity.

Figure (1) shows the main form of the program that use to generate data.

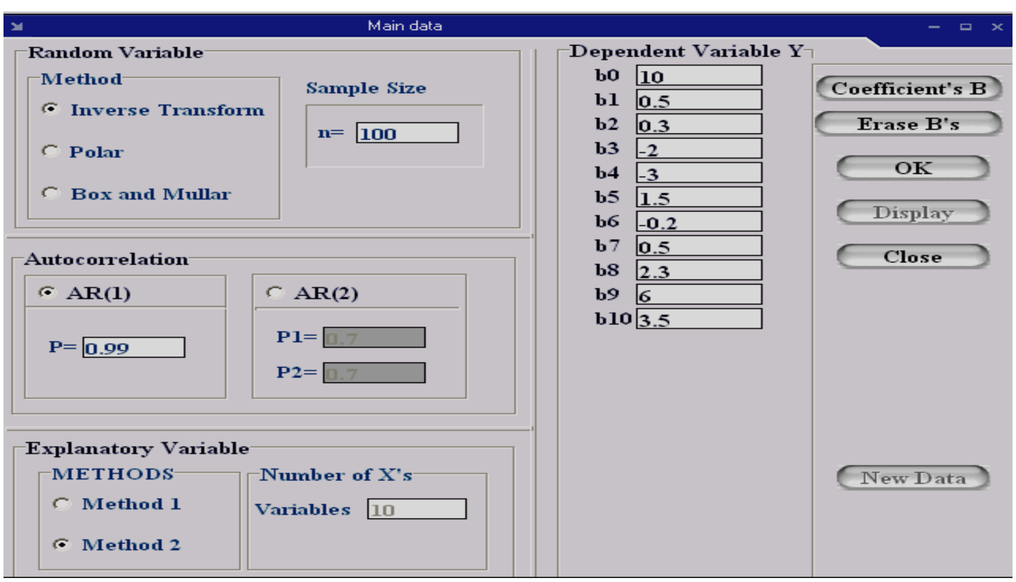

Figure 1 : *The main form use to generate data*

Models Criteria:

The package designed to appropriate, different sample sizes from 2 to infinity, number of explanatory variables from 2 to 45 and it's more flexible in choosing the autocorrelation coefficients.

In order to make this study more inclusive we choose the following criteria to the models:

Sample size:

We choose the following sizes as small samples sizes

$$n = 5,10,20$$

We choose the following sizes as large samples sizes

$$n = 30,50,100,200$$

Number of explanatory variables:

We choose different number of explanatory variables as follows :

$$p = 2,5,10,20$$

Dependent variable Y:

According to a number of explanatory variables p we compute Y where :

$$Y_t = B_0 + B_1 X_{1t} + B_2 X_{2t} + \ldots\ldots + B_p X_{pt} + U_t$$

$$p = 2,5,10,20$$

The models are :

$$Y_t = -2 + 0.5 X_{1t} + 0.7 X_{2t}$$

$$Y_t = 5 + 8X_{1t} - 9X_{2t} + 3X_{3t} - 6X_{4t} + 0.7X_{5t}$$

$$Y_t = 10 + 0.5X_{1t} + 0.3X_{2t} - 2X_{3t} - 3X_{4t} + 1.5X_{5t} - 0.2X_{6t} + 0.5X_{7t} + 2.3X_{8t} + 6X_{9t} + 3.5X_{10t}$$

$$Y_t = 2 + 3X_{1t} + 5X_{2t} - 0.2X_{3t} - 0.3X_{4t} - 0.5X_{5t} + 1.2X_{6t} + 1.3X_{7t} + 1.5X_{8t} + 0.02X_{9t} + 0.03X_{10t} + 0.05X_{11t} - 2X_{12t} - 3X_{13t} - 5X_{14t} + 12X_{15t} + 13X_{16t} + 15X_{17t} - .04X_{18t} - 0.06X_{19t} + 0.03X_{20t}$$

Autocorrelation Coefficients:

We choose the following values for the autocorrelation coefficients:

AR(1) : $r = \pm 0.99$, ± 0.7 , ± 0.3

AR(2) : $r_1, r_2 = \pm 0.99$, ± 0.7 , ± 0.3

Conclusions:

From the results obtained in this study , we can conclude that :

Dealing with autocorrelation from data that suffered from multicollinearity :

Multicollinearity increases when the error term follows first or second order autoregressive scheme.

Multicollinearity decreases if the model has a few explanatory variables.

When the error term follows first order autoregressive scheme :

MSE for ridge regression and ordinary least squares methods decreasing rapidly as sample size increase when an autocorrelation coefficient greater than or equal to |0.99|.

MSE for ridge regression method increases as the number of explanatory variables increase when an autocorrelation coefficient is less than or equal to |0.3| .

When the sample size is too large the OLS and ridge regression method smallest S^2 and largest R^2 .

Ordinary least squares method has largest S^2 and smallest R^2 when an autocorrelation coefficient is greater than or equal to |0.99|.

For small and large sample sizes the significance of the models estimated by OLS and ridge regression becomes more strong as the sample size and the number of explanatory variables increase .

When the error term follows second order autoregressive scheme :

MSE for ordinary least squares method decreases as sample size increases when an autocorrelation coefficient is greater than or equal to |0.7| .

MSE for ordinary least squares method decreases as number of explanatory variables increase when an autocorrelation coefficient less than or equal to |0.7| .

Ordinary least squares models are not significant.

The S^2 for Ordinary least squares method decreases as sample size increases.

When the sample size is too large the models estimated by OLS and ridge regression are not significant .

When an error term follows first or second order autoregressive scheme

Ordinary least squares method has largest *MSE*

When the sample size is too large the *MSE* for OLS and ridge regression increases , whereas, it decreases when the explanatory variables are too much .

The *MSE* for ordinary ridge regression greater than the *MSE* for generalize ridge regression, whereas, the opposite occurs when the sample size is too large .

The variance for ordinary ridge regression (generalize ridge regression) greater than the $biased^{\wedge 2}$ for ordinary ridge regression (generalize ridge regression).

Among the types of ridge regression method , if we take the *MSE* as a criterion of comparison we find that ordinary ridge is the best when the sample size is too large , otherwise , generalized ridge is the best one.

References

1. **Draper, N.R. & Smith, H. (1980)** " Applied Regression Analysis " , John Willey and Sons INC., New York .
2. **Durbin , J. (1970)** "Testing for Serial Correlation in Least Squares Regression when Some of the Regressors are Lagged Dependent Variables", Econometrica , Vol. 38, PP. 410-429.
3. **Durbin , J. (1970)**"An alternative to Bounds Testing for Serial Correlation in Least Squares Regression ", Econometrica ,Vol. 38, PP.422-429.
4. **Dwived , T.D. (1980)** "Finite Sample properties of Ridge Estimators",Technometrics , Vol. 22 , PP. 205-212.
5. **El Any, A.T. (1987)** "A comparison Study of Some Biased Estimation Methods", M.Sc. Thesis, Mustansiria University , Baghdad,IRAQ
6. **Ghufran, I.K.(1990)** " A comparison Study of Biased Estimation Methods Under wrong assumptions using simulation" , M.Sc Thesis, Baghdad University , Baghdad,IRAQ.
7. **Haldance, J. B. S. (2000)** "Autocorrelation Models" , British Scientist.
8. **Gajarati,D.N.(1995)** "Basic Econometrics",3rd ed., McGraw Hill Book company , London.
9. **Hoerl, A. E. & Kennard, R.W. (1970)** "Ridge Regression Application to Nonorthognal Problems", Technometrics , Vol. 12, PP. 69-82.
10. **Hoerl, A. E. & Kennard, R.W.(1970)** " Ridge Regression Biased Estimation for Nonorthognal Problems",Technometrics,Vol.12, PP.55-67
11. **Hoerl, A. E.,Kennard, R.W.& Baldwin , K.F.(1975)** "Ridge Regression: Some Simulations",Commun. Stat.,Vol.4, PP.105-123.
12. **John Nester,William & Kulner, H. (1985)** "Applied Linear Statistical Models" ,2nd ed. IRWIN.

13. **Johnston, J. (1972)** "Econometric Methods", 2nd ed. McGraw Hill Book company , New York.
14. **Kadiyala K. R.(1968)** "Transformation used to Circumvent the Problem of Autocorrelation", Econometrica , Vol. 36 , PP. 93-96.
15. **Kadiyala K. R.(1980)** "Some finite sample properties of Generalized Ridge Regression Estimators", The Canadian Journal of Statistics , Vol.8 , No. 1,PP. 47-57.
16. **Koutsoyannis, A. (1977)** "Theory of Econometrics " 2nd ed. Mac., Millan.
17. **Marquardt, D.W.(1970)**"Generalized Inverses: Ridge Regression Biased Linear Estimation and Nonlinear Estimation",Technometrics, Vol. 12, PP. 591-612.
18. **Myres , Rymond, H.(1986)** "Classical and Modern Regression with Application", Boston ,Duxburg Press C.
19. **Naylor T.H. & Others (1966)** " Computer Simulation Technique ", John Willey and Sons INC. New York
20. **Ronald, U. Mendoza (2002)** "Serial Autocorrelation" , Fordam Dep. of Economics.
21. **Shanon R.E.(1975)** "System Simulation:The Art and Science" , New Jersey : Printice – Hall.
22. **Webster ,J.T. ,Gunt ,R.F. & Mason ,R.L.(1974)** "Latent Roots Regression Analysis" , Technometrics, Vol. 16 ,PP.513-522.
23. **Younis,B.,Mosa,A.&Ameer,A.(2002)**"Econometrics",1st ed.,Dar Azza for publishing , Khartoum Sudan.