

مقارنة طرق تقدير معالم نموذج الانحدار في حالة ظهور مشكلة التعدد الخطي والقيم الشاذة

م. غفران اسماعيل كمال

أ.م.د. نزار مصطفى جواد

كلية الادارة والاقتصاد
جامعة بغداد

المستخلص

Ridge RLAV Ridge LSE
MM Ridge MM WRID
MM Ridge RMM RMM
RMM

ABSTRACT

A simulation study is used to examine the robustness of some estimators on a multiple linear regression model with problems of multicollinearity and non-normal errors, the Ordinary least Squares (LS), Ridge Regression, Ridge Least Absolute Value (RLAV), Weighted Ridge (WRID), MM and a robust ridge regression estimator MM estimator, which denoted as RMM this is the modification of the Ridge regression by incorporating robust MM estimator. Finally, we show that RMM is the best among the other estimators.

1- المقدمة وهدف البحث

$$\sum_{i=1}^n (y_i - x_i' \beta)^2 = \sum_{i=1}^n (r_i)^2$$

β

$$\hat{\beta}_{LS} = (X'X)^{-1} X'Y \text{-----(1)}$$



Ridge -2

$(X'X)$

β

Kennard Horel (1970) Ridge β LS

Conditional number (CN)

: CN $\cdot (X'X) X$

$$CN = \frac{\lambda_{\max}}{\lambda_{\min}} \geq 1 \text{ -----(2)}$$

$\cdot (X'X)$

λ

Belsley (1980)

CN

10

5

. 100

30

CN

Horel and Kennard

(a,b) (1970)

$(K > 0)$

:

$X'X$

$$\beta^{\wedge}(k) = (X'X + kI)^{-1} X'Y \text{ -----(3)}$$

:

$(p \times p)$

I

$k > 0$

b)

k

: k

Kennard ,Horel (1970)

$$k = \frac{pS_{LS}^2}{\beta'_{LS}\beta_{LS}} \text{ -----(4)}$$

$$S_{LS}^2 = \frac{(Y - X\beta_{LS}^{\wedge})'(Y - X\beta_{LS}^{\wedge})}{n - p} \text{ -----(5)}$$

p

$$\beta_{LS}^{\wedge} = \beta_{RID}^{\wedge} \quad k=0$$

$$\beta_{RID}^{\wedge} \quad k>0$$

$$\beta_{RID}^{\wedge} \rightarrow 0 \quad k \rightarrow \infty$$

Kennard & Horel (1970 a)

$$MSE_{\beta_{RID}^{\wedge}} < MSE_{\beta_{LS}^{\wedge}}, \quad k > 0$$



Edgeworth (1887)

M

Huber (1973) (LAV)

Rousseeuw (LMS)

(1987)

Yohai

(1987) 50%

(LTS)

.MM

:

. 50%

θ_0

T_0

-

-

$$r_i(T_0) = y_i - T_0'x_i \text{ -----(6)}$$

$$1 \leq i \leq n$$

M-

$$S_n = S(r(T_0)) \text{ ----- (7)}$$

Huber (1973)

M -

ρ_0

:

$$\frac{c}{d} = 0.5 \text{ -----(8)}$$

$$d = \max \rho_0(u)$$

M -

:

M -

ρ_1

-

$$\rho_1(u) \leq \rho_0(u) \text{ -----(9)}$$

$$\sup \rho_1(u) = \sup \rho_0(u) = d \text{ -----(10)}$$

ρ

$\psi(t)$

:

$$\psi(t) = \rho'(t)$$

$$\rho(t), \psi(t)$$



T_1 MM -

$$\sum_{i=1}^n \psi_i \left(\frac{r_i}{S_n} \right) x_i = 0 \text{ -----(11)}$$

$$S(T_1) \leq S(T_0) \text{ -----(12)}$$

$$S(\theta) = \sum_{i=1}^n \rho \left(\frac{r_i}{S_n} \right) \text{ -----(13)}$$

Ridge -4

$\hat{\beta}_{RID}$

Ridge

Ridge

Ridge

Ridge

Ridge

Montgomery

Askin (1980)

Ullah ,Vinod (1981)

Ridge

pfaffenberger, Dielman (1984)

(WRID)

.(RLAV)

Ridge -4-1

(1980) Askin ,Montgomery

:

$$\hat{\beta}_{wls} = (X'WX)^{-1} X'WY \text{ -----(14)}$$

W

OLS

$\hat{\beta}_{WLS}$

$$x_i \rightarrow \sqrt{w_{ii}} x_i$$

$$y_i \rightarrow \sqrt{w_{ii}} y_i$$

$$w_i = \frac{\psi(y_i - x_i' \beta)}{(y_i - x_i \beta)} \text{ -----(15)}$$



WRID

$$\hat{\beta}_{WRID} = (x'wx + kI)^{-1} x'wy \text{ -----(16)}$$

Ridge

Ridge (RLAV) , Lawrence (1984) Ridge

$$\hat{\beta}_{RLAV} (x'x + k^* I)^{-1} x'y \text{ -----(17)}$$

k (19) (18)

k^*
:

$$k^* = \frac{\rho S_{LAV}^2}{\hat{\beta}'_{LAV} \hat{\beta}_{LAV}} \text{ -----(18)}$$

$$S_{LAV}^2 = \frac{(y - x\hat{\beta}_{LAV})'(y - x\hat{\beta}_{LAV})}{n - p} \text{ ----- (19)}$$

$\hat{\beta}_{LAV}$

$$\min_{\beta} \sum_{i=1}^N |y_i - x'_i \beta|$$

-5

RMM

(25,50)

:

0.5 0.8 0.9

. 3

t

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

$$\beta_0 = \beta_1 = \beta_2 = 1$$



$$x_{ij} = (1 - \rho^2)z_{ij} + \rho z_{ij} \quad \begin{matrix} i=1,2,\dots,n \\ j=1,2 \end{matrix}$$

z_{ij}

(1000)
(SE)

(RMSE)

$$\begin{aligned} Bias &= E(\hat{\beta}) - \beta \\ &= \bar{\beta}_j - \beta_j \end{aligned}$$

$$\bar{\beta}_j = \frac{\sum_{j=1}^L \beta_j}{L} \quad L=1000$$

$$MSE(\hat{\beta}_j) = E(\hat{\beta} - \beta)^2 = (\bar{\beta}_j - \beta_j)^2 + \frac{1}{L} \sum_{j=1}^n (\beta_j^{\wedge} - \bar{\beta}_j)^2$$

$$[Var(\hat{\beta}_j) + (Bias)^2]^{\frac{1}{2}} \quad (MSE(\hat{\beta}_j))^{\frac{1}{2}}$$

β_1, β_2

β_0

(5,3,1)

(6,4,2)

تحليل نتائج تجربة المحاكاة

			LS		(1)
				LS	(2)
	OLS	RMM		.LS	
RMM	RLAV	Ridge		MM WRID LS	
			(1)		(2)
		RMM			LS
					RMM
	RLAV	Ridge		RMM	(2)
				WRID , MM , LS	



MM

(3,4)

RMM

RMM, RLAV , WRID

(3)

t -

RMSE

SE

(5)

(5,6)

RMSE

RMM

RMM

(6)

Ridge

MM

RMM

ρ^2



الاستنتاجات

RMM
 RMM WRID ,RLAN ,RMM Ridge
 RLAV ,WRID

المصادر

(1990) -1

- 1- Askin,R.G.and Montgomery,D.C.,(1980)."Augmented Robust Estimators". Technometrics,22,333-341.
- 2- Belsley,D.A.,(1980)."Regression Diagnostics, Identifying influential data and Sources of collinearity". New York, John Wiley.
- 3- Edgeworth, F.Y., (1887) "On observations Relating to Several Quantity, Journal Hermathena , 6,279-285.
- 4- Hoerl, A.E.and Kennard,R.W., (1970 a) ."Ridge regression : iterative Estimation of the Biasing parameter, Communications in statistics : A. Theory methods ,5,77-88.
- 5- Hoerl,A.E. and Kennard ,R.W.(1970,b)"Ridge regression :Applications to nonrthogonal problems, Technometrics, 12,69-82.
- 6- Huber ,P.J.(1973) ."Robust Regression :Asymtotics, conjectures and Monte Carlo . The Annals of statistics,1, 799-821.
- 7- Lawrnce ,K.D. and Arthur ,J.L. (1990) ."Robust Regression :Analysis and Application". New York, Marcel Deker.
- 8- Pfaffenberger ,R.C. and Dielman ,T.E., (1984)."A modified Ridge Regression Estimator Using the Least Absolute Value criterion in the Multiple Linear Regression model ,Proceedings of the American Institute for Decision , Toronto,791-793.
- 9- Vinod ,H.D., and Ullah, A.(1981)"Recent Advances in Regression Methods "New York, Marcel Dekker.



جدول (1) التحيز وجذور متوسطات الخطأ والخطأ المعياري لطريقة المربعات الصغرى وطرق التقدير الأخرى للتوزيع الطبيعي

$$\hat{\beta}_1$$

Values of ρ^2	0.0			0.5			0.95		
Method	Bias	RMSE	S.E	Bias	RMSE	S.E	Bias	RMSE	S.E
LS	-0.154 .0022	.223 .1445	.2189 .1657	-0.089 -0.041	.3321 .2276	.3365 .2341	.0443 -0.0522	3.0011 2.0453	3.0006 2.0564
RIDGE	-0.0645 -0.0231	.2221 .1453	.2341 .1456	-0.0432 -0.0166	.3213 .2210	.3211 .2321	-0.0122 -0.0221	1.5664 1.0543	1.5674 1.0785
RLAV	-0.0124 -0.0200	.2223 .1465	.2144 .1442	-0.0432 -0.0199	.2540 .2231	.3215 .2119	-0.0421 -0.0265	1.5668 1.0774	1.5764 1.0764
WRID	-0.0151 -0.0023	.2744 .1987	.2344 .1654	-0.0122 -0.0088	.4211 .3421	.4321 .3422	.0665 -0.0643	3.5647 2.4567	3.7845 2.3490
MM	-0.0154 -0.0030	.2344 .1786	.3443 .1339	-0.0112 -0.0010	.4332 .2887	.4876 .2887	.0498 -0.0453	3.7856 2.8976	3.7856 2.8976
RMM	-0.0611 -0.0198	.1766 .1433	.2346 .1445	-0.0432 -0.0223	.3765 .2123	.3422 .2987	.0210 -0.0101	1.5678 1.0762	1.534 1.0563

$$\hat{\beta}_2$$

Values of ρ^2	0.0			0.5			0.95		
Method	Bias	RMSE	S.E	Bias	RMSE	S.E	Bias	RMSE	S.E
LS	-0.0053 .0056	.2234 .1438	.2256 .1456	.0096 .0095	.3456 .2376	.3452 .2346	-0.0576 .0577	3.0654 2.0765	3.0111 2.0786
RIDGE	-0.0453 -0.0453	.2245 .1477	.2216 .1478	-0.0234 -0.0076	.3178 .2140	.3454 .2145	.0034 .0056	1.5764 1.0664	1.5342 1.0649
RLAV	-0.0453 -0.0123	.2267 .1456	.2217 .1453	-0.0256 -0.0087	.3096 .2185	.3459 .2134	.0134 .0245	1.5683 1.0885	1.5332 1.0756
WRID	-0.0123 .0034	.2645 .1876	.2675 .1876	.0123 .0134	.4321 .2765	.4387 .2687	-0.0978 .0569	3.5664 2.4533	3.5764 2.3442
MM	-0.0056 .0023	.2678 .1756	.2687 .1711	.0099 .0123	.4123 .2756	.4765 .2548	-0.0443 .00645	3.7694 2.4769	3.7890 2.4768
RMM	-0.0543 -0.0231	.2267 .1456	.2210 .1456	-0.0452 -0.0078	.3078 .2344	.3656 .2534	-0.0476 .0065	1.5340 1.0756	1.5764 1.0701



جدول (2) الكفاءة النسبية للمعاملات التقديرية للتوزيع الطبيعي

Estimator 1vs	Estimator2	β_1^{\wedge}			OF	ρ^2		β_2^{\wedge}
		0.0	0.5	Values 0.95		0.0	0.5	
RMM	LS	1.04	0.97	0.26	0.99	0.78	0.29	
		1.00	0.80	0.27	0.95	0.88	0.27	
	RID	1.01	1.01	1.01	1.01	0.98	0.99	
		1.00	1.00	1.00	1.02	1.00	0.95	
	WRID	0.66	0.55	0.16	0.68	0.55	0.18	
0.63		0.63	0.23	0.75	0.63	0.22		
RLAV	1.01	1.01	0.98	0.98	1.01	0.99		
	1.00	1.00	1.00	0.99	1.00	0.95		
MM		0.68	0.58	0.16	0.76	0.66	0.17	
		0.71	0.64	0.19	0.63	0.63	0.19	
MM	LS	1.54	1.42	1.52	1.38	1.54	1.54	
		1.33	1.44	1.44	1.44	1.33	1.48	
	RID	1.52	1.87	5.52	1.32	1.87	5.62	
		1.44	1.63	5.44	1.36	1.63	5.74	
WRID	0.97	0.97	1.11	0.98	0.97	1.01		
	0.80	0.96	1.10	0.88	0.96	1.00		
RLAV	1.54	1.91	5.62	1.34	1.83	5.62		
	1.43	1.65	5.44	1.36	1.54	5.44		
RLAV	LS	1.01	0.97	0.26	1.02	0.78	0.29	
		1.00	0.81	0.28	0.99	0.80	0.27	
	RID	1.03	1.01	1.01	1.01	1.01	1.01	
1.00		1.00	1.00	1.00	1.00	1.0		
WRID	LS	0.66	0.56	0.16	0.76	0.56	0.29	
		0.63	0.53	0.20	0.63	0.53	0.27	
	RID	1.55	1.54	1.32	1.54	1.54	1.54	
1.65		1.43	1.24	1.48	1.43	1.43		
WRID	RID	1.52	1.92	0.29	1.44	1.87	5.62	
		1.74	1.73	0.27	1.53	1.63	5.44	
RID	LS	1.04	0.78	0.29	1.02	0.76	0.29	
		1.00	0.80	0.27	0.99	0.83	0.27	



جدول (3) التحيز وجذور متوسطات الخطأ والخطأ المعياري لطريقة المربعات الصغرى وطرق التقدير الأخرى لنوزيع كوشي

$$\hat{\beta}_1$$

Values of ρ^2	0.0			0.5			0.95		
	Bias	RMSE	S.E	Bias	RMSE	S.E	Bias	RMSE	S.E
LS	2.786 .0123	122.564 18.564	125.908 18.786	2.564 .345	60.564 32.896	60.664 32.453	6.278 5.786	263.786 340.785	263.786 340.765
RIDGE	.089 -.432	44.765 7.895	44.786 7.786	.756 -.231	27.785 9.765	27.454 8.765	-1.076 .456	50.764 73.342	50.734 73.987
RLAV	-.534 -.578	.756 .786	.576 .534	-.534 .567	.786 .776	.665 .576	-.234 -.345	1.856 1.056	1.987 1.007
WRID	-.290 -.276	.576 .423	.456 .389	-.124 -.189	.655 .433	.576 .456	.189 -.134	3.376 1.459	3.564 1.456
MM	.043 -.002	.489 .256	.576 .342	.006 -.027	.789 .453	.634 .476	.576 -.256	6.342 3.997	6.785 3.897
RMM	-.564 -.755	.789 .786	.534 .568	-.455 -.546	.756 .790	.556 .589	-.126 -.345	1.534 .956	1.564 .908

$$\hat{\beta}_2$$

Values of ρ^2	0.0			0.5			0.95		
	Bias	RMSE	S.E	Bias	RMSE	S.E	Bias	RMSE	S.E
LS	3.976 -.238	144.432 17.432	143.897 17.564	1.234 -.645	78.897 36.896	78.876 36.984	-5.342 -5.564	245.897 344.765	245.908 345.876
RIDGE	.756 -.345	41.876 6.985	42.776 6.954	-.123 -.365	21.563 9.563	21.876 9.564	1.786 -.876	46.765 74.908	45.907 74.786
RLAV	-.576 -.534	.7543 .776	.553 .566	.566 -.544	.776 .779	.554 .564	-.232 -.334	1.908 1.098	1.987 1.094
WRID	-.234 -.299	.556 .455	.433 .432	-.234 -.231	.554 .590	.559 .522	-.267 -.121	3.443 1.432	3.432 1.987
MM	-.012 .023	.487 .254	.445 .231	-.022 .032	.633 .443	.645 .498	-.564 -.238	6.786 3.908	6.321 3.943
RMM	-.534 -.588	.732 .775	.654 .687	-.543 -.587	.765 .786	.543 .556	-.349 -.228	1.786 1.897	1.654 .998



جدول (4) الكفاءة النسبية للمعاملات التقديرية لتوزيع كوشي

Estimator 1 vs	Estimator2	β_1^{\wedge}			Values	OF	ρ^2		β_2^{\wedge}
		0.0	0.5	0.95			0.0	0.5	
RMM	LS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	
	RID	0.00	0.01	0.01	0.00	0.00	0.01	0.01	
		0.001	0.07	0.00	0.014	0.06	0.00	0.00	
	RLAV	1.00	0.94	0.67	1.00	0.95	0.67	0.67	
1.00		1.00	0.90	1.00	0.98	0.87	0.87		
WRID	2.25	1.55	0.21	2.28	1.26	0.21	0.21		
	2.99	2.44	.47	2.39	2.44	0.07	0.07		
MM	3.18	1.56	0.06	3.06	1.26	0.07	0.07		
	7.71	2.32	0.09	7.23	2.83	0.09	0.09		
MM	LS	0.00	0.00	0.00	0.00	0.00	0.00		
		0.00	0.00	0.00	0.00	0.00	0.00		
	RID	0.00	0.001	0.012	0.00	0.001	0.018		
		0.01	0.002	0.003	0.002	0.002	0.003		
RLAV	0.32	0.73	10.87	0.33	0.76	10.77			
	0.13.	0.34	13.77	0.14	0.34	13.66			
WRID	0.54	1.21	3.62	0.74	1.33	3.32			
	0.43	0.85	7.44	0.36	0.84	7.44			
RLAV	LS	0.00	0.00	0.00	0.00	0.00	0.00		
		0.00	0.00	0.00	0.00	0.00	0.00		
	RID	0.00	0.001	0.001	0.00	.001	.001		
1.00		0.007	0.000	0.014	.006	0.00			
WRID	2.66	1.56	0.36	2.76	1.56	0.29			
	2.63	2.53	0.50	2.63	2.53	0.57			
WRID	LS	0.00	0.00	0.00	0.00	0.00	0.00		
		0.00	0.00	0.00	0.00	0.00	0.00		
RID	0.00	0.00	0.00	0.00	1.87	0.005			
	0.00	0.00	0.00	0.005	0.002	0.00			
RID	LS	0.14	0.28	0.05	0.08	0.08	0.04		
		0.18	0.07.	0.06	0.17	0.05	0.05		



جدول (5) التحيز وجذور متوسطات الخطأ والخطأ المعياري لطريقة المربعات الصغرى وطرق التقدير الأخرى لتوزيع t- بدرجة حرية 3

$$\hat{\beta}_1$$

Values of ρ^2	0.0			0.5			0.95		
Method	Bias	RMSE	S.E	Bias	RMSE	S.E	Bias	RMSE	S.E
LS	-.0012 -.0056	.3765 .2355	.3452 .2785	.0123 -.0453	.5764 .3452	.5566 .3224	.2331 -.1342	5.6543 3.564	5.1324 3.7856
RIDGE	-.0987 -.0564	.3247 .2332	.3490 .2543	-.0342 -.0554	.4432 .3498	.4332 .3432	.0987 -.0453	2.7645 1.8958	2.5433 1.7653
RLAV	-.1234 -.0972	.3421 .2112	.3298 .2112	-.0121 -.0543	.4231 .3897	.4338 .3555	.1234 -.2314	1.7856 1.3490	1.7654 1.1233
WRID	-.0022 -.01223	.3442 .2217	.2998 .2113	-.0543 -.0004	.4879 .3221	.4432 .3324	.1453 -.0231	3.8754 2.6545	3.8976 2.7623
MM	.0034 -.0765	.3445 .2765	.3421 .2218	-.0054 -.0088	.5341 .3987	.5443 .3321	.2341 -.0875	4.7856 2.8976	4.2312 2.9987
RMM	-.1342 -.0543	.3765 .2786	.3329 .2221	-.0983 -.0453	.4875 .3287	.3998 .3221	.1121 -.0564	1.8976 1.8977	1.7654 1.2212

$$\hat{\beta}_2$$

Values of ρ^2	0.0			0.5			0.95		
Method	Bias	RMSE	S.E	Bias	RMSE	S.E	Bias	RMSE	S.E
LS	-.0007 -.0013	.3455 .2345	.3456 .2987	.0087 .0123	.5764 .3762	.5764 .3789	-.2341 .1564	5.1434 3.5644	5.1433 3.4566
RIDGE	-.0934 -.0542	.3667 .2667	.3564 .2134	-.0645 -.0345	.4563 .3256	.4576 .3121	-.1456 .0234	2.3454 1.2987	2.5433 1.6754
RLAV	-.0176 -.0056	.3978 .2134	.3342 .2221	-.0723 -.0398	.4571 .3121	.4567 .3111	-.1356 -.0045	1.7543 1.3324	1.7756 1.1543
WRID	-.0453 -.0056	.3775 .2456	.3987 .2887	-.0023 .0003	.4765 .3112	.5762 .3452	-.1234 .0934	3.6554 2.5996	3.4589 2.4766
MM	.0033 .0043	.3998 .2111	.3122 .2776	.0034 .0156	.5763 .3121	.4378 .3211	-.2234 .0945	4.5633 2.9876	4.2223 2.8343
RMM	-.0123 -.0234	.3448 .2658	.3514 .2113	-.0765 -.0633	.4235 .3116	.4456 .3214	-.1234 -.0015	1.7123 1.1124	1.7433 1.1133



جدول (6) الكفاءة النسبية للمعاملات التقديرية لتوزيع t

Estimator 1 vs	Estimator2	β_1^{\wedge}			β_2^{\wedge}		
		0.0	0.5	Values 0.95	OF 0.0	ρ^2 0.5	0.95
RMM	LS	0.82	0.49	0.11	0.87	0.53	0.11
		0.99	0.69	0.10	0.94	0.67	0.10
	RID	0.92	0.84	0.47	0.97	0.87	0.22
		0.92	0.92	0.44	0.94	0.94	0.23
	RLAV	1.00	0.94	0.97	1.00	0.95	0.97
	1.00	1.00	0.90	1.00	0.98	0.87	
WRID		1.25	0.66	0.21	1.28	0.66	0.25
		1.39	1.07	0.22	1.39	0.94	0.26
MM		1.18	0.56	0.15	1.16	0.66	0.15
		1.71	1.02	0.16	1.23	1.00	0.15
MM	LS	0.61	0.80	0.74	0.78	0.79	0.74
		0.74	0.67	0.70	0.60	0.67	0.70
	RID	0.69	1.38	3.04	0.84	1.27	3.07
		0.72	0.90	2.97	0.67	0.92	2.99
RLAV		0.73	1.73	6.87	0.83	1.55	6.12
		0.74	0.99	6.13	0.64	0.94	6.66
WRID		0.94	1.21	1.32	1.04	1.13	1.32
		1.03	1.05	1.44	0.96	0.94	1.44
RLAV	LS	0.83	0.50	0.12	0.88	0.55	0.12
		0.99	0.69	0.11	0.94	0.67	0.11
	RID	0.93	0.50	0.49	0.95	.88	.94
0.97		0.96	0.48	0.96	.93	0.94	
WRID		1.36	0.76	0.22	1.26	0.77	0.29
		1.33	1.03	0.24	1.43	0.98	0.27
WRID	LS	0.63	0.72	0.55	0.73	0.71	0.56
		0.69	0.64	0.47	0.65	0.70	0.47
RID		0.71	1.23	2.23	0.77	1.13	2.30
		0.68	0.85	1.87	0.76	0.99	2.00
RID	LS	0.89	0.58	0.24	0.92	0.62	0.24
		1.08	0.76	0.23	0.98	0.73	0.25