

AFOSR-TR-78-0732

2

3

AD A 0536

MILKING DATA THROUGH BIASED REGRESSION TECHNIQUES

Webster and R.F. Gunst, Southern Methodist University

be

The recent interest in biased estimation procedures in multiple linear regression arises from the large variances of the least squares estimators (unbiased) of the regression coefficients when multicollinearities are present. The biased estimation procedures greatly reduce this variance at the cost of some bias. It is the purpose of this paper to look at this bias with reference to the nature of the specific problem being investigated.

For a structure upon which to base this discussion let the model be of the form

$$Y_{i} = \beta_{o} + \beta_{1}X_{i1} + \beta_{2}X_{i2} + \dots + \beta_{p}X_{ip} + \epsilon_{i}$$

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

or $Y = 1 \beta_{p} + X \beta + \epsilon$ (1)

where $\underline{Y}' = [Y_1, Y_2, \dots, Y_n]; \underline{1}$ is an nxl vector of ones; β_0 and $\beta' = [\beta_1, \beta_2, \dots, \beta_p]$ are unknown parameters; $X = \{X_{ij}\}$, an nxp matrix of known values of rank p with $X'l = \phi$ and the diagonal elements of X'X equal to one; and $\varepsilon' =$

$$[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]$$
 with ε_i a random variable, $E(\varepsilon_i) = 0, E(\varepsilon_i^2) = \sigma^2$ and $E(\varepsilon_i \varepsilon_i) = 0, i \neq i'$.

There are basically two questions addressed in a multiple linear regression analysis: i) The estimation of β , and ii) the estimation of Y, a future value, when $X_1, X_2 \dots X_p$ are given. The answers are not necessarily the same as will become apparent in our discussion. We shall in essence be looking for a solution to i) but also show that a poor solution to i) may still be a good solution to ii) in a restricted sense.

A Decomposition of β

When problems with multicollinearities arise, a tool often used is to express β as a linear function of the latent vectors of X'X (e.g. Hocking et al. [1976]):

$$\underline{B} = \gamma_1 \underline{\alpha}_1 + \gamma_2 \underline{\alpha}_2 + \dots + \gamma_p \underline{\alpha}_p \qquad (2)$$

where a is the latent vector of X'X associated with the latent root λ_{ρ} . Placing this form of β into model (1) yields

 $\underline{Y} = \underline{1}\beta_0 + \gamma_1\underline{z}_1 + \gamma_2\underline{z}_2 + \dots + \gamma_p\underline{z}_p + \underline{\varepsilon},$ where $\underline{z}_{\ell} = X \underline{\alpha}_{\ell}$. The \underline{z}_{ℓ} are mutually orthogonal hence the $\gamma_{\ell} = \underline{z}_{\ell}^{\prime} \underline{y} / \underline{z}_{\ell}^{\prime} \underline{z}_{\ell}$ are uncorrelated, unbiased estimators of γ_{l} with variance σ^{2}/λ_{l} . A consequence of multicollinearities is some small λ_{ℓ} (Mason et al. [1975]); hence, some γ_{ℓ}

will have large variances. The least squares estimator of β will also have some elements with large variances as can be seen from

$$\overline{\mathbf{p}} = (\mathbf{x}, \mathbf{x})_{-\mathbf{x}} \mathbf{x}, \overline{\mathbf{x}} = \lambda^{1} \overline{a}^{1} + \lambda^{5} \overline{a}^{5} + \cdots + \lambda^{5} \overline{a}^{5} \mathbf{b}^{2}; \quad (3)$$

giving
$$\operatorname{Var}(b_j) = \sigma^2 \left[\frac{a_{j1}}{\lambda_1} + \frac{a_{j2}}{\lambda_2} + \ldots + \frac{a_{jp}}{\lambda_p} \right].$$
 (4)

One method of handling this problem is to use a biased estimator which in essence is the same as (3) except replace γ_{l} with $W_{l}\gamma_{l}$ as the coefficient of $a_1 (0 \le W_1 \le 1)$. Then the corresponding term in (4) is $W_{l}^{2} a_{jl}^{2} / \lambda_{l}$. The idea is to let W_{ℓ} be small if λ_{ℓ} is small. Some of the more commonly used biased estimators have "L of the following forms:

Estimator	WE				
Simple Ridge	$\lambda_{l}/(\lambda_{l}+k) k>0; l=1,2,,p$				
Generalized Ridge	$\lambda_{\ell}/(\lambda_{\ell}+k_{\ell}) k_{\ell}>0; \ell=1,2,,p$				
Principal Components	0 or 1; <i>t</i> =1,2,,p				
Shrunken	c 0 <c<1; f="1,2,,p</td"></c<1;>				

The bias of this estimator of β would then

$$\sum_{\ell=1}^{p} (1-W_{\ell}) \gamma_{\ell} \alpha_{\ell}.$$

In addition to selecting the W, so that the variances of the resulting estimators are not extremely large, we wish to choose the Wg so that the bias is not great. Since $\mathfrak{g}_{\mathfrak{l}}$ is known and

not null, in fact $\sum_{j=1}^{\infty} \alpha_{jk}^2 = 1$, information is j=1

needed on the γ_{g} in order to insure that these biases are not so large that the benefits of variances reduction are counteracted.

From (4) we see that we wish to choose W, small if λ_{f} is small but from (5) this would inflate the bias unless $\gamma_{\underline{\ell}}$ is also small. Keep in mind that when λ_{ℓ} is small γ_{ℓ} has a large variance, $\sigma^2/\lambda_{\underline{l}}$; hence $\gamma_{\underline{l}}$ is of limited value in determing whether γ_{e} is small. This will be illustrated in the following data.

The Mesquite Problem

This data was furnished by R.J. Freund of Texas A & M University. The purpose of the study was in part to determine an estimation equation for the total production of aerial photosynthetic biomass of mesquite from easily measured parameters of the plant. An identification of the variables is given in Table A.1 and the data from 20 trees in Table A.2.

The concept of volume suggests a multiplicative model; hence an analysis of the logarithms of the variables would be appropriate. A linear regression of the raw data yielded approximately the same coefficient of determination, R^2 , as a linear regression of the logarithms. For the purpose of this paper we will use the analysis of the raw data since the interpretation of the regression coefficients is less confusing.

The residuals of the least squares fit of the data indicate the second tree to be an out-

DISTRIBUTION STATEMENT A

approved for public release Distribution Unlimited

Reprinted from the 1977 Statistical Computing Section Proceedings of the American Statistical Association • 7

for peaced

release

+1m11m. publie

lier. Physically this tree also stands out from the others with $X_5 = 9$ primary stems. We discarded this tree as an outlier and continued our analysis using the remaining 19 data points. These had a coefficient of determination of $\mathbb{R}^2 = .868$.

Table A.3 contains the standardized X'X matrix. The latent roots of this matrix are $\lambda_1 = 3.33$, $\lambda_2 = .89$, $\lambda_3 = .57$, $\lambda_4 = .12$, $\lambda_5 = .09$, indicating two multicollinearities,

although they are not as severe as one often encounters. The variance inflation factors, diagonal of $(X'X)^{-1} = (5.90, 4.58, 6.02, 5.03, 1.37)$, also point out the existence of some multicollinearities and further indicate that plant density, X_c, is not involved in them.

The "total variance" of b, the sum of the variances of these five estimators, is

 σ^2 [the sum of the diagonals of $(X'X)^{-1}$]=22.90 σ^2 .

σ

$$\sum_{\ell=1}^{2} \lambda_{\ell}^{-1}$$

From this second form we note that if $W_4 = W_5 = 0$, the resulting biased estimator of $\underline{\beta}$ has total variance

$$\sigma^2 \sum_{t=1}^{3} \lambda_t^{-1} = 3.46 \sigma^2.$$

Thus there is room for a considerable reduction in variance using an estimator of the form

$$\underline{B} = Y_{1}\underline{a}_{1} + Y_{2}\underline{a}_{2} + Y_{3}\underline{a}_{3} + W_{4}Y_{4}\underline{a}_{4} + W_{5}Y_{5}\underline{a}_{5}.$$
 (6)

With this in mind, the question then is how much bias would be introduced. The latent vectors $\underline{\alpha}_4$ and $\underline{\alpha}_5$ are given in Table A.4. Note that the fifth element in $\underline{\alpha}_4$ and $\underline{\alpha}_5$ are both relatively small; thus the introduction of small W₄ and W₅ will have little effect upon $\hat{\beta}_5$. Also the variance inflation factor of 1.37 indicates that there is little room for improvement on b_5 .

3. An Investigation of the Bias

The bias is affected by both γ_4 and γ_5 . We have already noted that $\gamma_1 = \underline{Y} \cdot \underline{z}_1 / \underline{z}_2 \cdot \underline{z}_1$ is an unbiased estimator of γ_1 with variance σ / λ_1 . Since γ_4 and γ_5 are both quite small, γ_4 and γ_5 both have exceedingly large variances. What this means can be shown graphically.

Due to the orthogonality of the $\underline{z}_{\underline{k}}$, and since $\underline{z}_{\underline{k}}'$ $\underline{1} = 0$,

 $\underline{\mathtt{Y}}'\underline{\mathtt{z}}_{\underline{k}}=(\underline{\mathtt{Y}}-\underline{\mathtt{J}}\overline{\mathtt{Y}}-\hat{\mathtt{Y}}_{1}\underline{\mathtt{z}}_{1},\ldots,\hat{\mathtt{Y}}_{\underline{k}-1}\underline{\mathtt{z}}_{\underline{k}-1}-\hat{\mathtt{Y}}_{\underline{k}+1}\underline{\mathtt{z}}_{\underline{k}+1}-\ldots,\hat{\mathtt{Y}}_{\underline{p}}\underline{\mathtt{z}}_{\underline{p}})'\underline{\mathtt{z}}_{\underline{k}},$

The long expression in parentheses on the right is \underline{Y} adjusted for the intercept and all the z's except z_g . Call this partial residual vector

 $\underline{r}_{\underline{\ell}}$. Then $\gamma_{\underline{\ell}} = \underline{r}_{\underline{\ell}} \underline{z}_{\underline{\ell}} / \underline{z}_{\underline{\ell}} \underline{z}_{\underline{\ell}}$ is the least squares slope of $r_{\underline{\ell}}$ versus $z_{\underline{\ell}}$, a slope that is easily

visualized through a plot of r_{ℓ} vs. z_{ℓ} . Figures Al through A5 show this for the five z's of this example.

FIGURE A1. PARTIAL RESIDUAL PLOT OF FIRST PRINCIPAL COMPONENT



FIGURE A2. PARTIAL RESIDUAL PLOT. OF SECOND PRINCIPAL COMPONENT



FIGURE A3. PARTIAL RESIDUAL PLOT OF THIRD PRINCIPAL COMPONENT



FIGURE A4. PARTIAL RESIDUAL PLOT OF FOURTH PRINCIPAL COMPONENT



For a moment ignore the source of this data and consider that each of these five plots was brought to you separately and you were asked to find a predictor for each. We believe that for Figures A4 and A5 you would quite politely say that there is not enough information to establish a relationship between the two variables. There may or may not be a relationship but data is needed outside the range of \pm .2. For example



FIGURE A5. PARTIAL RESIDUAL PLOT

OF FIFTH PRINCIPAL COMPONENT

look at Figures A6 and A7. (Figure A6 is the composite of Figures A1 and A5. Figure A7 is the composite of the reflection of Figure A1 and Figure A9. Both of these are compatable with Figure A5 and either one <u>could</u> be the result of data for z_5 outside of \pm .2. The crux of the matter is that the available data gives us little or no information on γ_4 and γ_5 .

FIG. A6. PARTIAL RESIDUAL PLOT OF FIFTH PRINCIPAL COMPONENT AUGMENTED WITH POINTS FROM FIGURE A1



9

FIG. A7. PARTIAL RESIDUAL PLOT OF FIFTH PRINCIPAL COMPONENT AUGMENTED WITH REFLECTED

POINTS PROM FIGURE AL



There is, however, one other available source of information: the problem itself and the experimenter knowledgeable in the area. Because of the simplicity of the measurements in this problem we can perhaps wear his hat, keeping in mind that the variable for prediction is the amount of leaves.

Looking at the nature of X_1 through X_4 we see that all four are size measurements and increasing any one should indicate more leaves. Thus β_1 through β_4 should all be positive. Further, the change of scale using the standardized X's suggest that β_1 and β_2 may be of the

same magnitude, as should perhaps β_3 and β_4 .

Now let us look at \underline{a}_4 and \underline{a}_5 . Multiplying both sides of (2) on the left by \underline{a}_1^* yields

$$\gamma_{g} = \underline{\alpha}_{g} p$$
,
so that $\gamma_{4} = .62\beta_{1} - .47\beta_{2} - .52\beta_{3} + .33\beta_{4} + .03\beta_{5}$

$$5 = -.44\beta_1 + .42\beta_2 - .54\beta_3 + .55\beta_4 + .17\beta_5$$

From what this problem suggests about these β 's we would not expect either of these γ 's to be large; i.e. note that the signs on β_1 and β_2 (and β_3 and β_4) are opposite, yet we expect them² to have the same sign and roughly equal magnitude-their effects should therefore cancel in γ_1 and γ_5 . Thus to estimate β using (6) with small W₄ and W₅ would greatly reduce the variances with the introduction of little bias.

. The Pitprop Data

Jeffers [1967] presented data on the maximum compressive strength of timbers used in mines. A description of the variables is given in Table B.1 and the standardized X'X matrix in Table B.2. This X'X matrix has three small latent roots

$$\lambda_{13} = .0387, \lambda_{12} = .0415, \lambda_{11} = .0506$$

Their corresponding latent vectors are given in Table B.3

For this example wearing the hat of the experimenter is more difficult so let us concentrate on X_1 and X_2 , the top diameter of the prop and the length of the prop. One would expect the greater the diameter, the greater the strength; hence, β_1 should be positive. With the same line of thinking, the longer the prop the less its supporting strength; hence β_2

Now look at γ_{12} and γ_{13} :

$$\gamma_{12} = -...39 \beta_1 + ...41 \beta_2 + ...$$

$$Y_{13} = -.57 \beta_1 + .58 \beta_2 + ...$$

From the contribution of β_1 and β_2 one would expect γ_{12} and γ_{13} not to be small. Thus small values of W_{12} and W_{13} in $\hat{\beta}$ could lead to considerable bias in the estimator of $\hat{\beta}$.

5. The Effect of W, on the Prediction of Y

Up to this time we have been concerned with the estimation of β . Now let us consider the use of this estimate to predict Y at a point X. (Note that the X's in this vector are scaled in the same manner as those in our X matrix.) Writing β in the form

$$\hat{\underline{\beta}} = \underline{W}_1 \hat{Y}_1 \underline{\underline{\alpha}}_1 + \underline{W}_2 \hat{Y}_2 \underline{\underline{\alpha}}_2 + \dots + \underline{W}_p \hat{Y}_p \underline{\underline{\alpha}}_p$$

the predictor will be

 $\mathbf{Y} = \mathbf{\overline{Y}} + \mathbf{W}_1 \mathbf{\gamma}_1 \mathbf{\underline{X}'} \mathbf{\underline{\alpha}}_1 + \mathbf{W}_2 \mathbf{\gamma}_2 \mathbf{\underline{X}'} \mathbf{\underline{\alpha}}_2 + \dots + \mathbf{W}_p \mathbf{\gamma}_p \mathbf{\underline{X}'} \mathbf{\underline{\alpha}}_p.$ (7) The bias in the predictor is

$$(1-w_1)\gamma_1\underline{x}'\underline{q}^+(1-w_2)\gamma_2\underline{x}'\underline{a}_2^+\dots^+(1-w_p)\gamma_p\underline{x}'\underline{a}_p$$
.

The bias will be small if for each small W_g either γ_g or $\underline{X}' \underline{\alpha}_g$ is small. Let us say that λ_p was small, indicating a multicollinearity with coefficients identified in $\underline{\alpha}_p$. W_p is then chosen to be small in order to reduce the variance of $\hat{\beta}$ but, unfortunately, suppose γ_p is not near zero. The pth term in the bias will still contribute little if $\underline{X}' \underline{\alpha}_p$ is near zero, conteracting the nonzero γ_p . Thus, if the point of prediction

satisfies this multicollinearity of the original data, a small W will induce little bias in the predictor. P

To summarize this last paragraph, consider that the $W_{\underline{k}}$ are chosen small only when $\lambda_{\underline{k}}$ is

small. Then, regardless of the magnitude of the $Y_{\underline{x}}$ (within reason), the \hat{Y} of (7) will have little

bias if the point of prediction, X, satisfies the multicollinearities of the original data. For example a prediction equation from the pitprop data using small W_{12} and W_{13} could be quite satisfactory for props with multicollinearities designated by $\underline{\alpha}_{12}$ and $\underline{\alpha}_{13}$. However interpretation of the $\hat{\beta}_j$ in this equation could be misleading. On the other hand interpretation of the β_i from the mesquite data using small W_4 and W_5 should be informative.

6. Conclusion

In biased linear regression techniques, small W_{l} are desired in order to reduce the variance of β in the presence of multicollinearities. This will introduce considerable bias in β if the corresponding γ_{l} are large. The basic point of this paper is that when λ_{l} is small the data itself generally gives little information on γ_{l} . This is well illustrated by the graphs of the mesquite data. Some information may be available, however, if γ_{l} is expressed as a linear function of the colorests of β and the metaget

of the elements of β and the nature of the specific problem is carefully analyzed.

We feel that in practice a complete investigation of the properties of a set of regression data, as suggested above, is sometimes hampered by the lack of computer programs. Robert Pierce, while at SMU, has written a program, REGRESS, which will simultaneously do a least squares, ridge, latent root and principal component analysis as well as furnish a shrunken estimate. The latent roots and vectors of X'X are a portion of the printout. A copy of this program is available upon request.

ACKNOWLEDGEMENT

This research was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-75-2871. The United States Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

- Hocking, R.R., Speed, F.J. and Lynn, M.J. (1976), "A class of biased estimators in linear regression," <u>Technometrics</u>, <u>18</u>, 425-437.
- Jeffers, J.N.R. (1967), "Two case studies in the application of principal component analysis," <u>Applied Statistics</u>, <u>16</u>, 225-236.
- Mason, R.L., Webster, J.T. and Gunst, R.F. (1975), "Sources of multicollinearity in regression analysis," <u>Communications</u> in Statistics, 4, 277-292.

TABLE A.1

ESOUITE	DATA:	RESPONSE	AND	PREDICTOR	VARIABLES
---------	-------	----------	-----	-----------	-----------

Variable	Description
Y (LEAFWT)	Total Weight (GRAMS) of Photosyn- thetic Material
X1 (DIAM1)	Canopy Diameter (METERS) Measured Along the Longest Axis of the Tree Parallel to the Ground
X ₂ (DIAM2)	Canopy Diameter (METERS) Measured Along the Shortest Axis of the Tree Parallel to the Ground
X ₃ (TOTHT)	Total Height (METERS) of the Mesquite Bush
X4 (CANHT)	Canopy Height (METERS) of the Mesquite Bush

X₅(DENS) Plant Unit Density (NUMBER OF PRIMARY STEMS/PLANT UNIT)

TABLE A.2

MESQUITE DATA

Bush						
Number	DIAMI	DIAM2	TOTHT	CANHT	DENS	LEAFWT
1	2.50	2.30	1.70	1.40	5	723.0
2	5.20	4 00	3.00	2 50		4052 0
-	2 00	1.60	1 70	1 40	,	345 0
-	2.00	1.60	1.70	1.40	+	345.0
4	1.60	1.60	1.60	1.30	1	330.9
5	1.40	1.00	1.40	1.10	1	163.5
6	3.20	1.90	1.90	1.50	3	1160.0
7	1.90	1.80	1.10	.80	1	386.6
8	2.40	2.40	1.60	1.10	3	693.5
9	2.50	1.80	2.00	1.30	7	674.4
10	2.10	1.50	1.25	.85	1	217.5
11	2.40	2.20	2.00	1.50	2	771.3
12	2.40	1.70	1.30	1.20	2	341.7
13	1.90	1.20	1.45	1.15	2	125.7
14	2.70	2.50	2.20	1.50	3	462.5
15	1.30	1.10	.70	.70	1	64.5
16	2.90	2.70	1.90	1.90	1	850.6
17	2.10	1.00	1.80	1.50	2	226.0
18	4.10	3.80	2.00	1.50	2	1745.1
19	2.80	2.50	2.20	1.50	1	908.0
20	1.27	1.00	.92	.62	1	213.5

TABLE A. 3

CORRELATION	MATR	IX OF	PRED	ICTOR	AND	RESPONSE
VARIABI	LES,	MESQU	ITE D	ATA (1	n =	19)

	x,	×2	x ₃	×4	x ₅
×1	1.00	. 88	.73	.68	. 32
×2		1.00	.64	. 58	.20
×3			1.00	. 88	. 39
×4				1.00	.22
x,5					1.00

TABLE A.4

SMALLEST TWO LATENT ROOTS AND CORRESPONDING LATENT VECTORS OF X'X, MESQUITE DATA (n = 19)

$\lambda_5 = .0904$	×4 =	.1155
---------------------	------	-------

VARIABLE	<u>a</u> 5	º4	VIF
×1	444	.624	5.9
×2	. 420	472	4.6
×3	544	522	6.0
×4	.551	.339	5.0
x ₅	.165	. 026	1.4

TABLE B.1

PITPROP DATA:	RESPONSE AND PREDICTOR VARIABLES
VARIABLE	DESCRIPTION
¥	Maximum Compressive Strength of Prop
x ₁	Top Diameter of the Prop (INCHES)
x2	Length of Prop (INCHES)
x ₃	Moisture Content of Prop (% OF DRY WEIGHT)
×4	Specific Gravity of the Timber (AT TIME OF TEST)
×5	Oven-dry Specific Gravity of the Timber
× ₆	Number of Annual Rings at Top of Prop
×7	Number of Annual Rings at Base of Prop
×8 .	Maximum Bow (INCHES)
×9	Distance of the Point of Maximum Bow from Top of Prop (INCHES)
×10	Number of Knot Whorls
×11	Length of Clear Prop from Top of Prop (INCHES)
x ₁₂	Average Number of Knots per Whorl
X13	Average Diameter of Knots (INCHES)

TABLE	B.2	
-------	-----	--

• -

	COR	RELATION	MATRIX	OF PREDI	CTOR AND	RESPONS	E VARIAB	LES, PIT	PROP DAT	A		
×1	×2	×3	×4	×5	×6	×7	x ₈	x ₉	×10	×11	×12	×13
x, 1.00	. 95	. 36	. 34	13	.31	. 50	. 42	. 59	. 55	.08	02	.13
x ₂	1.00	. 30	. 28	12	. 29	. 50	. 42	.65	.57	.08	04	.14
x ₃		1.00	.88	15	.15	03	05	.13	08	.16	.22	.13
×4			1.00	.22	. 38	.17	06	.14	01	.10	.17	.02
x ₅				1.00	. 36	.30	.00	04	.04	09	15	21
× ₆					1.00	.81	.09	.21	.27	04	.02	33
×7						1.00	. 37	.47	.68	11	23	42
×e							1.00	.48	. 56	.06	36	20
×9								1.00	.53	.09	13	08
×10									1.00	32	37	29
×_11										1.00	.03	.01
×12											1.00	.18
×13												1.00
¥42	34	73	54	.25	.12	.11	25	24	10	06	12	15

TABLE B.3

SMALLEST THREE LATENT ROOTS AND CORRESPONDING LATENT VECTORS OF X'X, PITPROP DATA

	$\lambda_{13} = .0387$	$\lambda_{12} = .0415$	$\lambda_{11} = .0506$	
VARIABLE	<u>a</u> 13	<u>a</u> 12	<u><u>a</u>11</u>	VIF
x,	572	392	005	13
x,	. 582	.411	054	14
x ₃	.408	527	.117	12
X	383	. 585	017	12
x	.118	202	.005	3
×6	.057	080	537	7
x,	.002	.036	.764	12
X	.018	.053	.026	2
X	058	054	051	2
xin	.004	060	318	5
x,,	007	005	048	2
x12	.004	002	.047	1
×13	009	013	.045	2



SECURITY CLASSIFICATION OF THIS PAGE (Whan Data Entered) **READ INSTRUCTIONS** REPORT DOCUMENTATION PAGE BEFORE COMPLETING FORM 2. JOVT ACCESSION NO. 3. PECIPIENT'S CATALOG NUMBER 78-1732 AFOSR TYPE OF REPORT & PERIOD COVERED ILKING DATA THROUGH BIASED REGRESSION TECHNIQUES Interim PERFORMING ORG. REPORT NUMBER 7. AUTHOR(s) 8. CONTRACT OR GRANT NUMBER(S) R. F. Gunst J. T// Webster AFOSR-75-2871 10. PROGRAM ELEMENT, PROJECT, TASK PERFORMING ORGANIZATION NAME AND ADDRESS Southern Methodist University Department of Statistics 611 02 Dallas, Texas 75275 1. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research /NM 1977 Bldg. 410, Bolling AFB 7 Washington, D.C. 20332 4. MONITORING AGENCY NAME & ADDRESS(II dillerent from Cont olling Office) 15. SECURITY CLASS. (of this report) UNCLASSIFIED 154. DECLASSIFICATION/DOWNGRADING SCHEDULE 16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distribution Unlimited 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) B 18. SUPPLEMENTARY NOTES 1977 Statistical Computing Sec Proceedings Am Statistical Assoc p7-13 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Multicollinearity **Regression Analysis** Principal Components 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The recent interest in biased estimation procedures in multiple linear regression arises from the large variances of the least squares estimators (unbiased) of the regression coefficients when multicollinearities are present. The biased estimation procedures greatly reduce this variance at the cost of some bias. It is the purpose of this paper to look at this bias with reference to the nature of the specific problem being investigated. EDITION OF I NOV 65 IS UBSOLETE UNCLASSIFICATION DE THIS PAGE (When Dete Entere DD 1 JAN 73 1473