DOCUMENT RESUME

ED 073 164 TM 002 413

AUTHOR	Beaton, Albert E.; And Others
TITLE	The Acceptability of Regression Solutions: Another Look at Computational Accuracy.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RB-72-44
PUB DATE	Sep 72
NOTE	43p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Computer Programs; *Data Analysis; *Mathematical Models; *Measurement Techniques; *Multiple Regression Analysis; Technical Reports
IDENTIFIERS	*Longley (J W)

ABSTRACT

Longley proposed a set of data for use in testing regression programs. This paper shows that the numerically accurate solution is likely to be an unreasonable estimate of the regression coefficients for this problem. This is true because the accuracy of the data and appropriateness of the model may affect the solution more than the computational method. An easily computed index is derived that can be used to indicate such computational instability. The basic conclusion is that a concern about highly accurate computational methods must be tempered with a concern for whether the data are accurate enough to make the results of such computation meaningful. (Author)

RESEARCH BULLETIN

RB-72-44

THE ACCEPTABILITY OF REGRESSION SOLUTIONS: ANOTHER LOOK AT COMPUTATIONAL ACCURACY

Albert E. Beaton
Donald B. Rubin
and
John L. Barone

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the authors. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
September 1972

THE ACCEPTABILITY OF REGRESSION SOLUTIONS:
ANOTHER LOOK AT COMPUTATIONAL ACCURACY

Abstract

Longley proposed a set of data for use in testing regression programs. This paper shows that the numerically accurate solution is likely to be an unreasonable estimate of the regression coefficients for this problem. This is true because the accuracy of the data and appropriateness of the model may affect the solution more than the computational method. An easily computed index is derived that can be used to indicate such computational instability. The basic conclusion is that a concern about highly accurate computational methods must be tempered with a concern for whether the data are accurate enough to make the results of such computation meaningful.

THE ACCEPTABILITY OF REGRESSION SOLUTIONS: ANOTHER LOOK AT COMPUTATIONAL ACCURACY

Albert E. Beaton, Donald B. Rubin and John L. Barone

Educational Testing Service

1. Introduction

Multiple regression is an extremely popular and powerful method of data analysis. Recently, there has been increasing concern about the numerical accuracy of common computer programs now available and in use. The paper of Longley (1967) is perhaps the most startling paper on this subject in the recent statistical literature. Longley took what seems to be a reasonable set of economic data and performed a six-variable multiple regression analysis using several different programs on several different computers. He found that different regression programs resulted in very different solutions including differences in sign and first significant digit. This finding seems to indicate that one should be very careful about the program and machine he uses.

We feel that the computer program is often not the most important factor in computing a regression analysis, and that the best thing a program can do for some problems is to refuse to complete the calculations. Numerical experiments in this paper will show that the computationally accurate solution to this regression problem--even when computed using 40 decimal digits of accuracy--may be a very poor estimate of regression coefficients in the following sense: small errors beyond the last decimal place in the data can result in solutions more different than those computed by Longley with his less preferred programs. The computationally accurate solution is shown to be nowhere near the center of the distribution of a large number of presumably equally possible solutions.

The solution to a regression problem is affected by the data, the statistical model, and the program. This paper will explore the actual accuracy of the data used and show in what sense they are inadequate for the solution of this model. A reduced statistical model will be fit under which the results are not seriously affected by small errors in the data or the particular programming algorithm. Various algorithms are shown to be sufficiently accurate for most practical purposes if a regression model has a reasonably stable solution.

We then show how knowledge of the error variance in the independent variables can be used to compute a simple "perturbation index" which indicates the stability of the computed solution over the range of possible true data sets.

2. The Longley Problem

At first glance the Longley problem seems very much like a typical multiple regression analysis of a time series in which one "dependent" variable Y is regressed on six "independent" variables. The variables are

- Y Total Derived Employment (in thousands)
- X1 Gross National Product Implicit Price Deflator (in tenths)
- X2 Gross National Product (GNP) (in millions)
- X3 Unemployment (in thousands)
- X4 Size of Armed Forces (in thousands)
- X5 Noninstitutional Population 14 Years of Age and Over (in thousands)
- X6 Year

Longley also presented seven components of Total Derived Employment which are discussed in Section 4. Longley fit the following regression model

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \epsilon_i$$

$$i = 1,2,\ldots,16 \quad .$$

The Longley analysis was computed on the data available for the 16 years from 1947 to 1962. The basic data are shown in Table 1 along with the means,

----

Insert Table 1 about here

----

standard deviations, ratios of the means to standard deviations, and intercorrelations.

The means and standard deviations do not seem to indicate any particular difficulty for analysis. A careful research person might try to improve computational accuracy through standardizing each variable by subtracting its mean and dividing by its standard deviation, thus converting the raw data matrix to a matrix of standard scores (Golub, 1969). The ratio of the mean to the standard deviation is an indicator of the sort of computational problem discussed by Neely (1966); although the ratios here are not zero, as would be the case with standard scores, the ratios, except for X6, do not seem unduly large. The high intercorrelations among the independent variables portend a conditioning problem[1] since no fewer than five of the 15 unique off-diagonal correlations are greater than .99 and a sixth is nearly .98. We feel that most statisticians would advise a client not to fit a model with such high intercorrelations. Nevertheless, Table 2 gives the usual regression solution for this problem produced by the DLSSQ program discussed in detail later.

Insert Table 2 about here

Longley fit the model with these data by a high-precision desk calculator method and by a number of different regression programs on different machines. The results of some of the programs used by Longley and two programs used in this paper are shown in Table 3.

Insert Table 3 about here

The first two lines are the solution of the regression model by the high precision desk calculator method and by an IBM 1401 program that performs calculations using 40 decimal-digit accuracy. We have inserted the calculated solution of two programs (DORTHO and DLSSQ) which were written for the experiments performed in this pape . All four solutions agree to at least seven decimal places and thus may be considered identical for most practical purposes. This vector of regression coefficients will be referred to as the "unperturbed" solution without regard to the method of calculation.

T remaining part of this table has the solutions computed by Longley with eight programs on four different machines. The variations are striking. Some programs generate regression coefficients different in sign and in most significant digit from the unperturbed solution. The results of the ORTHO program are closest to the unperturbed solution. The ORTHO algorithm has been published by Walsh (1962) and used in the OMNITAB program (Hilsenrath et al., 1966) of the National Bureau of Standards.

3. Is the Unperturbed Solution a Good Solution for This Sample?

Although calculation to very high precision is satisfying, we wish to explore the unperturbed solution further. We do not question that the unperturbed solution is a possible solution to this regression problem, but there are a large number of other solutions, each in a sense as likely to be the correct solution as any other or as the unperturbed solution.

Taking an extremely conservative position, we cannot avoid the likelihood that the 1947 value of variable X1, GNP Implicit Price Deflator, was not exactly 83.0 but some number between 82.5 and 83.499..; that the Gross National Product is not precisely 234,289,000,000, but some number between 234,288,500,000.00 and 234,289,499,999.99. All of the variables, X1 through X5, are subject to this type of deviation. Even X6, Year, is not an exact variable, although it may have a smaller error than the other independent variables. For our purposes here, we will ignore errors in the dependent variable, Total Derived Employment, even though that measure is clearly not exact either. We presume, then, that the data in Table 1 are absolutely accurate as far as they go, but do not go as far as possible.

The error introduced by such rounding would seem to be trivial since the data are presented with three to six digits of accuracy. To investigate this assumption, we have performed a numerical experiment by taking a random sample of possible exact values to see if these perturbed data sets would result in a solution similar to the unperturbed solution. A sample of 1,000 plausible sets of six independent variables was generated by adding a rectangularly distributed random number[2] between -.5 and +.499... in the digit after the last published digit. All data sets would be exactly the same as the published set if rounded to the published number of digits, and in this sense each of these data sets is as likely to be the exact data as any other or as the published data.

We next computed a thousand regression analyses using these perturbed values and the DORTHO subroutine.[3] The results are shown in Table 4. The

----

Insert Table 4 about here

----

results of this experiment are very striking indeed. Looking first at the highest and lowest values of the regression coefficients, these miniscule variations in the values of the independent variables have resulted in changes in the computed regression coefficients from -232.2792 to 237.0467 for B1, and equivalently elsewhere. There are differences in sign and magnitude for all regression coefficients.

One might hope that nearly all possible solutions would agree with the unperturbed solution to at least one significant digit. Not so. For all variables except X4, the unperturbed regression coefficients agreed to a single significant digit with a perturbed solution in about 2% of the cases; X4 agreed to at least one place in about 95% of the cases. Not one of the 1,000 sets of estimated regression coefficients agreed with the unperturbed to one decimal place in all seven coefficients.

Perhaps the unperturbed solution is at least near the center of the thousand perturbed solutions. But no. The mean and median of the thousand solutions are shown in Table 4. For B_1 and B_2, the mean and median differ in sign from the unperturbed solution; only for B_4 do the unperturbed solution and the mean agree to one significant digit. Assuming that the unperturbed solution is the true mean of all possible samples and that the sample means are normally distributed about the unperturbed solution, then we can apply the standard student t -test to the hypotheses that the unperturbed values are the true population means. The t -statistics are also shown. The hypothesis that the unperturbed solution is the average of all solutions for these equally likely data sets is entirely implausible with the absolute values of the t -statistics ranging from about 22 to 116.

If the number of observations (N = 16) were large, we would expect the average of these thousand solutions to be as indicated in the row labeled P-lim in Table 4a. The reason will be discussed later. For now, notice that the average is much closer to P-lim than to the unperturbed and that P-lim is not at all close to the unperturbed.

These results are shown graphically in Figure 1 which depicts histograms

---

Insert Figure 1 about here

---

of each regression coefficient including the intercept, B_0 . Each histogram is centered at the mean of the perturbed solutions and includes the range from three standard deviations below to three standard deviations above the mean. The vertical line with an encircled U represents the unperturbed solution; the abscissa has been extended to include this point wherever necessary. The line with an encircled P represents the P-lim value.

The effect of these perturbations on the squared multiple correlation is shown in Table 4b. All R^2 's are high, but the unperturbed R^2 is not near the center of the distribution. It is in fact 65 standard errors away from the mean R^2 . The estimated values of Y for the unperturbed solution and the mean of a thousand perturbed solutions are shown in Table

4c. In all cases the predictions agree to several places, but it is not true that the unperturbed estimates are near the average of the perturbed estimates. The absolute values of the t -statistics for the differences between the unperturbed and average perturbed solution range from 3.5 to 82.

We conclude, therefore, that it is extremely unlikely that the unperturbed solution is the "correct" solution of this problem. Assuming uniform rounding error in the independent variables, it is highly likely that two of the unperturbed coefficients are incorrect in sign and all but one are not correct to one significant digit. The unperturbed multiple correlation and estimated values, although close enough to the average perturbed values for most practical purposes, are nevertheless significantly different.

The unperturbed solution is, therefore, in this sense totally unsatisfactory. Since regression analysis is a combination of model, data and algorithm, we will now explore these individually to see what effect each has had in the analysis.

4. The Data

Since a regression analysis can be sensitive to small perturbations of the data, we think it worthwhile to make some comments about the actual accuracy of the Longley data. In fact, the error is often many orders of magnitude larger than the error we have introduced. In general, data are subject to errors in sampling, measurement, calculation, copying, and to a host of other factors and inconsistencies. A thorough analysis of the error in these data is far beyond the scope of this paper, but we will point out a few salient facts about each variable. This discussion may seem to be focusing on trivia, but we have seen that smaller errors may have enormous effects.

The Longley data come from four government reports:

1. The United States Department of Labor "Manpower Report of the President and a Report on Manpower Requirements, Resources, Utilization, and Training," March 1963, Table A-4, p. 141. (USDLA)
2. The United States Department of Labor "Employment and Earnings," Vol. 10, #3, September 1963, Tables A-1 and B-1. (USDLB)
3. The United States Department of Commerce, Office of Business Economics "Survey of Current Business," July 1963, p. 12. (USDC)
4. Council of Economic Advisors, <u>Economic Report of the President</u>, January 1964, Table C-6, p. 214. (CEA)

USDLA is drawn from the Department of Labor Monthly Labor Force Survey. Presumably, this is the same monthly survey described in USDLB, which is described next.

The USDLB data are drawn from several sources; the prime resources are:

1. A monthly labor force survey sampling of 35,000 households in 357 areas in the country.
2. Payroll employment statistics supplied monthly by a sample of industrial, commercial, and government establishments employing collectively about 25 million workers.
3. Unemployment contribution reports filed by employers subject to state unemployment insurance laws. These are considered bench-mark data and cover about 75% of the

total nonfarm employees. Other bench-mark data are collected from various government agencies.

4. Labor turnover statistics, supplied by a sample of manufacturing, mining, and communication industries.

The USDC bench-mark data come primarily from census data. Also, the Bureau of Census provides an annual sample survey of manufacturers which are used as a basis for estimating a number of GNP statistics. Comprehensive annual reports of government agencies and annual data from private sources are also used.

CEA was used for the estimated implicit price deflators which are based primarily on data from the Consumer Price Index of the Bureau of Labor Statistics. These data are supplemented by information drawn from the Agricultural Marketing Service, Department of Commerce, Interstate Commerce Commission, and various other government statistics.

Clearly, data from such sources may be subject to errors of many different types. Let us look at some characteristics of individual variables.

Y. Total Derived Employment (in thousands). The dependent variable is the sum of the estimates of

Y1. Agricultural Employment (from USDLB)

Y2. Self-Employment (from USDLA)

Y3. Unpaid Family Workers (from USDLA)

Y4. Domestic Workers (from USDLA)

Y5. Nonagricultural Private Workers (this was computed by subtracting the total government workers from the total nonagricultural) (from USDLB)

Y6. Federal Workers (from USDLB)

Y7. State and Local Government Workers (from USDLB)

Both data sources warn that there are several periods of noncomparability over time in these components because of the introduction of data from the 1950 census and the admission to statehood of Alaska and Hawaii. Total employment figures were increased by 350,000 and 300,000 respectively. There is also a systematic difference in the interpretation of figures since variables Y2, Y3, and Y4 have not been adjusted for a change in the definitions of employment and unemployment adopted in 1957, whereas Y1 and (we think) Y5, Y6, and Y7 have been adjusted. The change in definition involved a decrease of 250,000 in total number of employed. The overall effect of these errors is in the hundreds of thousands, perhaps millions.

X1. Gross National Product Implicit Price Deflator (in tenths). The implicit price deflator index is the ratio of GNP in current prices to GNP in constant prices. We have not been able to estimate its error. However, the publication "U.S. Income and Output, Supplement to Survey of Current Business" (1958, p. 52) notes:

> ...we called attention to the shortcomings of price inflation. These stem from lack of price information directly applicable to many components of the current-dollar product flow; from the fact that, generally speaking, available price information cannot take adequate account of premiums, discounts, and bargain sales; and from the even more basic problems encountered in pricing items subject to significant quality change, or whose physical units are not nearly definable for other reasons.

X2. Gross National Product (in millions). The Gross National Product is also subject to many kinds of error. Actually, GNP can be computed from either the nation's input or output, and both calculations should result in identical figures. In practice, the estimates do not, and over the years from 1947 to 1962 the discrepancy between the two measures ranges from +3.5 to -3.0 billion dollars.

X3. Unemployment, X4. Size of Armed Forces, X5. Noninstitutional Population (in thousands). These variables come from the same table (USDLB) and thus have the same properties. The data from 1947 to 1950 have been adjusted to reflect changes in the definition of employment and unemployment. These variables are not comparable over time because of the introduction of data from the 1950 census and the introduction of Alaska and Hawaii.

X6. Year. The year is perhaps the most difficult of these variables to understand. It is a catchall variable, and it is difficult to describe just what it purports to measure. Gross National Product is the sum of a number of things over a year, whereas the population figures (X3, X4, and X5) are values that fluctuate during the year and perhaps can be considered average values. To what does year refer? Is it a calendar year or fiscal year? Are the different employment figures collected at the same point in time? We do not know.

All in all, one has the feeling that these data are good to a little over two significant digits except for X6. One also has the feeling that the various government agencies have gone to great pains to make the data as accurate as possible. Of course, there are other types of error that we cannot estimate here. Although these data may be sufficiently accurate for some purposes, they clearly are not sufficiently accurate to "support" the unperturbed solution.

5. The Model

The selection of a mathematical model is a very important part of a regression analysis. Ideally, a research person specifies a model, then collects data to estimate parameters, choosing the independent variables carefully to minimize the interindependent variable correlations and avoiding the problem of multicollinearity. But persons working with observational data often cannot avoid high correlations. Thus, if the research person is really interested in performing linear regression using all his variables, the results of Section 3 indicate that he may have to collect his data with extreme precision to generate reasonable estimates of regression coefficients. Such precise measurement is seldom possible.

On the other hand, many studies of observational data do not have a strong causal basis which dictates a model; in fact, many persons use the data and a stepwise regression program to construct a model. The question to which we address ourselves here is: Would a different model have been as sensitive to such minor errors in the data?

As indicated above, we doubt that many statisticians would approve of a regression analysis including such highly correlated independent variables. We have therefore decided to try a different model

$$Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \epsilon$$

which deletes variables such that no X's are correlated higher than .95. The highest correlation among these variables is .62 between X_1 and X_3.

We first fit this regression model without perturbation using the DORTHO routine. The results are shown in Table 5 which also contains the results for the same thousand sets of perturbed data.[4]

Insert Table 5 about here

First, the perturbed results always agreed with the unperturbed in sign and in at least the first significant digit. The lowest and highest values are not far enough apart to change interpretation. The t -statistics indicate that it is not unreasonable to presume that the solutions from the DORTHO routine would be the average of all perturbed solutions. The P-lim row indicates that, in contrast to the six-variable problem, the average perturbed solution in large samples would be very close to the unperturbed solution.

We conclude, then, that this model is not affected much by minor perturbations of these data. In fact, as we shall see shortly, even very poor programs compute good solutions to this model.

6. The Program

The question of computational accuracy involves both algorithm and precision and both must be related to cost. Longley's experiment indicates that different programs do indeed yield different results, and thus one might assume that a research person is obliged to use the numerically best program for all problems. Our experiment shows that the effects of minor errors in the data are greater than differences due to program; in fact, many of the coefficients estimated from the poor programs are closer to the mean of the perturbed solutions than the unperturbed solution. But are we usually in trouble using less stable programs?

Many numerical analysts prefer the modified Gram-Schmidt method for regression analysis. The algorithm avoids computing a cross-products matrix, thus does not "square" its problems; that is, the condition number of the data matrix is the square root of the condition number of the raw cross-products matrix, thus almost certainly better conditioned. Most packaged regression programs do compute a cross-products matrix and solve the normal equations using a matrix inversion subroutine. All the programs in Table 3 that disagreed (and some of those that agreed) with the unperturbed solution tried to solve the normal equations.

Why do programmers fail to use the modified Gram-Schmidt algorithm? Basically, modified Gram-Schmidt is too expensive for a general program: it requires a pass over the data for each independent variable. The ORTHO routine is more efficient, requiring only two passes. If all data can be kept in computer memory, then the extra multiplications to calculate regression coefficients required by Gram-Schmidt can be justified, especially if residuals are to be calculated, but, if the data matrix exceeds computer storage, then the number of logical rewinds--whether on tape or disk--will ordinarily discourage programmers and users.

To avoid rewinds, one might use less preferred methods with double precision. The Longley paper seems to indicate that double precision does not help since the IBM double precision solution was only trivially better than single precision. However, this finding is not general but due instead to a subtle bug in the IBM program.[5]

To investigate the effects of algorithm and accuracy, we have programmed two subroutines, DLSSQ (Double precision Least Squares) and WRYLG (the Worst Routine You are Likely to Get) in addition to DORTHO. The DLSSQ is not a "good" program but is about what one might expect in packaged regression programs. The WRYLG is the single precision version of DLSSQ and is about as bad a program as one might come upon, excluding those programs with real bugs. Bugs in programs can usually be caught by methods such as those suggested by Longley (1967) and by Mullet and Murray (1971).

Both DLSSQ and WRYLG use fairly standard computing procedures. The cross-products are computed without subtracting means. The cross-products $\Sigma X_i X_j$ are centered by the notorious algorithm

$$\sum_{i=1}^{N} (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k) = \Sigma X_{ij} X_{ik} - \frac{\Sigma(X_{ij})(\Sigma X_{ik})}{N}$$

where N is the number of observations and j and k index any pair of variables. The inverse is computed by the Gauss-Jordan method. Pivoting is done in order without reordering by size. DLSSQ requires that a pivot be greater than .01 of the original variance; WRYLG requires only that a pivot be positive, no matter how small. Aside from this, the only difference between these routines is that DLSSQ is in double precision on the IBM 360.

We submitted the unperturbed Longley data to both of these subprograms. The DLSSQ routine agreed quite closely with the desk calculator solution as shown in Table 3. The WRYLG is not shown in Table 3 since it could not compute a solution because of a negative pivot. In fact, the single precision representation of the augmented cross-products matrix has a negative eigenvalue and determinant.

The 1,000 sets of perturbed data were submitted to both DLSSQ and WRYLG. The DLSSQ and DORTHO programs are compared in Table 6. There is an average

----

Insert Table 6 about here

----

of over 7.5 significant digits of agreement for all coefficients. In the worst case the two programs agree to 4.34 digits. Apparently, the perturbations affect both programs similarly.

A summary of the 1,000 solutions by WRYLG are shown in Table 7. WRYLG

Insert Table 7 about here

----

rejected the problem because of a nonpositive pivot in fully 65% of the samples. The analyses that went to completion tend to vary more than those from DORTHO. As with the DORTHO routine, the coefficient of X_4 agreed with the unperturbed solution to one significant digit in about 95% of the cases in which a solution was produced. We might view this program as poorer because of the increased variability in results; however, we might also consider WRYLG superior because it indicated that the problem has no solution in 65% of the samples.

The six-variable problem is one in which we know that small errors have major effects, but is the WRYLG adequate for the three-variable problem? DLSSQ and DORTHO agree to an average of over 13 digits for the 1,000 data sets where three variables are used. A comparison of the WRYLG and DORTHO is shown in the bottom of Table 5. WRYLG and DORTHO agree to an average of 3.5 to 4 significant digits,[6] even though the t-statistic indicates that the average difference is significantly different from zero. The means and standard deviations of the differences, and the largest and smallest differences, are also shown.

We summarize, then, that the DLSSQ routine agrees substantially with the DORTHO routine in every comparison. Single precision routines including WRYLG are more variable in the poorly conditioned six-variable problem. The WRYLG agrees with DORTHO to about 3.5 digits on the average in the three-variable problem.

7. A Perturbation Index

It clearly would be helpful if the standard output of regression programs included some indicator of the effect of perturbations on the calculated coefficients. A possible candidate is the list of standard errors or t -tests of the coefficients; these, however, do not indicate such instability but rather the instability of the coefficients over repeated sampling of new Y's for the same X's. Thus the gross instability of the solution might surprise some researchers.

The instability should not be surprising to numerical analysts. Wilkinson (1965, 1967) has developed a system of error analysis which places bounds on the effects of perturbation on calculated solutions due to the finite word length of computers. The Longley data are so ill-conditioned that the Wilkinson procedure does not give an error bound even with double precision on the IBM 360.[7] As Wilkinson says

> It would be more reasonable to ascribe the loss of accuracy in such a case to be apparent sensitivity of the equations rather than to speak of "severe accumulation of rounding errors."

Our motivation here is to indicate how much the solution can vary due to perturbations beyond the last supplied digit even assuming an infinite word length computer. To do this we first consider the large sample limit of the equally likely solutions. This limit, called P-lim in the previous tables, can roughly be thought of as the center of the distribution of the equally likely solutions assuming a large sample size. Since, in large samples, almost all perturbed solutions are extremely close to this central value, it is the summary of interest.

Consider the regression model

$$y = T\beta + \epsilon$$

where y is an $N \times 1$ vector of dependent variables, T is an $N \times m$ matrix of correct values of the m independent variables, β is an $m \times 1$ vector of parameters we wish to estimate, and ϵ is the $N \times 1$ catchall vector for the unpredictable portion of y. Letting X represent the actual observed data and E the difference between X and T, we have $T = X + E$. We assume[8] that the expectation of each component of E is zero and that the individual errors are independent with known variances d_i:

$$\mathcal{E}(E) = 0 \ , \ \frac{1}{N}\mathcal{E}(E'E) = D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_p \end{bmatrix} , \ \mathcal{E}(E'X) = 0 \ .$$

If we knew the true values of the independent variables the usual least squares estimator would be

$$\hat{\beta} = (T'T)^{-1}T'y = [(X + E)'(X + E)]^{-1}(X + E)'y$$

$$[X'X + E'X + X'E + E'E]^{-1}(X + E)'y \ .$$

As the sample size N gets larger $(N \to \infty)$ the values of quantities in the sample approach their population values; more formally, assuming $\frac{X'X}{N} = C_{xx}$, a fixed $m \times m$ matrix, and $\frac{X'y}{N} = C_{xy}$, a fixed m-vector for all N

$$\text{P-lim } \hat{\beta} = [C_{xx} + D]^{-1} C_{xy} \quad .$$

Now the usual least squares estimator $\hat{\beta}_x$ based on the observed values is

$$\hat{\beta}_x = (X'X)^{-1}X'y = C_{xx}^{-1} C_{xy} \quad .$$

The difference between $\hat{\beta}_x$ and P-lim $\hat{\beta}$ is an indicator of the effect of the error perturbation on the calculated solution. This difference is

$$\hat{\beta}_x - \text{P-lim } \hat{\beta} = C_{xx}^{-1} C_{xy} - (C_{xy} + D)^{-1} C_{xy}$$

$$= [C_{xx}^{-1} - (C_{xx} + D)^{-1}] C_{xy}$$

$$= [I - (I + C_{xx}^{-1} D)^{-1}] C_{xx}^{-1} C_{xy}$$

$$= [I - (I + C_{xx}^{-1} D)^{-1}] \hat{\beta}_x \quad .$$

If the matrix $C_{xx}^{-1} D = 0_{mxm}$, then $\hat{\beta}_x = \text{P-lim } \hat{\beta}$. If C_{xx}^{-1} is "small" compared to I, then $\hat{\beta}_x \simeq \text{P-lim } \hat{\beta}$. The phrase "small compared to I" really means that all of the eigenvalues of $C_{xx}^{-1} D$ are close to zero. The largest eigenvalue of a matrix is commonly used to measure its size and is called the "spectral norm." A simpler measure is the trace of $C_{xx}^{-1} D$, which equals the sum of the eigenvalues which are nonnegative because C_{xx}^{-1} is positive and D is nonnegative. If the trace of $C_{xx}^{-1} D$ is substantially less than 1, we can be confident that $\beta_x \simeq \text{P-lim } \hat{\beta}$ because the largest eigenvalue must be substantially less than unity. Equivalently, if the "perturbation index"

$$PI = tr(C_{xx}^{-1} D)$$

is substantially less than 1, we can be confident that the effect of perturbation error on the computer solution will be small, especially for large N .

Alternatively, the perturbation index can be written

$$PI = \sum_{j=1}^{m} C_{xx}^{jj} d_j = N \sum_{j=1}^{m} (X'X)^{jj} d_j$$

where C_{xx}^{jj} and $(X'X)^{jj}$ are the diagonal elements of C_{xx}^{-1} and $(X'X)^{-1}$, respectively. The diagonal elements of C_{xx}^{-1} can be written $C_{xx}^{jj} = \frac{1}{(1 - R_j^2)s_j^2}$, where s_j^2 is the marginal variance of X_j and R_j^2 is the squared multiple correlation between X_j and all the other X variables.[9] Since D is a diagonal matrix with elements d_j , the diagonal elements of $C_{xx}^{-1}D$ are $\frac{d_j/s_j^2}{(1 - R_j^2)}$, and hence the perturbation index can be written

$$PI = \sum_{j=1}^{m} \frac{d_j/s_j^2}{(1 - R_j^2)} \quad .$$

The numerator of each term of PI , d_j/s_j^2 , is the ratio of the rounding error variance to the marginal variance. The denominator of each term of PI , $1 - R_j^2$, has been standard output for some time in some regression routines as a measure of collinearity.[10] The calculation of the order of magnitude of PI is trivial given an estimate of the original rounding error variances and either the diagonal of the inverse matrix, or the marginal variances and collinearity indices.

Since in our experiment the random numbers were rectangular, the value of d_j was 1/12 for all variables except X_1 for which the value was 1/1200. To compute the P-lim solution for the Longley problem the values of Nd_j were added to the diagonal of X'X before using the DLSSQ regression subprogram. The results are shown in Tables 4a and 6. Although the P-lim solution is much closer to the center of the distribution than the unperturbed solution, it is still significantly different from the mean of the perturbed solutions in all coefficients; perhaps this is not surprising because $N = 16$ is rather small for our assumptions.

The values of $(X'X)^{jj}$ and d_j , and the collinearity indices for both the three- and six-variable problems are shown in Table 8. The perturbation indices are 2.86×10^{-5} and 2.9777 respectively.

We note in the three-variable problem that none of the individual elements of the perturbation index are larger than 10^{-4}, which is very small. In the unstable six-variable problem, the component associated with variable 6 is much greater than unity so that we would expect instability. We reran the problem excluding variable 6 with the results that the perturbed and the ordinary least squares solutions agreed to at least three decimal places and all the components of the perturbation index became small.

Insert Table 8 about here

----

8. Summary

The purpose of this paper has been to show that regression coefficients can fluctuate wildly as a function of seemingly minor errors in the data as well as by choice of algorithm, precision of the computer program or the model. Indeed, under the assumptions stated above, computations using very high accuracy may result in a solution very unlikely to be close to the most likely solution. The results of these experiments are summarized in Table 9.

Insert Table 9 about here

The first line in the table compares the regression solution with three variables to the solution with six variables. The marked effect of the addition of highly correlated independent variables should be no surprise to statisticians. In none of the seven coefficients was there a single significant digit of agreement. Two coefficients, B_0 and B_1 , differed by orders of magnitude. Since the three-variable model assumes that the coefficients B_2 , B_5 , and B_6 are identically zero, the number of digits of agreements for these is exactly zero.

Then the thousand perturbed solutions using DORTHO were compared with the unperturbed solutions for both the three- and six-variable problems. The perturbations affected the regression coefficients after the second significant digit for the three-variable problem. For the six-variable problem, no regression coefficient averaged a single significant digit of agreement with the unperturbed solution. Three coefficients averaged differences in orders of magnitude. The effect of perturbation is, therefore, very serious in the highly collinear six-variable model.

The next comparison is between the DORTHO algorithm and the DLSSQ algorithm, both of which are double precision. The same thousand perturbed sets of data were given to both programs and the results compared. We note that the algorithms agreed to an average of over 13 significant digits for the three-variable problem and over 7 significant digits in the six-variable problem. We conclude, therefore, that the choice of algorithm is not important if sufficient precision is used.

The last comparison is between the double precision (DLSSQ) and single precision (WRYLG) least squares algorithms. It is clear that precision must at some point affect a solution. For the three-variable problem, the effect of single precision versus double precision was, on the average, after the third place or approximately an order of magnitude less than the effect of perturbation. For the six-variable problem, 650 solutions were rejected because of negative pivots, and the remaining 350 comparisons of sets of regression coefficients did not average a single significant digit of agreement; in fact, four coefficients were off by an order of magnitude.

We cannot know what the "true" solution to the Longley problem is, but it is clear that the use of stable algorithms and high precision is not likely to yield a valid answer without more accurate data. If data are sufficiently accurate, then the additional labor of special algorithms may be worth the trouble, but we feel that in many cases the attempt at estimating regression coefficients in highly collinear problems cannot be justified statistically. We propose that the perturbation index discussed in section 7 be used routinely to indicate the existence of severe instability in regression solutions.

References

Berkson, J. Are there two regressions? Journal of the American Statistical Association, 1950, 45, 164-180.

Cochran, W. G. Errors of measurement in statistics. Technometrics, 1968, 10, 637-666.

Golub, G. H. Matrix decompositions and statistical calculations. In R. C. Milton and J. A. Nelder (Eds.), Statistical computation. New York: Academic Press, 1969.

Hilsenrath, J. et al. Omnitab, a computer program for statistical and numerical analysis. U.S. Department of Commerce, National Bureau of Standards Handbook 101. Washington, D.C.: U.S. Government Printing Office, March 4, 1966.

IBM Corporation, System/360 Scientific Subroutine Package, 1968. White Plains, N.Y., form # H20-0205-3.

Longley, J. W. An appraisal of least squares programs for the electronic computer from the point of view of the user. Journal of the American Statistical Association, 1967, 16, 319.

Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Mullet, G. M., & Murray, T. W. A new method for examining rounding error in least squares regression computer programs. Journal of the American Statistical Association, 1971, 66, 235.

Neely, P. M. Comparison of several algorithms for computation of means, standard deviations and correlation coefficients. Communications of the ACM, 1966, 9:7, 497-499.

Tausworthe, R. C. Random numbers generated by linear recurrence modulo two. *Mathematics of Computation*, 1965, *19*, 201-209.

United States Department of Commerce. A supplement to the survey of current business. Washington, D.C.: Office of Business Economics, 1958.

Walsh, P. J. Algorithm 127. *Communications of the ACM*, 1962, *5*, 511-513.

Wampler, R. H. A report on the accuracy of some widely used least squares computer programs. *Journal of the American Statistical Association*, 1970, *65*, 330.

Wilkinson, J. H. *The algebraic eigenvalue problem*. Oxford: Clarendon Press, 1965.

Wilkinson, J. H. The solution of ill-conditioned linear equations. In A. Ralston and H. S. Wilf (Eds.), *Mathematical methods for digital computers*, Vol. 2. New York: Wiley, 1967.

Footnotes

[1]However, moderate simple correlations do not guarantee good conditioning.

[2]The uniform pseudo-random numbers were produced by a double precision Tausworthe (1965) generator.

[3]The ORTHO algorithm which Longley found most satisfactory for fixed word-length computers was programmed to investigate the Longley data. Our subroutine is called DORTHO and is programmed in double precision for the IBM 360/65. Although the algorithm is classical Gram-Schmidt in nature, it avoids numerical instability by a second-stage correction. DORTHO produces regression coefficients that agree with Longley's hand calculated solution in every published place. We have been very careful in coding this routine since we wish there to be no question about its accuracy, even though the following arguments call for belief in only one or two significant digits.

[4]The lines involving WRYLG will be discussed in the next section.

[5]The IBM solution was computed with two programs (CORRE and MULTR) in the IBM Scientific Subroutine Package (SSP). These subroutines were designed for single precision (24-bit mantissa), but the comments in the program instruct the user that, to make the routine double precision, one need only make some arrays double precision. A statement using one of these arrays is

R(JK) = R(JK) + D(J)*D(K)

where R is a double precision matrix in which sums of squares and cross-products are accumulated and D(J) and D(K) are single precision data. Unfortunately, in IBM FORTRAN the product of two single precision numbers is a single precision number, thus FORTRAN compiles instructions to multiply

D(J) by D(K) , then truncates the product to single precision; since the sum of single precision and double precision addends is double precision, the program fills out the least significant part of D(J)*D(K) with zeros before adding. These sums of squares and products are, of course, but little better than single precision. This FORTRAN idiosyncracy could have been avoided by the statement

```
R(JK) = R(JK) + DBL(D(I))*D(J)
```

and the resulting regression solution would have been considerably closer to Longley's.

[6] We have used the method of calculating significant digits suggested by Jordan (1968) and used by Wampler (1970).

[7] To compute the Wilkinson error bound the quantity

$$2^{-t} n^{.5} k(A)$$

must be less than unity, where $t = 56$ is the number of bits in the mantissa of a computer word, $n = 7$ is the number of equations being solved, and $k(A) = 2.361 \times 10^{19}$ is the spectral norm. For this problem the result is approximately 8.7.

[8] Note that these assumptions correspond to the Berkson case (see Berkson, 1950, or Cochran, 1968) and are not the usual assumptions of errors of measurement (see Lord & Novick, 1968 or Cochran, 1968).

[9] Actually this relation holds only if the vector of constant ones is included (i.e., if the regression plane is not forced through the origin) and does not hold for the diagonal corresponding to the constant variable. More generally, s_j^2 should be interpreted as the raw sum of squares of X_j and R_j the cosine of the angle between X_j and the hyperplane spanned by the other independent variables.

[10]This index is called t_j by Longley. We note that this index is also subject to computational error but our numerical experiments indicate that such error is not important.

TABLE 1

Longley Data and Descriptive Statistics

	Y Total Derived Employ.	X_1 GNP Price Deflator	X_2 Gross Nat. Pro.	X_3 Unemployment	X_4 Size of Armed Forces	X_5 Noninst. Pop. 14 Yrs. & Over	X_6 Time
RAW DATA	60,323	83.0	234,289	2,356	1,590	107,608	1947
	61,122	88.5	259,426	2,325	1,456	108,632	1948
	60,171	88.2	258,054	3,682	1,616	109,773	1949
	61,187	89.5	284,599	3,351	1,650	110,929	1950
	63,221	96.2	328,975	2,099	3,099	112,075	1951
	63,639	98.1	346,999	1,932	3,594	113,270	1952
	64,989	99.0	365,385	1,870	3,547	115,094	1953
	63,761	100.0	363,112	3,578	3,350	116,219	1954
	66,019	101.2	397,469	2,904	3,048	117,388	1955
	67,857	104.6	419,180	2,822	2,857	118,734	1956
	68,169	108.4	442,769	2,936	2,798	120,445	1957
	66,513	110.8	444,546	4,681	2,637	121,950	1958
	68,655	112.6	482,704	3,813	2,552	123,366	1959
	69,564	114.2	502,601	3,931	2,514	125,368	1960
	69,331	115.7	518,173	4,806	2,572	127,852	1961
	70,551	116.9	554,894	4,007	2,827	130,081	1962
Mean	65317.0000	101.6813	387698.4375	3193.3125	2606.6875	117424.0000	1954.5000
S.D.	3511.9684	10.7916	99394.9378	934.4642	695.9196	6956.1016	4.7610
Mean/S.D.	18.5984	9.4222	3.9005	3.4172	3.7456	16.8807	410.5229
CORRELATIONS	1.0000	0.9709	0.9836	0.5025	0.4573	0.9604	0.9713
	0.9709	1.0000	0.9916	0.6206	0.4647	0.9792	0.9911
	0.9836	0.9916	1.0000	0.6043	0.4464	0.9911	0.9953
	0.5025	0.6206	0.6043	1.0000	-0.1774	0.6866	0.6683
	0.4573	0.4647	0.4464	-0.1774	1.0000	0.3644	0.4172
	0.9604	0.9792	0.9911	0.6866	0.3644	1.0000	0.9940
	0.9713	0.9911	0.9953	0.6683	0.4172	0.9940	1.0000

TABLE 2

Unperturbed Solution and Related Statistics (DLSSQ)

Squared Multiple Correlation = 0.9955
Standard Error of Estimate = 304.8541

Variable	Regression Coefficients B_j	Standard Error of Regression Coefficients SE_{B_j}	Student's t for $H_o: \beta_j = 0$	Contribution to Multiple Correlation
X_0	-3482258.6348	890420.3836	-3.9108	———
X_1	15.0619	84.9149	0.1774	0.0000
X_2	-0.0358	0.0335	-1.0695	0.0006
X_3	-2.0202	0.4884	-4.1364	0.0086
X_4	-1.0332	0.2143	-4.8220	0.0117
X_5	-0.0511	0.2261	-0.2261	0.0000
X_6	1829.1515	455.4785	4.0159	0.0081

TABLE 3

Some Calculated Least Squares Solutions to Longley Data*

Program	B_0	B_1	B_2	B_3	B_4	B_5	B_6
Desk Calculator	-3482258.6330	15.061872271	-0.035819179	-2.020229803	-1.033226867	-0.051104105	1829.15146461
IBM 1401	Not Calculated	15.061872271373	-0.035819179292	-2.020229803816	-1.033226867173	-0.051104105653	1829.151464613355
DORTHO (IBM 360)	-3482258.634597	15.061872271761	-0.035819179293	-2.020229803818	-1.033226867174	-0.051104105653	1829.151464614112
DLSSQ (IBM 360)	-3482258.634818	15.0618723573	-0.035819179316	-2.020229804106	-1.033226867254	-0.051104105500	1829.151464718495
Program D (IBM 7074)	-269126.040	-36.81780	0.05905	-0.59308	-0.60657	-0.34354	183.73361
D.T.S. (GE 235)	-3442970.0	13.9439	-0.0345996	-2.00228	-1.02763	-0.0548529	1809.05
IBM (IBM 360, single)	3483380.0000	15.00159	-0.03581	-2.02009	-1.03331	-0.05138	1829.74341
IBM (IBM 360, double)	3480987.91882	15.04414	-0.03578	-2.01066	-1.03308	-0.05124	1828.50156
BMD (IBM 7094)	146263.53125	27.08203	-0.03194	-1.94616	-0.98694	-0.01299	1653.37500
NIPD (IBM 7094)	-5.12128699	-0.04135583	0.54787309	-0.06653126	-0.06401523	-3.54369950	30.99588871
NIPD' (IBM 7094)	-5.22	-0.0387	0.536	-0.0680	-0.0641	-3.44	31.5
ORTHO (IBM 7094)	-34823280.0	15.065641	-0.035821413	-2.0202494	-1.0332373	-0.051100187	1829.1871

*For details on all programs except DORTHO and DLSSQ see Longley (1967). DORTHO and DLSSQ were included for comparison and are discussed in text.

TABLE 4a

Results of Numerical Experiment on 6 Variable Perturbed Data
(DORTHO)
N=1000

	B_0	B_1	B_2	B_3	B_4	B_5	B_6
Unperturbed Solution	-3482258.6348	15.0619	-0.0358	-2.0202	-1.0332	-0.0511	1829.1515
Mean	-1152648.3737	-26.4404	0.0344	-0.9637	-0.7209	-0.2804	637.0983
Median	-1185031.8499	-31.6441	0.0363	-0.9384	-0.6978	-0.3065	652.6028
S.D.	637918.2596	59.7325	0.0243	0.3507	0.1498	0.1608	325.9498
Lowest Value	-3483280.6948	-232.2792	-0.0880	-2.4234	-1.3259	-0.9393	-1706.8911
Highest Value	3452563.4777	237.0467	0.1964	1.7670	0.3633	0.4805	1800.8851
Percent Agreement *	2.00	3.90	0.70	6.4	94.6	0.90	1.00
t-value for mean=Longley	115.4220	-21.9604	91.3060	95.2157	65.8909	-45.0682	-115.5895
P-lim	-806545.7479	-32.4549	0.0449	-0.8104	-0.6794	-0.3148	460.0103
t-value for mean=P-lim	-17.155	3.[illegible]	-13.661	-13.822	-8.760	6.764	17.178

*The percent of calculated solutions t[illegible] [illegible]gree with unperturbed Solution to at least one digit.

Table 4[illegible]
Squared Multiple Correlations

	Y Unperturbed	Y Perturbed	s Perturbed	Low	High	t-value for Perturbed = Unperturbed
	[illegible]	[illegible]	[illegible]	0.9874	0.9960	-65.9623

TABLE 4e
Predicted Values

	Y Unperturbed	Y Perturbed	s Perturbed	Low	High	t-value for Perturbed = Unperturbed
19[illegible]	600[illegible]	600[illegible].5894	143.8050	59669.7157	60613.7569	3.4957
1948	[illegible]	61231.5866	129.4594	60819.6701	61600.3754	3.7517
[illegible]	[illegible]	60092.1840	149.9160	59617.2684	60576.5399	-6.8555
1950	[illegible]	61573.3563	151.5326	60978.3381	61900.3515	-5.5736
1951	[illegible]	63407.[illegible]	189.7322	62858.6714	63809.6771	82.5856
1952	63860.[illegible]	64065.5677	174.3887	63447.7752	64454.0388	32.1252
[illegible]	65253.0409	64914.3581	150.2856	64425.8586	65360.1665	-50.1980
1954	6377[illegible].1804	63628.9526	144.7665	63170.0848	64078.4692	-37.7048
1955	66024.6952	65954.6363	136.8382	65489.8427	66428.2652	-11.3618
1956	67401.6059	67120.1842	203.0546	66779.7320	67861.3419	-43.8040
1957	68186.26[illegible]	67919.9428	180.2461	67441.2040	68499.7698	-46.6991
1958	66552.0550	665[illegible]	129.1780	66135.5024	66919.3946	-4.8895
1959	6881[illegible]	68923.6105	160.5254	68398.7555	69309.6653	22.2697
1960	6964[illegible]	6957[illegible]	166.7460	69076.5119	70092.5064	-14.8304
[illegible]	[illegible].0695	6[illegible].3086	147.3297	68766.5898	69659.9349	32.4476
[illegible]	[illegible]	709[illegible]	132.2663	[illegible]	71217.9229	40.7887

TABLE 5

Four Variable Reduced Model

		B_0	B_1	B_3	B_4
Unperturbed Data	DORTHO	30393.4125	398.0758	-1.0725	-.8165
	WRYLG	30389.2344	398.1814	-1.0735	-.8178
DORTHO 1000 Perturbations	Mean	30393.6397	398.0705	-1.0725	-0.8164
	Median	30393.2430	398.0951	-1.0725	-0.8166
	Standard Deviation	32.2661	0.5418	0.0056	[illegible]
	Lowest Value	30289.7507	396.2621	-1.0894	[illegible]
	Highest Value	30486.0691	399.7687	-1.0532	[illegible]
	Percent* Agreement	100.0	100.0	100.0	[illegible]
	t-value for mean=unperturbed	.2226	-.3092	-.1477	[illegible]
	P-Lim	30393.9794	398.0647	-1.0724	[illegible]
	t-value for mean=P-lim	-.3328	.3384	-.3921	[illegible]
Difference Between DORTHO and WRYLG 1000 Perturbations	Mean	0.7602	-0.0043	-0.0001	[illegible]
	Standard Deviation	3.0381	0.0606	0.0003	[illegible]
	Largest Negative Difference	-8.9499	-0.1961	-0.0017	[illegible]
	Largest Positive Difference	9.9655	0.1873	0.001	[illegible]
	Avg. Number of Agreeing Sig. Dig.	4.2386	4.0646	3.583	[illegible]
	t-value for mean difference=0	7.909	-2.243	-6.320	[illegible]

TABLE 6

Comparison of DORTHO and DLSSQ

N=1000

	B_0	B_1	B_2	B_3	B_4	B_5	B_6
Avg. # of Sig. Dig.	7.7341	7.9117	7.7358	8.0529	8.5755	8.2149	7.7653
Fewest Significant Digits	6.8812	4.3409	4.7483	6.4017	6.6976	4.5182	7.0044

TABLE 7

Results of Numerical Experiment on 6 Variable Perturbed Data

WRYLG
N=350*

	B_0	B_1	B_2	B_3	B_4	B_5	B_6
Mean	-1278608.1341	-597.7108	0.2348	1.4467	-0.8938	-2.6038	827.3905
Median	-381896.2187	-233.3836	0.1288	0.2372	-0.6665	-1.1866	289.2565
Standard Deviation	4248967.8715	1619.9556	0.4915	5.5652	1.3985	6.7117	2554.1080
Lowest Value	-6072246u.uuuu	-21880.4805	0.0756	-0.3648	-24.0115	-93.2490	36.4833
Highest Value	34900.2070	-56.6455	6.7673	77.3966	-0.4950	-0.4939	36405.9688
Percent Agreement**	2.00	0.00	0.00	0.00	95.43	0.00	2.86

*650 samples were rejected due to a nonpositive pivot.
**The percent of 350 calculated solutions that agree with unperturbed Solution to at least one digit.

TABLE 8

Perturbation Indices

Four Variable Reduced Model

	$Diag[(X'X)^{-1}]$	d_i	$Diag(X'X)^{-1}16d_i$
B_0	$.7569 \times 10^{1}$	0	
B_1	$.2092 \times 10^{-2}$	$\frac{1}{1200}$	$.2789 \times 10^{-4}$
B_3	$.2259 \times 10^{-6}$	$\frac{1}{12}$	$.3011 \times 10^{-6}$
B_4	$.3193 \times 10^{-6}$	$\frac{1}{12}$	$.4256 \times 10^{-6}$

Full Model

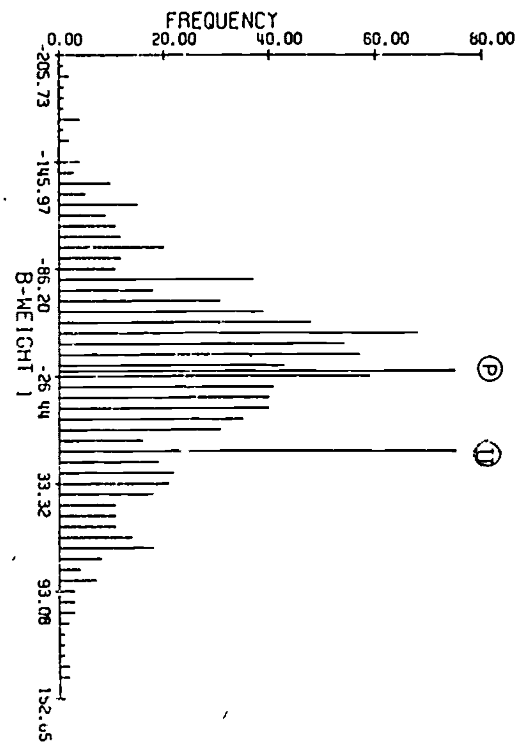
	$Diag[(X'X)^{-1}]$	d_i	$Diag(X'X)^{-1}16d_i$
B_0	$.8531 \times 10^{7}$	0	
B_1	$.7759 \times 10^{-1}$	$\frac{1}{1200}$	$.1034 \times 10^{-2}$
B_2	$.1207 \times 10^{-7}$	$\frac{1}{12}$	$.1608 \times 10^{-7}$
B_3	$.2567 \times 10^{-5}$	$\frac{1}{12}$	$.3422 \times 10^{-5}$
B_4	$.4940 \times 10^{-6}$	$\frac{1}{12}$	$.6586 \times 10^{-6}$
B_5	$.5499 \times 10^{-6}$	$\frac{1}{12}$	$.7331 \times 10^{-6}$
B_6	$.2232 \times 10^{1}$	$\frac{1}{12}$	$.2976 \times 10^{0}$

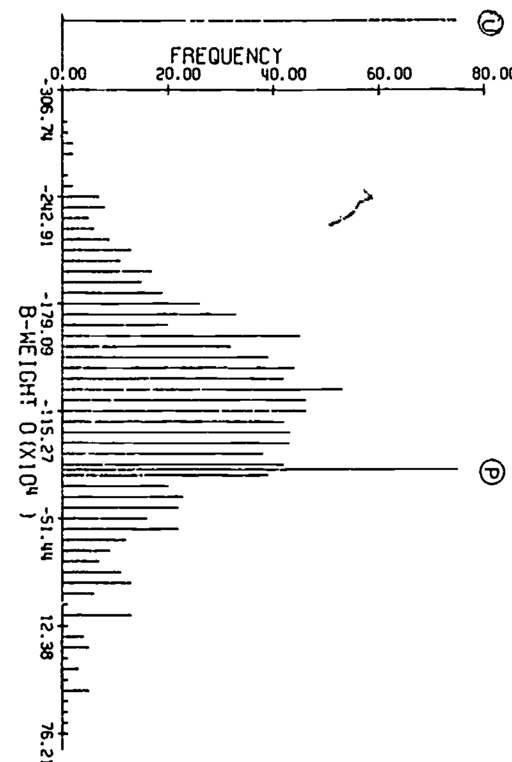
TABLE 9

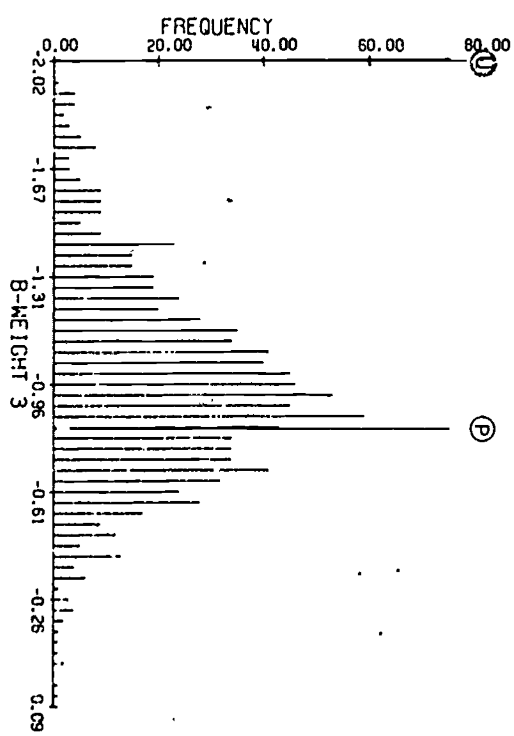
Average Number of Significant Digits Between Pairs of Regression Coefficients

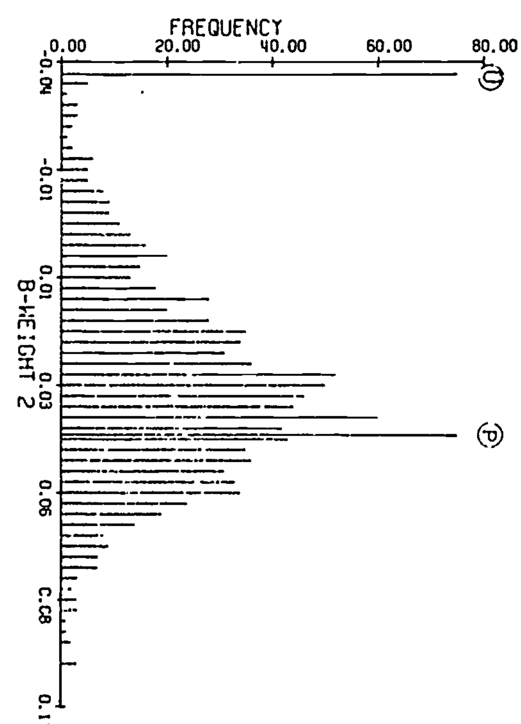
Differences Due to	N	B_0	B_1	B_2	B_3	B_4	B_5	B_6
Model	1	- 0.0038	- 1.4053	0.0000	0.3287	0.6783	0.0000	0.0000
Three Variable: Perturbation	1000	3.2482	3.1475	*	2.5409	2.3695	*	*
Algorithm	1000	13.9414	13.7571	*	13.2317	13.0691	*	*
Precision	1000	4.2386	4.0646	*	3.5830	3.4194	*	*
Six Variable: Perturbation	1000	0.1941	- 0.4632	-0.2559	0.3118	0.5850	- 0.5920	0.2037
Algorithm	1000	7.7341	7.9117	7.7358	8.0529	8.5755	8.2149	7.7653
Precision	350	0.1214	-0.8496	-0.6104	-0.1823	0.9189	-0.6461	0.1783

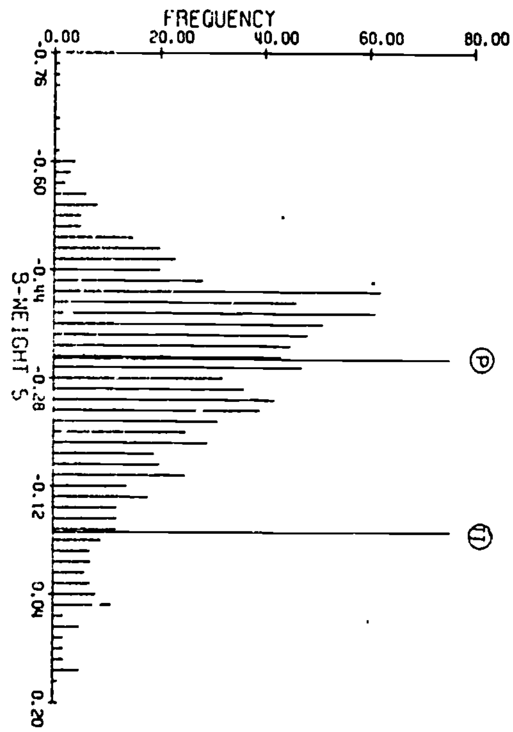
* Word length of computer
Negative numbers indicate differences in orders of magnitude.

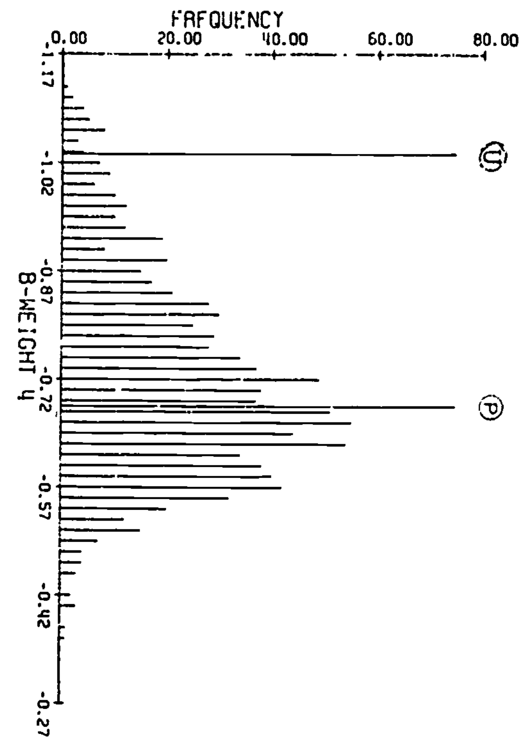

FREQUENCY
0.00 20.00 40.00 60.00 80.00
-205.73 -145.97 -86.20 -26.44 33.32 93.08
B-WEIGHT 1
P
U

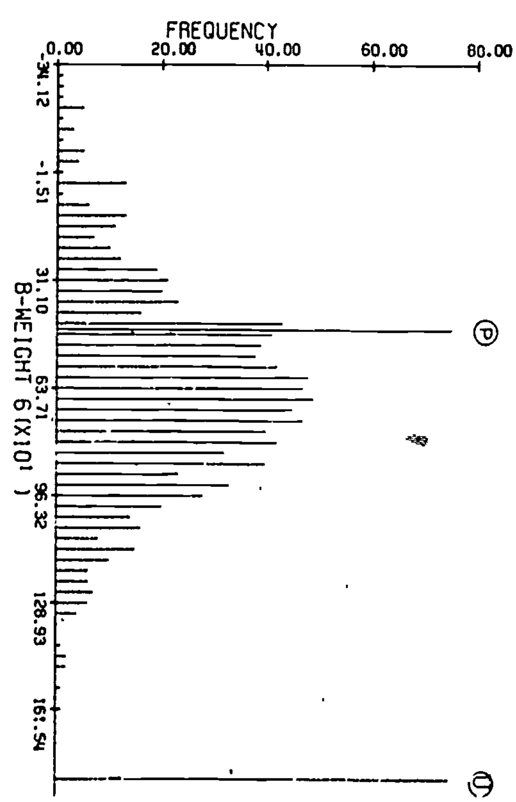

FREQUENCY
0.00 20.00 40.00 60.00 80.00
-306.74 -242.91 -179.09 -115.27 -51.44 12.38 76.21
B-WEIGHT 0(X10⁴)
U
P


FREQUENCY
0.00 20.00 40.00 60.00 80.00
-2.02 -1.67 -1.31 -0.96 -0.61 -0.26
B-WEIGHT 3
U
P


FREQUENCY
0.00 20.00 40.00 60.00 80.00
-0.04 -0.01 0.01 0.03 0.06 0.08 0.11
B-WEIGHT 2
U
P


FREQUENCY
0.00
20.00
40.00
60.00
80.00
-0.76
-0.60
-0.44
-0.28
-0.12
0.04
0.20
B-WEIGHT 5
P
T


FREQUENCY
0.00
20.00
40.00
60.00
80.00
-1.17
-1.02
-0.87
-0.72
-0.57
-0.42
-0.27
B-WEIGHT 4
U
P


FREQUENCY
0.00
20.00
40.00
60.00
80.00
-34.12
-1.51
31.10
63.71
96.32
128.93
161.54
B-WEIGHT 6 (X10¹)
P
U