

AD-A167 601

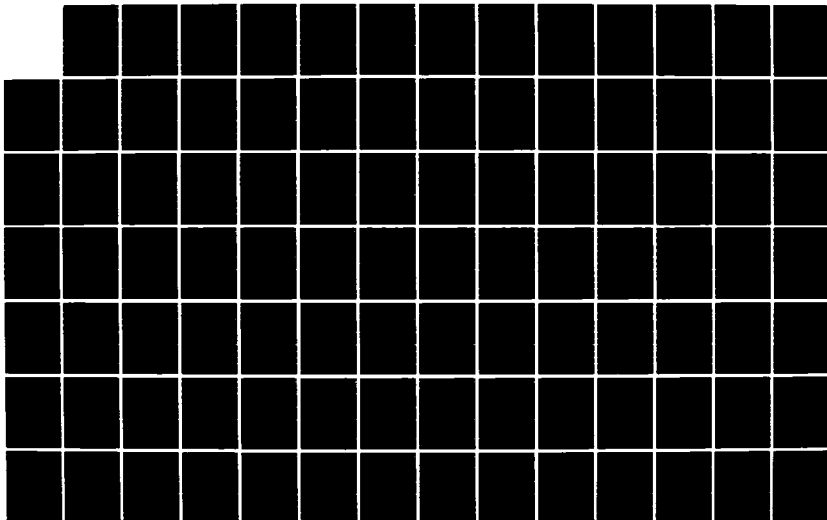
IMPROVED ALGORITHMS FOR ESTIMATION PREDICTION AND  
CONTROL(U) VISTA RESEARCH CORP TUSCON AZ J G CALDWELL  
26 JAN 86 N00014-85-C-0014

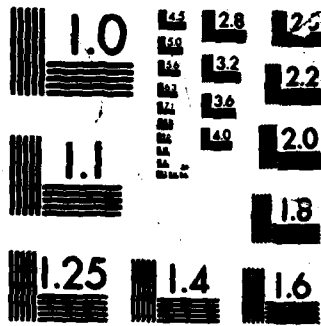
1/2

UNCLASSIFIED

F/G 9/2

NL





MICROCOPY

CHART

12

AD-A167 601

IMPROVED ALGORITHMS  
FOR ESTIMATION,  
PREDICTION AND CONTROL

Final Report

January 26, 1986

Contract No. N00014-85-C-0814

DTIC FILE COPY

DTIC  
SELECTED  
MAY 5 1986

This document has been approved  
for public release and sale; its  
distribution is unlimited.

VISTA RESEARCH CORPORATION

86 4 15 092

12

IMPROVED ALGORITHMS  
FOR ESTIMATION,  
PREDICTION AND CONTROL

Final Report

January 26, 1986

Contract No. N00014-85-C-0814

Submitted to:

Dr. Neal D. Glassman  
Office of Naval Research  
Department of the Navy  
800 N. Quincy Street  
Arlington, Virginia 22217-5000

Submitted by:

Vista Research Corporation  
4540 Cerco del Corazón  
Tucson, Arizona 85718  
(602) 299-0286

DTIC  
ELECTE  
S MAY 5 1986 D  
A

## **DISCLAIMER NOTICE**

**THIS DOCUMENT IS BEST QUALITY  
PRACTICABLE. THE COPY FURNISHED  
TO DTIC CONTAINED A SIGNIFICANT  
NUMBER OF PAGES WHICH DO NOT  
REPRODUCE LEGIBLY.**

TABLE OF CONTENTS

Section	Page
Foreword . . . . .	i
I. Introduction and Summary . . . . .	1
II. Background . . . . .	11
III. Project Approach . . . . .	24
IV. Simulation Results . . . . .	37
V. Conclusions and Recommendations. . . . .	41
References. . . . .	44
Appendices	
A. Computer Program Source Code Listings. . . . .	A-1
B. Computer Program Output. . . . .	B-1

Requested For	
By	
Approved	<input checked="" type="checkbox"/>
Disapproved	<input type="checkbox"/>
Special	<input type="checkbox"/>
<i>Attaches on file</i>	
By	
Approved	
Disapproved	
Special	
First Special	
A-1	23



## FOREWORD

This report was prepared under Contract No. N00014-85-C-0814 with the Office of Naval Research (ONR), under Phase I of the Small Business Innovation Research (SBIR) Program. Vista Research Corporation expresses its appreciation to ONR, to the SBIR Program, and to Dr. Neal D. Glassman, ONR Scientific Officer, for their support in sponsoring this work. The author of this report was Dr. J. George Caldwell.

## I. INTRODUCTION AND SUMMARY

### A. Study Purpose

This report describes the results of a study to examine the feasibility of developing fast algorithms for estimation, prediction and control. The objective of the study was to assess the likelihood of finding procedures which could be faster, in terms of computer running time, than the classical least-squares method of parameter estimation, used extensively to develop models of stochastic phenomena. While the least-squares method has proved its worth in over a century of use, it has some serious drawbacks, stemming from the fact that it requires the inversion of matrices. For "large" problems such as the problem of tracking many missiles or processing data from multiple intelligence sensors in real-time, the computational burden of the least-squares method can overwhelm even today's powerful computing systems. Previous attempts to solve this problem have centered on the development of faster computers (e.g., array processors), the improvement of algorithms for matrix inversion, or the simplification of the model to produce a matrix that is easier to invert. In general, these approaches have not been successful in solving the problem. Modern sensor exploitation systems, for example, still cannot operate in real-time or even near-real-time.

The present study proposed to adopt a totally different approach to the problem. In particular, it was proposed to investigate methods which would avoid the computationally-intensive process of matrix inversion. Avoiding this procedure could reap tremendous benefits. For example, the problem of tracking a missile can involve the inversion of a nine-dimensional matrix at each instant that a radar pulse is received, if a nine-component state vector is used to represent missile position, velocity, and acceleration.

This project was supported as a Phase I project of the Small Business Innovation Research (SBIR) program. The objective of a Phase I SBIR project is to assess the feasibility of a proposed concept and to develop a plan for developing the concept. If, based on the Phase I effort, the proposed



concept appears to have merit, then development of the concept may be funded under Phase II of the SBIR program.

## **B. Study Results**

This Phase I project has successfully accomplished all of the tasks identified in the proposal, and established the feasibility of the proposed concept. Seven tasks were proposed to be accomplished in this study:

- Task 1. Development of Criteria for Comparing Alternative Estimation Schemes
- Task 2. Development of Test Cases
- Task 3. Implementation of the Single-Variable Linear-Model Case
- Task 4. Extension to the Multiple-Variable Linear-Model Case
- Task 5. Comparison of Methods
- Task 6. Design of Phase II Work Plan
- Task 7. Preparation of Final Report

The results of each of the seven study tasks are summarized in the paragraphs that follow.

### **Task 1. Development of Criteria for Comparing Alternative Estimation Schemes**

A total of seven criteria were identified and retained as useful for describing the performance of alternative estimation algorithms. These criteria address the following performance aspects:

- o computational speed
- o computer storage requirements
- o precision of the model parameter estimates
- o bias of the model parameter estimates
- o accuracy of the model parameter estimates
- o precision of model-based predictions

## o numerical stability of the algorithm

"Computational speed" refers to the time required to analyze the data and produce estimates of the model being used to describe the data. "Computer storage requirements" refers to the total amount of direct-access memory ("core") required to implement the algorithm. "Estimate precision" refers to the amount of variability of the estimates in repeated data samples. "Estimate bias" refers to the difference between the expected (average) value of the parameter estimates in repeated data samples and the true values of the parameters. "Estimate accuracy" is a combined measure of precision and bias. "Prediction precision" refers to the closeness of the model-based predictions to actual future values. "Numerical stability" refers to the ability of an algorithm to converge to a desired answer.

Specific measures were determined for each of the preceding concepts. Because of project resource limitations, however, it was not possible to develop computer software to determine numerical values for all of the measures.

### Task 2. Development of Test Cases

It was decided to test the performance of alternative algorithms on sixteen data sets. All of these data sets involve a single dependent variable ("y") and a number (m) of independent variables ("x's"). In each case the model used to generate the data is of the form:

$$y_j = b_0 + \sum_i b_i x_{ij} + e_j$$
$$= b_0 + \underline{x}_j' \underline{b} + e_j$$

This model, called a linear statistical model, specifies the relationship of the dependent variable y to m independent, or explanatory, variables ( $x_{1j}, x_{2j}, \dots, x_{mj}$ ). The variables  $b_0$  and  $b_1, b_2, \dots, b_m$  are constants, called regression coefficients. They are specified when generating the test data, and are to be estimated by the estimation algorithm. The  $e_j$  are called model error terms, or model residuals. They are a sequence of uncorrelated random variables with mean zero and standard deviation SIG. The m x's are also random variables with zero mean and variance matrix  $\Sigma * \text{SIGMA}^2$  (where the asterisk denotes multiplication). The test cases vary in the ratio of SIG to SIGMA (models for which the ratio SIG/SIGMA is low are

easier to estimate), and the strength of the correlations among the x's (if the x's are uncorrelated, i.e.,  $\Sigma$  is the identity matrix, the estimation is easy; if the x's are highly correlated, the estimation is difficult). The values of SIG, SIGMA,  $\Sigma$ ,  $b_0$ , and  $\underline{b}$  are collectively referred to as the "model parameters."

The test cases considered were as follows:

Test Cases 1-4:  $m=1$  x, value of SIG/SIGMA varies from low to high

Test Cases 5-8:  $m=3$  x's, correlation of x's varies from none to high

Test Cases 9-12:  $m=6$  x's, correlation of x's from none to high

Test Cases 13-16:  $m=10$  x's, correlation of x's from none to high

The number ( $m$ ) of independent variables (x's) represents the "dimensionality" of the estimation problem. In the classical least-squares approach, it is necessary to invert a matrix of order  $m$  in order to estimate the model parameters.

In social science applications, the dimensionality of a linear regression model can be quite high, e.g.,  $m = 25$  to  $50$  (e.g., there can be a model coefficient corresponding to every possible response to a socioeconomic or demographic question included in a survey questionnaire). In industrial and scientific applications, the number of explanatory variables may vary from small to large, but the researcher is often able to specify the values of the independent variables, so that even if there are many of them, special procedures (e.g., fractional factorial experimental designs) can be used to avoid the explicit inversion of a matrix. In military applications, the value of  $m$  is often moderate or small. For example, in the application of tracking a missile, the position, velocity, and acceleration of the missile can be specified by a nine-component state vector, and the estimation of the parameters of the statistical model (called a Kalman filter) requires inversion of a nine by nine matrix. The test cases specified above include the usual dimensionality range of interest for many military applications. For all test cases, the number of observations generated (i.e., the sample size) was  $n = 100$ . Later study should examine the effect of varying sample size; with the resources available to this project, it was

not possible to examine a very large number of test cases, and the decision was made to hold sample size constant for all test cases.

A computer program, called SIMULA, was written to implement the generation of the test case data. A source code listing of that program is presented in Appendix A.

### Task 3. Implementation of the Single-Variable Linear-Model Case

In the 1940's and 1950's, some work was done in investigating model estimation procedures that were alternatives to the classical least-squares procedure. Two of these procedures are referred to as the "Wald" method and the "Bartlett" method. They were designed for application to the case in which there was a single explanatory variable (i.e.,  $m = 1$ ). A computer program was written in this project to implement both of these procedures.

As one of the first steps in this study, it was proposed to compare the performance of the Wald and Bartlett estimation procedures to the performance of the classical least squares method, in order to illustrate the utility of the criteria and associated performance measures proposed to compare alternative estimation procedures.

This comparison demonstrated that several of the suggested criteria could be applied to measure algorithm performance. For the single-explanatory-variable example, however, the comparison is not very revealing. The estimation of parameters for single-explanatory-variable ( $m = 1$ ) models is very easy and fast with any of the methods, so that there is almost no variation in the performance measures.

It was not possible to implement all of the performance measures in this project, because of resource limitations. For example, one problem that arose was that, for the microcomputer software used in this study, the system timer could not be accessed by the FORTRAN compiler. It was decided not to allocate project resources to the development of an assembly-language timer that could be linked to the FORTRAN-compiled object modules. Instead, timing measurements were made externally (manually, by direct visual observation), and they are hence approximate. For other performance measures (e.g., the parameter accuracy measures), it would have been necessary to replicate a large number of sample cases to determine numerical estimates of these measures. Once again, project resources were not sufficiently ample to accomplish this. Although it was not

possible to implement all of the suggested performance measures in this Phase I effort, it is quite feasible to do so with some additional resources, and this should be done as part of Phase II, if Phase II is funded.

#### Task 4. Extension to the Multiple-Variable Linear-Model

Task 4 was the central task of the proposed study. The objective of this task was to determine fast algorithms for estimating parameters of models containing more than one explanatory variable. In this project, we synthesized and analyzed a number of algorithms that represent extensions of the Wald-Bartlett methods. They are iterative methods, and will be referred to as "iterative Wald-Bartlett" methods. Several variations of this method were considered, and the one that worked best was selected for detailed examination. This method is described in detail in this report, and all of the performance assessments that are presented in this report are for this method.

The thrust of this project was to compare the performance of new estimation techniques to the performance of the classical least-squares technique. In order to do this, a computer program was required that could perform the classical least-squares computations. We implemented the least-squares estimation procedure by using the Gauss-Jordan method of solving the normal equations (i.e., inverting the correlation matrix). This procedure is described in Chapter 3 of Cooley and Lohnes, Multivariate Procedures for the Behavioral Sciences (Reference 8), and is the basis for the least-squares estimation procedure presented in the IBM Scientific Subroutine Package (Reference 9). The least-squares algorithm used in this project was based on subroutines available from the IBM Scientific Subroutine Package, adapted to the Microsoft (R) FORTRAN compiler that was available to the microcomputer used in this project. Thorough documentation of the subroutines, including commented code, is included in Reference 9. A source code listing of the algorithm used in this project is presented in Appendix A. The listing in Appendix A does not include any comments in the subroutines, in order to avoid possible copyright infringement.

The computer program source code for the iterative Wald-Bartlett method and the classical least-squares method is also presented in Appendix A. These programs are written in FORTRAN II, an early, unstructured version of FORTRAN. That version was used since it was the version implemented by the Microsoft FORTRAN compiler used on this project.

## Task 5. Comparison of Methods

Task 5 was concerned with comparison of the performance of the iterative Wald-Bartlett method and the classical least-squares method, applied to estimate the parameters of the sixteen test cases developed in Task 2. The results were very interesting.

First, the performance of both the iterative Wald-Bartlett and the classical least-squares method depends on the nature of the problem. For "easy" problems, in which there are few x's or they are uncorrelated, both methods work well. Second, in problems of low to moderate difficulty, there does not appear to be an appreciable speed difference between the classical least-squares algorithm and the alternative algorithm synthesized in this study.

Third, both the classical least-squares and the iterative Wald-Bartlett methods have difficulties with very difficult problems (m large and the x's highly correlated). The really significant result that was observed in this case was that, whereas the classical method may fail catastrophically, producing totally-absurd results, the iterative Wald-Bartlett method is not particularly fast, but it determines an estimated model that produces reasonably close predictions. (Note: The original number of test cases planned to be examined was 12. It was after observing this result that we added four more test cases, representing singular covariance matrices, to examine this phenomenon in greater detail.)

The implications of this result are very significant, and probably outweighs the importance of speed in most applications, particularly since the processing speed differences between the iterative Wald-Bartlett and the classical least squares algorithm are not great. For example, in an embedded-computer application (e.g., a tracker on an unmanned missile) it may be very desirable to have a "robust" estimation or prediction procedure -- one that does not fail catastrophically. Moreover, the problem causing the catastrophic failure of the least-squares method (the failure of the algorithm to be able to invert a "nearly-singular" matrix) is one that has plagued data analysts for years -- ever since the widespread use of digital computers to implement the least-squares methodology. Matrix inversion problems in statistical analysis were not severe or widespread prior to the 1960's, when most computations were done using mechanical calculators. Most statistical calculations were done by masters-level mathematicians, and the problems were kept

small or designed to avoid the inversion of large matrices. Starting in the 1960's, however, computer packages and computing resources were widely available, and were being applied in many cases by researchers who had little or no appreciation of the matrix-inversion problem inherent in the least-squares method. With the large number of non-statisticians using statistical linear-model packages (e.g., multiple regression packages) and the low precision of many microcomputers, this problem arises frequently. Consulting statisticians are often called in on regression analysis investigations when regression analysis packages produce meaningless results due to a matrix-inversion problem caused by a near-singularity in a correlation matrix or a linear dependency in the x's. With the growing number of microcomputers, and the increasing number of non-statisticians using regression analysis programs (and the many other statistical procedures that require matrix inversion), there is a need for statistical estimation procedures that do not fail when near-singularities or linear dependencies are present in the data. The demand for "robust" statistical estimators probably represents a far greater commercial value than the demand for high-speed algorithms, since the demand for real-time processing represents only a small portion of the total demand for statistical estimation. This aspect should be explored further in Phase II.

In general, Task 5 succeeded in demonstrating the feasibility of the proposed approach. Although the iterative Wald-Bartlett method does not appear to be substantially faster than the classical least-squares method, it appears to be much more "robust" than the classical least-squares method. Also, since the iterative Wald-Bartlett method is based on "order statistics," it would be much less sensitive to data errors, such as "outliers," than the classical method. The results of the Phase I study suggest that additional exploration of this area would be very beneficial.

#### Task 6. Design of Phase II Work Plan

The results of this Phase I study have demonstrated the feasibility of determining fast, robust estimators. This study centered on the synthesis of a particular class of estimator, however, and that estimator certainly does not represent a final solution to the general problem. Although the iterative method considered in this study does not fail catastrophically in difficult or ill-conditioned problems, as does the classical least-squares method, it is not very fast in such cases. The present feasibility study has

revealed the promise of the proposed approach, but substantial development effort is necessary to produce an estimation method that works well (i.e., is both fast and robust) in all cases. We believe that the resources available in Phase II can accomplish that development effort, and have outlined a research plan for implementing the effort.

It is proposed that the Phase II effort be redirected from the goal originally proposed for Phase I. The Phase I research was directed solely to the problem of finding fast algorithms. Based on the Phase I results, however, it appears that there may be more potential (both in terms of project success and commercial value) in attempting the development of robust algorithms. The serendipitous discovery that the synthesized iterative method could produce solutions to problems for which the classical least-squares method failed catastrophically probably outweighs the promise of a fast algorithm, in terms of both military and commercial/industrial significance. Embedded processors in military weapon systems require software that is robust, i.e., does not fail catastrophically under certain circumstances. The current effort has demonstrated that it is indeed possible to develop such estimators. The growing use of microcomputers in commercial and industrial applications will create a growing demand for estimation procedures that do not fail in ill-conditioned problems, and do not require the participation of a professional statistician to assure their convergence to a correct solution. These applications include not only multiple linear regression models of the sort addressed in this study, but the whole range of modern-day data analysis procedures (multivariate analysis of variance, factor analysis, canonical correlation analysis, discriminant analysis, and time series analysis models), since all of these methods involve matrix inversion, which is the source of the slowness and potential for catastrophic failure of these methods. The availability of robust algorithms would be particularly beneficial in time series applications (e.g., "Box-Jenkins" analysis), where iterative estimation methods are often employed, and convergence problems are encountered by non-statistician users.

It is proposed that Phase II be a two-year effort, staffed at approximately 2-5 persons per year. A description of the proposed tasks to be addressed in the Phase II effort is included in this report.

#### Task 7. Preparation of Final Report



This document describes the activities, results, and conclusions of the Phase I study. In addition, it identifies the effort that should be implemented in Phase II, in order to complete the development of fast, robust algorithms for estimation, prediction, and control.

This Phase I study has accomplished each of the seven tasks identified and described in the proposal, in the proposed six-month time frame. We believe that our success in accomplishing all of the proposed tasks, on schedule, augurs well for the accomplishment of the Phase II objectives as well.

### C. Organization of the Report

The remainder of this report consists of four additional sections and two appendices. Section II ("Background") describes the motivation for the proposed research effort, both relative to the requirement for fast algorithms and for robust algorithms. Section III ("Project Approach") describes the methodology for conducting the Phase I investigation. The methodology consists of the specification of performance criteria, the simulation of test-case data with which to test the performance of alternative algorithms, the synthesis of one or more candidate "fast algorithms," and the comparison of the performance of the synthesized method to the classical least-squares method using the specified criteria and the test-case data. Section IV ("Simulation Results") presents the results of this simulation study. Section V ("Conclusions and Recommendations") summarizes the study conclusions and describes the additional research and development that is needed to develop and implement the proposed concept. The Appendices contain source code listings of all computer programs used in this study (Appendix A), and detailed output listings (Appendix B).

## II. BACKGROUND

### A. General Motivation for the Proposed Study

This report describes the results of a research study to assess the feasibility of developing fast algorithms for real-time estimation, prediction and control. Such algorithms would provide a solution to a critical problem faced in both industrial and military applications -- the fact that the algorithms used to implement state-of-the-art statistical estimation, prediction and control techniques are far too slow for many real-time or near-real-time applications of high interest, even using the fastest computers.

The slowness of statistical correlation/tracking techniques such as the Kalman-Bucy filter was one of the principal reasons for the failure of the ballistic missile defense program of the 1960's. This problem has still not been satisfactorily solved, even with substantial improvements in computer processing speed and direct-access storage capability. Modern command, control, communications, and intelligence (C<sup>3</sup>I) systems such as those intended to support the AirLand Battle concept, the Air Force's Tactical Air Control Center, or Naval tactical data systems require multisensor correlation/fusion to be performed in real time or near real time. The success of such systems is in jeopardy to the extent that they rely on data processing by traditional statistical algorithms.

In addition to military applications, the availability of fast prediction and control algorithms would assist control of rapid, time-varying processes occurring in industry and commerce (e.g., hot steel finishing mills and air traffic control), for which available prediction and control methods are generally heuristic, because of the current inability to conduct on-line system identification (and thereby use model-based predictors/controllers).

### B. The Requirement for Fast Prediction/Control Algorithms

Over the past two decades, tremendous advances have been made in the development of powerful statistical estimation techniques. The applications associated with these techniques cover a wide variety of fields, including

scientific, economic, industrial, and military applications. In general, present-day statistical estimation procedures represent extensions of the work of Gauss, who developed the least-squares method of estimating the position of an asteroid, based on observations which contained measurement errors. Gauss's method, or the method of least squares, consists in determining a set of model parameters in such a way that the sum of squares of the differences between the actual measurements and the position estimates based on the model is minimized.

A salient characteristic of the method of least squares is that it requires computation of a "cross-products" matrix (or a covariance matrix or a correlation matrix), and it requires the inversion of a matrix whose order depends on the number of parameters being estimated. In many applications, this presents no problems, since statistical analysis need not be done in "real-time." Instead, in most applications the analyst may collect the observed data, and then determine the parameter estimates "off-line" in the hours, days, or weeks that follow. In some applications, however, such as those involving the tracking of fast-moving objects (such as missiles, airplanes, or satellites), or the control of rapidly-changing systems (such as certain industrial processes), it is necessary to estimate the parameters "on-line," in real-time or near-real-time. In problems in which attention centers on only one or a few processes at a time (e.g., a small number of objects are being tracked, or a small number of electromagnetic emitters are being monitored), the computational burden is not severe. If numerous tracks (or emitters) are involved, or the underlying system changes "too fast," the method of least-squares breaks down -- the computational requirements may saturate even the most powerful (fastest, largest) computers available.

In an attempt to remedy this problem, a tremendous amount of effort has been expended on the development of computationally efficient algorithms for determining least-squares estimates. In 1960, Kalman and Bucy developed a recursive solution to the least-squares estimation problem, now called the Kalman filter (or Kalman-Bucy filter). Kailath (Reference 1) provided a comprehensive survey of over 600 references on filtering. Aggarwal (Reference 2) describes the problem of developing least-squares algorithms that are numerically stable and computationally efficient.

One approach to reducing the computational complexity of tracking algorithms is linearization. While this procedure

helps, it is not sufficient. As an example of the inadequacy of linearization, it is noted that the ballistic missile trackers proposed in the 1960's were in fact simplified nine-component Kalman filters in which the equations of motion had been linearized, and the covariance matrix radically simplified. Yet this approach still failed, because of the tremendous computational requirements of the general linear model and the least-squares estimates.

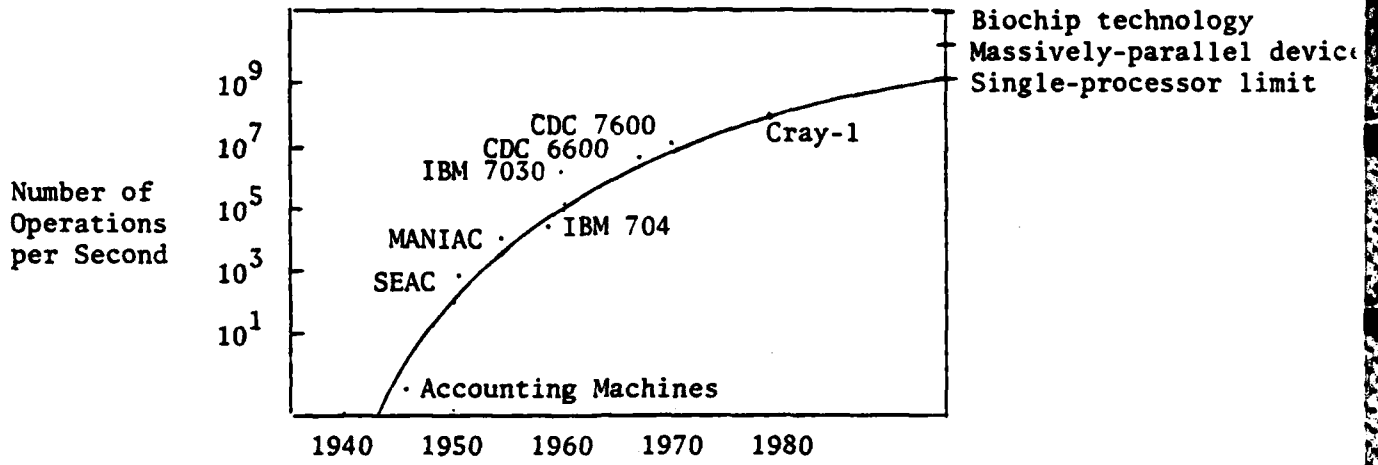
As a further example of the inadequacy of linearization, it is noted that neither the Kalman filter nor the Box-Jenkins (autoregressive-integrated-moving-average time series) models could outperform the heuristic alpha-beta tracker, in air traffic control studies of the early 1970's. As a final example, it appears that in complex correlation/tracking problems (e.g., satellite ocean surveillance, intelligence analysis of electromagnetic emissions for unit identification), heuristic nonlinear "algorithmic" procedures work better than procedures derived from linear-model theory.

The potential exists to reduce the computational burden of prediction/control algorithms by a factor of several orders of magnitude. Furthermore, this reduction can probably be achieved at very modest cost -- a few person-years of research effort. This investment is negligible, when compared to the massive research investment that will be required to develop even a hundred-fold increase in computer speed, through the development of an operational large-scale parallel-processor or bio-chip technology.

In addition to improvements in the computational efficiency of least-squares algorithms, tremendous gains have been made in the speed of the computers which perform the least-squares computations. Advances such as Very Large-Scale Integration (VLSI) technology have increased computer processing speeds and direct-access storage capabilities by a factor of one thousand since the mid-1960's, when the feasibility of performing least-squares tracking of incoming ballistic missiles was seen to exceed available computational capabilities.

Despite the tremendous computational gains of the last fifteen years, however, it appears that further advances may be elusive. Computer processing technology is now running up against physical limits, such as the time required for electronic signals to propagate along the wires inside the computer. As reported in a recent issue of Defense Science (Reference 3), advances in computational speed are leveling off (see Figure 1). In order to achieve processing

Figure 1. The Exponential Rate of Growth in Computer Processing Speed is Slowing



Sources: Defense Science, April 1984  
High Technology, February 1984

speed increases of two orders of magnitude or more with physical devices, it appears that parallel processing architectures will be required. Unfortunately, formidable problems are associated with the development of large-scale parallel processors and associated software, and a substantial amount of basic research and development will be required.

To achieve further increases will probably require the development of "bio-chips," in which synthetic organic molecules perform the binary switching functions of present-day physical switches such as silicon or gallium arsenide field-effect transistors. In principle, the use of bio-chip molecular switches could lead to circuit elements one thousand times smaller than may be achieved by conventional semiconductors. Developments in this area may be slow, however, since even the most promising of the bio-chips -- the soliton switch -- is still theoretical.

In summary, it does not appear likely to significantly improve the computational speed of least-squares algorithms, and the likelihood of realizing substantial increases in computer speed very soon is not high. In order to achieve substantial gains in processing speed in the near term, then, it appears that the most promising approach is to develop estimation procedures which impose a substantially reduced computational burden.

### C. Early Accomplishments in the Area of Low-Compute Estimation Procedures

#### 1. Experimental Design Applications

In the 1930's and 1940's, tremendous advances were made in the development and application of the general linear statistical model, to solving problems in statistical experimental design. In those days, however, no computers (other than human beings) were available for solving large-scale systems of linear equations, and ingenious methods were developed to determine designs for which the estimators could be determined without the need for explicit matrix inversion.

The advantage in the field of experimental design, of course, is that the statistician has control over the values of the explanatory variables of the model. The popular experimental designs developed in the 1930's and 1940's (randomized blocks, Latin squares, fractional factorial, partially-balanced incomplete blocks) were models in which the designer introduced various degrees of orthogonality

into the "design matrix" of the model, so that the equations of estimation could be easily solved.

The point to be recognized here is that, in the face of a strong requirement to develop low-compute estimation procedures in experimental-design applications, tremendous advances were made. The underlying theory was complex (Galois theory and projective geometries), but the computational algorithms that resulted were extremely simple, allowing for the rapid hand-solution of estimation problems containing large numbers of variables.

## 2. Regression-Model Applications

In regression analysis, the statistician does not always have control over the values of the explanatory variables, as is the case in experimental design. Nevertheless, in the period 1920-1950 significant advances were made in the field of developing low-compute procedures for solving regression problems. The Biometrika Tables for Statisticians, which were published in 1953 to "reduce the labour of statistical arithmetic," included tables of orthogonal polynomials, which vastly reduced the amount of computation required to produce estimates of linear contrasts. These tables were of invaluable aid to statisticians in computing estimates of regression model parameters until about 1960, when high-speed digital computers became generally available in research facilities. At that time, it appears that just about all effort directed toward computational simplicity ceased. A digital computer could, in minutes, invert large matrices that simply could not be inverted manually.

It is interesting to note that it was at just about this same time that Kalman and Bucy developed their recursive scheme for determining estimates and predictions for a linear time-series model. The computational requirements of their method were staggering from a manual perspective, but offered no difficulties when tackled by a digital computer. Over the next twenty years, advances were made in the development of more rapid or more precise algorithms for implementing the Kalman-Bucy filter, but these methods accepted the basic linear-model-estimation formulas of the recursive filter as a starting point. With this mind-set, the computational requirements of the Kalman-Bucy filter were never substantially reduced.

In a serendipitous fashion, however, it is interesting to observe that some advances were inadvertently made in the development of low-compute estimators for the non-time-series linear (regression) model. These

developments, by Wald and Bartlett, are described in the paragraphs that follow.

In its simplest form, the general linear statistical model (which forms the basis for modern estimation, prediction, and control algorithms) may be written as:

$$\underline{y} = X'p + \underline{e},$$

where

$\underline{y}$  = vector of observations;

$p$  = vector of parameters;

$X$  = matrix of explanatory variables ("data matrix");

$\underline{e}$  = error vector,

and the prime (') denotes matrix transposition. In correlation/tracking and fusion problems, the form of the equations changes somewhat (e.g., there are "model" and "observation" errors, and the representation is usually in terms of a state vector), but the elementary form given above will serve to illustrate the nature of the estimation algorithms.

The least-squares estimate of the parameter  $p$  is given by:

$$p = (XX')^*xy$$

where the asterisk denotes a conditional (generalized) inverse of a matrix. The key point to note with the least-squares estimate is the fact that it involves matrix products and matrix inversion. (As the model becomes more complex, the number of matrix operations increases.)

In the simple example of regression analysis (fitting a straight line), the above model reduces to:

$$y_i = p_1 + p_2x_i + e_i,$$

where the index  $i$  denotes the  $i$ -th observation ( $i = 1, 2, \dots, n$ ), and the least-squares estimates of the parameters (the intercept and slope of the line) are:

$$p_2 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$



and

$$p_1 = \bar{y} - p_2 \bar{x}$$

where  $\bar{x}$  and  $\bar{y}$  denote the means of the observed x's and y's, respectively.

In 1940, Abraham Wald (the "father" of statistical decision theory and sequential analysis) proposed (Reference 5) a much simpler estimator as an alternative to  $p_2$ :

$$p_2 = (\bar{y}_2 - \bar{y}_1) / (\bar{x}_2 - \bar{x}_1)$$

where  $\bar{x}_1$  and  $\bar{x}_2$  denote the means of the x-values above and below the median (of the x's) and  $\bar{y}_1$  and  $\bar{y}_2$  denote the means of the corresponding y-values (i.e., the y-values associated with the x-values). Wald originally proposed this estimator as a solution to the problem in which both the x variable and the y variable are subject to error. It is interesting to observe, however, that Wald's estimate requires substantially less computation than the least-squares estimator -- 4n additions and one division versus 4n additions, 2n multiplications and one division, where n denotes the number of observations. This represents a reduction in computer time by an order of magnitude.

Wald's estimator possesses the desirable statistical property of consistency (which the least squares estimate does not, in the errors-in-variables problem), but the sampling variance of the estimator is larger than for the least-squares estimate. This inefficiency may be overcome by taking a slightly larger sample, in which case the Wald estimate still has the computational advantage.

In 1949, M. S. Bartlett (Reference 6) modified Wald's estimator by dividing the ranked x-variable into three equal-sized groups, and forming the estimate

$$p_2 = (\bar{y}_3 - \bar{y}_1) / (\bar{x}_3 - \bar{x}_1)$$

where  $\bar{x}_1$ ,  $\bar{y}_1$  are the means corresponding to the low-value group of x's, and  $\bar{x}_3$ ,  $\bar{y}_3$  are the means corresponding to the high-value group of x's. Bartlett's estimator is more efficient (i.e., has lower sampling variance) than Wald's estimator, and requires 1/3 less computation.

In 1958, J. W. Hooper and H. Theil (Reference 7) extended the Wald/Bartlett method of grouping to the case of multiple linear regression (in which there is more than one x-variable). The method was judged somewhat tedious to implement, however, and was essentially abandoned. Note that this was about the time when high-speed digital computers (e.g., the IBM 650) were becoming generally available (at least in the major universities and research centers), and so there was at that time no longer an incentive to prefer the Wald-type estimators to the least-squares estimators on computational grounds. (For the errors-in-variables problem, for which the Wald's estimator was originally developed, a method by J. Durbin, introduced in 1954, was generally adopted as a preferred method. It is more (statistically) efficient than the Wald and Bartlett estimators, but is not relevant to the present problem because it requires the same amount of computation as the least-squares estimation procedure.)

It is interesting to observe that both the Wald and Bartlett estimators may be derived from the formula for the least-squares estimates, by replacing the values of the explanatory variables by +1's and -1's in the case of the Wald estimator, and by +1's, -1's, and 0's in the case of the Bartlett estimator. This procedure is analogous to the procedure of determining an experimental design: the x-values are set at values which enable the equations of estimation to be solved without explicit matrix inversion. (The difference is, of course, that in the case of experimental design, the specified x-values are actually used in the experimentation process, whereas in the regression case, the actual (continuous) x-value is replaced by the simpler discrete value.)

The point to the above is that, with a little ingenuity, it is possible to develop estimators that have drastically reduced computational requirements, over those of the least-squares estimates. The preceding example addresses the simplest situation (one dependent variable, one explanatory variable), with no matrix multiplications or inversions required, and yet a reduction in computational requirements of an order of magnitude were realized. In the general case of several or many variables (e.g., a Kalman filter for a nine-component state vector), involving many matrix multiplications and inversions, the potential for dramatic computational reductions is tremendous.

A review of the statistical literature of the past two decades reveals a fixation with least-squares estimation. To be sure, some new estimators have been introduced (e.g.,

jackknife estimators), but they are computationally similar to the least-squares estimates (requiring matrix multiplications and inversions), and have similar computational efficiencies. It appears that the effort to improve computational efficiency has been conditional on use of the least-squares approach, rather than on centering on novel estimation procedures. The criteria against which the procedures are invariably judged is the error-variance of the minimum-variance linear unbiased estimator. While restriction to this criterion may be reasonable for off-line statistical estimation, it is not reasonable to restrict attention to this single criterion for the large-scale on-line (real-time) estimation situation. We believe that, once the criteria against which the estimator are to be judged are appropriately modified, significant and substantial improvements will follow.

#### E. Parallel with Optimization Theory

In a sense, the situation with respect to the use of the general linear statistical model and the least-squares estimates parallels the situation that existed in the 1950's in the field of optimization theory. At that time, the principal optimization procedure was linear programming. The framework of linear programming did not suit many practical applications, however, and so alternative nonlinear programming methods were sought. In 1963, the Generalized Lagrange Multipliers method was introduced by H. E. Everett. This powerful method produced very fast solutions to very large optimization problems in which the objective function could be nonlinear, non-convex, and discontinuous. The GLM method had a tremendous advantage over previous optimization procedures, in that it did not require the solution of a large-scale system of equations. Unfortunately, the GLM method is restricted to problems in which the objective function is "separable," i.e., may be expressed as a certain sum. Furthermore, it was not possible to guarantee convergence of the method, and in some cases convergence could be slow.

In the late sixties, Fiacco and McCormick promoted the use of "quadratic penalty functions" to solve constrained optimization problems. This approach (called the Sequential Unconstrained Minimization Technique, or SUMT) worked well for a larger class of problems, but in general it did not possess the great speed of the GLM method. Also, the method fails in a fairly wide class of problems in which the Hessian matrix is "ill-conditioned" (not positive definite).

Finally, also in the late 1960's, the two methods were combined into what is now known as the "generalized Lagrangian method." Hestenes and Powell suggested adding a quadratic penalty function to the Lagrangian function, instead of the objective function, as is done in the SUMT method. The generalized Lagrangian method converges fast, and does not exhibit the ill-conditioning that frequently occurs in the original penalty-function method (SUMT).

Thus, in a span of about ten years, a tremendous leap forward was made in solving constrained optimization problems. The interesting fact to note, however, is that whereas the linear programming solution (the simplex method) had a very well-behaved theory associated with it, and was guaranteed to converge in a predetermined number of steps, the generalized Lagrangian methods possessed no such property -- the methods are defined as algorithms (for adjusting the values of Lagrange multipliers), and no definitive statement can be made about the rate of convergence.

The fact remains, however, that by moving out of the very restrictive linear-model framework, tremendous advances were quickly realized. Furthermore, satisfactory solutions were not achieved by "linearizing" nonlinear problems, but by developing heuristic algorithmic procedures, which were demonstrated to work well.

While the problems of estimation, prediction, and control are statistical in nature, they are also optimization problems, and it is reasonable to conjecture that a tremendous advance in speed may be realized by applying the ingenuity and heuristic methods that worked so well in the field of constrained optimization. The current situation -- a near-total dependence on the general linear statistical model -- is analogous to the situation in which the field of constrained optimization theory found itself twenty-five years ago. It would appear that much can be done.

#### F. The Need for a Robust Estimator

During the course of this Phase I study, it was observed that the algorithm that had been synthesized as a candidate "fast" algorithm possessed a remarkable property -- it produced reasonable results when applied to "difficult" problems, in which the classical least-squares method failed catastrophically. The "difficult" problems were ones in which the correlation matrix that had to be inverted in the classical least-squares method was "near-singular" (i.e., had a small determinant). This situation arises whenever

the explanatory variables of the model are highly correlated. In this case, inverting the correlation matrix (which is central to the classical least-squares method) is difficult (in a numerical analysis sense), particularly for large or even moderately large matrices (e.g.,  $m = 10$  explanatory variables). What happens is that roundoff errors (due to the truncated representation of real numbers in the computer) ruin the matrix inversion, and the matrix inversion algorithm fails to produce the desired inverse. The least-squares method in fact fails "catastrophically," in the sense that the produced results are generally totally wrong -- the resultant model produced by the method may predict even worse than the "trivial" model that predicts the mean value of the dependent variable for all values of the independent variables. This problem is ameliorated somewhat by using double precision arithmetic, but it is a particularly troublesome problem in a microcomputing environment, where the computer word length is short (and the precision of computation is low). The problem is a particularly insidious one, since many statistical analysis programs do not warn the user that the method has failed, and if the results do not appear to be patently absurd, they may be accepted as correct.

In addition to the problem of near-singular correlation matrices, the problem of catastrophic failure may arise if a problem is "ill-conditioned," or "ill-specified." This happens, for example, if there is a linear dependency among the independent variables (the  $x$ 's). This situation often arises in social science applications. In a survey questionnaire, for example, the respondent may be asked to select an answer in one of a number (e.g., five) of categories. In the analysis of the survey data, the response for the selected category is coded as "1," and the response for the other categories are coded as "0's." A social science researcher who is not aware of the matrix inversion to be done in a regression analysis may include all five responses as variables in the data base, and in a regression model. These variables are linearly dependent, however, since the sum of all five category responses must equal 1. The presence of this linear dependency in the dependent variable will cause the correlation matrix to be singular. Once again, the least-squares method, if applied to this problem, will fail catastrophically. This case is often not as troublesome as the "near-singular" case, however, since the presence of an exact linear dependency may result in a computed matrix determinant of exactly zero, and some computer programs check for this condition. Because of roundoff errors, however, the computer algorithm may fail to recognize the singularity, compute a nonzero

numerical value for the determinant, and produce a "solution," which, unfortunately, is totally wrong.

This situation is a very real problem. Many users of statistical program packages are not trained in the theory underlying the computations and are unaware of the pitfalls of the least-squares method. Researchers in social science are now trained in university curricula in how to apply the statistical procedure of multiple regression analysis, but they are not expected to know matrix algebra and are generally not trained in it as part of their introduction to statistics. As a professional consulting statistician, the author of this report has been retained on more than one occasion to "explain" absurd results obtained because of the problem of linear dependencies in regression analysis applications. In technical terms, the classical least-squares method is not "robust" with respect to the presence of linear dependencies or near-dependencies (high correlations) in the explanatory variables.

Statistical theory can handle the presence of linear dependencies, if they are recognized. In such cases, the inverse of the correlation matrix is replaced by a "generalized inverse" or "conditional inverse." Although this theory is generally taught to graduate statistics majors, however, it is not known to most data analysts. Moreover, most of the major statistical software packages do not offer this capability.

The present study began as an attempt to determine the feasibility of finding fast algorithms for estimation, prediction, and control. In the course of the study, an algorithm was developed that did not fail catastrophically in "difficult" problems. Upon observing this, it was decided to explore the performance of the algorithm in ill-conditioned problems containing linear dependencies. The method works well in such cases. The significance of this result could be very great from a commercial viewpoint, in view of the extensive use of statistical multiple regression analysis. Moreover, virtually all multivariable statistical analysis procedures involve matrix inversion. All of these methods are subject to catastrophic failure, and are candidates for application of a method that does not require matrix inversion.

### III. PROJECT APPROACH

#### A. Summary of Approach

The approach proposed for this study consisted of four major steps:

1. Development of Criteria for Comparing Alternative Estimation Algorithms
2. Generation of Test Cases
3. Synthesis of Candidate Algorithms
4. Comparison of the Performance of Candidate Algorithms to the Performance of the Classical Least-Squares Algorithm

Each of these steps is described in detail in the following subsections.

#### B. Development of Criteria for Comparing Alternative Estimation Algorithms

In order to assess the performance of alternative procedures for estimation, prediction and control, it is necessary to identify a number of quantitative, measurable descriptors of performance. Alternative procedures may differ in a number of respects, such as processing speed and accuracy. It is desirable to identify a set of performance measures, or criteria, which afford a relatively comprehensive description of algorithm performance, and yet is not overly redundant.

It is noted that the performance measures that are appropriate for an algorithm may vary, depending on whether the algorithm is intended for use for estimation, or for prediction, or for control. In an estimation problem, attention centers on estimating the model parameters ( $b$ ,  $\sigma$ , SIG, SIGMA) as closely as possible. In a prediction or control problem, attention centers on using the model for predicting a new value of  $y$  corresponding to a specified set of  $x$ -values. How close the parameter estimates are to the true values is of secondary interest in this situation. (In

a "prediction" problem, the x's are either passively observed or actively controlled; in a "control" problem, the x's are actively controlled. The problem of predicting the state of the economy is essentially a prediction problem; the problems of controlling a steel production process or directing a "smart" bomb to a target are examples of "control" problems.) The intended application of the model -- estimation, prediction or control -- should influence the sample design for the collection of the data from which the model is to be estimated. For example, if a model is to be used to predict how the dependent variable (y) will respond to forced changes in the independent variables (x's), then the data should correspond to the case in which forced changes are made in the x's. A model developed from passively-observed x's is not appropriate for predicting how the system will respond if forced changes are made in the x's (although this is often done, with disappointing results!).

After consideration of the various uses (estimation, prediction and control) of the models under study, and of the various properties of algorithms that may be of interest to model developers, it was decided that the following seven concepts characterized the algorithm performance in reasonably comprehensive fashion:

1. Computer running speed
2. Computer storage requirements
3. Precision of the parameter estimates
4. Bias of the parameter estimates
5. Mean squared error of the parameter estimates
6. Precision of model-based predictions
7. Numerical stability of the algorithm

Having decided on these concepts as comprising a relatively comprehensive characterization of the performance of an algorithm, it remained to determine quantitative measures of each concept. These concepts and their associated measures are described in the paragraphs that follow. Note that, although the preceding concepts were identified as part of the Phase I report, it was not possible with the Phase I resources to numerically determine values for all of the measures. That numerical determination can be accomplished



with some additional programming, and should be done as one of the first steps in the Phase II study.

### 1. Computer Running Speed

The primary motivation for the proposed study was to determine the feasibility of determining estimation, prediction and control algorithms that had faster running times than the classical least-squares method. The measure of speed that was used in this study was the total elapsed time required to read the data from the file on which it was stored, compute the parameter estimates, and print out the results. The time required to "set up" the run (e.g., specify the data file name, number of parameters, number of observations, etc.) was not included, since this time consists mainly of the time of the human operator to enter data through the microcomputer keyboard, and does not reflect algorithm performance.

All of the computer programming and processing on this project was done using a Radio Shack Model II microcomputer using a TRSDOS Version 2.0 operating system. This microcomputer utilizes a 4 MHz Z80 microprocessor, and has 64 kilobytes of direct access memory. The programming was done in FORTRAN II, using a Microsoft (R) FORTRAN compiler to produce the object code. Unfortunately, the available Microsoft FORTRAN did not permit access to the system timer, and so the timing was done manually (external to the program), not automatically (internal to the program). The running times presented in this report are hence approximate. In addition to algorithm processing time, they include the time required to print the results. The amount of printed output for the various methods is comparable. For the simpler cases examined, the algorithm processing time was very short compared to the print time, and so the total measured running time is not an accurate reflection of the processing time. For the more difficult cases, the algorithm processing time is large relative to the print time, and so the measured running time is a relatively valid indicator of algorithm processing time.

It is recognized that the speed measure used in this study is very crude, particularly for problems having a small number of explanatory variables (since the processing time for these problems is small compared to the data access and printout time). In Phase II, we propose to develop an assembly-language timer that can be called internal to the program, so that accurate measures of processing time can be determined.

The time measurements for the iterative algorithm were biased high in many cases, because the algorithm was forced to make at least ten iterations even if convergence had already occurred. In Phase II, a test for convergence should be developed, so that the algorithm stops when convergence occurs.

## 2. Computer Storage Requirements

The amount of direct-access computer memory required to implement an algorithm is a concern, since it determines how "large" a problem can be handled with available memory, or how much memory is required to handle a problem of a given size. The size of an estimation problem is determined primarily by two factors -- the number of observations and the number of explanatory variables (we are speaking here only the univariate case, in which there is but a single dependent variable). The classical least-squares method has an advantage in that the observations may be read and processed one-at-a-time, and do not need to be stored simultaneously in memory. The particular "fast algorithm" that was considered in this study required that all of the data be stored in memory.

Although the required amount of direct-access memory required depends on the problem size, the computer programs developed in this project did not dynamically adjust the memory requirement to the size of the problem. Instead, the "dimensions" of the program variables were set to allow for storage of all observations and variables, corresponding to the largest problem analyzed -- 100 observations and ten independent variables (and a single dependent variable). Under these conditions, the core requirements of the classical least-squares algorithm and the particular "fast algorithm" investigated in this study were as follows:

Least-squares algorithm: 20,984 bytes

Fast algorithm: 21,759 bytes,

i.e., the core requirements are approximately the same.

## 3. Precision of the Parameter Estimates

In an estimation problem, it is desired to determine ("estimate") the values of the model parameters as "closely" or "accurately" as possible. The model parameters are  $b$ ,  $\beta$ ,  $\epsilon$ , SIGMA, and SIG. With regard to estimation, however, we are generally interested only in  $b_0$ ,  $b_1$ , and SIG, not in  $\epsilon$  or SIGMA. The reason for this limitation is that the

parameters  $\Sigma$  and SIGMA are not generally of concern in an estimation problem once the x's are known. The usual objective in a data analysis is to estimate the values of  $b_0$ ,  $b$ , and SIG, or to predict the value of a new y given specified values of the x's. While the fact that the x's are random variables can affect some analysis procedures (e.g., hypothesis testing), the least-squares estimates are the same whether the x's are fixed numbers or random variables. For this reason, we shall restrict attention only to measures of precision of the estimates of  $b_0$ ,  $b$ , and SIGMA. This restriction applies also to the next two performance concepts considered (bias and accuracy).

Accuracy is usually reflected in two concepts -- precision and bias. The "precision" (or reliability) of an estimator (i.e., an estimation formula or algorithm) refers to the degree of reproducibility, or variation, of the estimates, if repeated data samples were selected and the estimator used to determine the estimate value for each sample. The usual measure of precision is the standard deviation (square root of the variance, or second central moment). The "bias" of a parameter estimator is the amount of systematic error in the estimate, measured as the difference between the average value obtained for the parameter estimate in repeated sampling and the true value of the parameter. (Note: in the preceding discussion, we have used the distinction that an "estimate" is a numerical value that represents our guess as to the value of the parameter, whereas the term "estimator" refers to the formula or algorithm for producing that numerical value. This distinction in usage of these terms is not strict, either in the field of statistics or in this report.)

In the present study, it was possible to estimate the precision of the parameter estimates for the classical least-squares method, but project resources did not permit the estimation of precision of the parameter estimates for the fast algorithm studied. Closed-form mathematical formulas are available for estimating the precision of the least-squares estimates, but such is not the case for estimating the precision of the fast algorithm estimates. Instead, the precision has to be measured empirically, by selecting many independent data samples and computing the standard deviation of the parameter estimates over these samples. Although it was not possible to implement this procedure in Phase I (because of resource limitations), the estimated standard deviation of the parameter estimates is considered to be an important measure of algorithm performance, and this procedure should be implemented in the Phase II effort.

While the precision and bias of the parameter estimates are of high concern in estimation problems, they are of secondary concern in prediction and control problems. In the latter types of problems, it is the accuracy of the prediction that is of primary concern. It is possible to have a model in which the parameter estimates are not very accurate, and yet the accuracy of the predictions based on that model is comparable to those obtained from a model in which the parameter estimates are substantially more accurate. (This may be the case, for example, if the explanatory variables are highly correlated.)

The precision of most statistical estimates increases as the sample size (number of data observations) increases, usually by the factor  $1/\sqrt{n}$  (for the standard deviation).

#### 4. Bias of Parameter Estimates

The bias of an estimate is the expected value (over repeated data samples) of the difference between the expected value of the parameter estimate and the true value of the parameter. The bias of an estimator is of greater concern in estimation problems than in prediction and control problems.

As was the case in the measurement of precision, Phase I project resources did not permit the generation an analysis of a large number of data samples (corresponding to the same model parameters) to estimate the bias.

The bias of an estimator may or may not decrease as the sample size increases.

#### 5. Mean Squared Error

The mean squared error is a measure of accuracy -- i.e., it is a "combined" measure of precision and bias. The definition of the mean squared error of a parameter estimator is:

$$\text{Mean Squared Error} = \text{variance} + \text{bias}^2$$

or

$$E(\hat{p} - p)^2 = E(\hat{p} - E(\hat{p}))^2 + (E(\hat{p}) - p)^2,$$

where  $p$  is the parameter value and  $\hat{p}$  is the estimate, i.e., the mean squared error is the expected value of the square

of the difference between the estimated parameter value and the true parameter value, over repeated data samples.

Once more, Phase I resources did not permit the numerical determination of the mean squared error, but this can be done in Phase II.

#### 6. Precision of Model-Based Prediction

The three preceding measures are concerned with measurement of the "closeness" of the parameter estimates to the true values, or to the amount of variability in the parameter estimates in repeated data samples. For prediction problems, the primary concern is how close the model-based predictions (of  $y$ , the dependent variable) corresponding to a specified set of  $x$ -values will be to a newly-sampled  $y$  corresponding to those  $x$ -values. An interesting fact is that there can be fairly substantial errors in the estimation of the model parameters, and yet predictions based on the (erroneous) model may be almost as good as those based on a much-more-nearly-correct model. In some applications (e.g., econometric modelling), attention centers very much on estimation of the model parameters. In other applications, the parameter values are of incidental interest -- all that matters is the error of prediction.

A standard indicator of the prediction error is the estimated variance or standard deviation of the model "residuals," or error terms (differences between the observed  $y$ -values and those predicted by the model). This is not the same as the standard deviation of the prediction error for a particular set of  $x$ 's, which depends also on the specified values of the  $x$ 's. The standard deviation of the prediction error is proportional to the standard deviation of the residuals, however, and so the latter is a good indicator of the predictive ability of the model.

Another indicator of the predictive power of a model is the reduction in variance between predictions based on the trivial model that predicts that each new  $y$  will equal the mean (of the  $y$ 's), and the variance of the predictions based on the estimated model. (This measure is valid only for application of the model to predict  $y$ -values from  $x$ -values that were produced in the same fashion -- e.g., passively observed, forcibly changed -- as were those from which the model parameters were derived. Also, its use assumes that the theoretical variance of  $y$  is finite, which is often not the case for time series data.) The standard error of the residuals is best estimated from a sample other than that from which the model parameters were derived, but as long as

the number of parameters is very small relative to the number of observations, it is common practice to use the same data set for both purposes.

The reduction in variance for predictions based on the "mean model" compared to predictions based on the estimated model is called the "coefficient of determination." It can be defined as:

$$CD = 1 - \frac{\text{(variance of predictions using "mean model")}}{\text{(variance of predictions using estimated model)}}$$

The coefficient of determination was determined both for the classical least-squares model and the fast algorithm algorithm studied in this project.

## 6. Numerical Stability of the Algorithm

The three preceding measures of algorithm performance are appropriate in most cases. In some situations, however, an algorithm may fail to operate as intended because of computer roundoff errors or because of an intrinsic weakness in the algorithm, such as a failure to converge to an answer close to the desired answer, or a failure to converge at all. The length of time required for convergence to a desired answer is considered under performance measure 1 (computer processing speed), if the process converges. (An additional measure of performance that is of interest in studying the performance of an iterative algorithm is the "rate" of convergence, or the number of iterations required for convergence.)

Some of the test cases examined in the project present difficulties, for both the classical least-squares and the fast algorithm. In some such cases, the algorithm may fail catastrophically, i.e., produce totally wrong results. In other such cases, the algorithm may produce an answer (i.e., set of parameter estimates) that is not very close to the correct answer (i.e., the true parameter values). Those cases are noted, with an indication of the nature of the failure. For example, the classical least-squares method may fail because of an inability to invert a near-singular matrix. Or, a "fast" algorithm may converge very slowly.

## 2. Generation of Test Cases

The performance of an estimation algorithm may vary, depending on the nature of the data set to which it is applied. We proposed to develop a set of test cases to which candidate algorithms could be applied, and to measure

their performance relative to each test case. Although there is an infinite variety of test cases that might be considered, it was possible in the present project to generate and analyze only a few. It was decided to examine sixteen cases in all. The nature of these test cases was described in Section I of this report. The test cases differ in terms of the number of variables included in the model, and in terms of the complexity of the model, as reflected in the covariance matrix of the explanatory variables (x's).

The general model considered in this study was of the following form:

$$y_j = b_0 + \underline{x}_j' \underline{b} + e_j \quad j=1,2,\dots,n$$

or

$$\underline{y} = b_0 \underline{1} + \underline{X}' \underline{b} + \underline{e}$$

where

$y_j$  = dependent variable

$\underline{y}' = (y_1, y_2, \dots, y_n)$  = vector of all  
observed y's

$\underline{x}_j' = (x_{1j}, x_{2j}, \dots, x_{mj})$  = vector of inde-  
pendent variables corresponding to the  
j-th observation

$\underline{X} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$  = data matrix

$\underline{e}' = (e_1, e_2, \dots, e_n)$  = vector of model error  
terms corresponding to all observations

$b_0$  = y-intercept

$\underline{b}' = (b_1, b_2, \dots, b_m)$  = vector of regression  
coefficients

$\text{var}(\underline{x}_j) = \Sigma \text{SIGMA}^2$

$\text{var}(\underline{e}) = I \text{SIG}^2$

$\underline{1}' = (1,1,\dots,1) = n\text{-component vector of all 1's}$

where  $\Sigma$  is an  $m \times m$  covariance matrix and  $I$  is an  $n \times n$  identity matrix. The various test cases correspond to four different values of  $m$  ( $m = 1, 3, 6, \text{ and } 10$ ), various values of SIGMA and SIG, and various covariance matrices  $\Sigma$ .

The  $x$ 's are random variables that are generated as follows. First,  $m$  independent random variables,  $f_1, f_2, \dots, f_m$ , normally distributed with mean 0 and variance SIGMA, are sampled. An  $m \times m$  coefficient matrix,  $C$ , is then specified:

$$C = \{c_{ij}\} .$$

The  $x$ 's are computed as

$$\underline{x} = C \underline{f} .$$

The covariance matrix of  $\underline{x}$  is

$$\text{var} \underline{x} = CC' \text{SIGMA}^2 = \Sigma \text{SIGMA}^2 ,$$

where  $\Sigma = CC'$ .

To generate  $y_j$  corresponding to a specified  $\underline{x}$  (say,  $\underline{x}_j$ ), a standardized normal deviate,  $e_j$ , is sampled, multiplied by SIG, and the result added to  $b_0 + \underline{x}'\underline{b}$ :

$$y_j = b_0 + \underline{x}_j' \underline{b} + \text{SIG } e_j .$$

From the preceding, it is seen that the variance matrix is not explicitly specified. Instead, it is the coefficient matrix  $C$  that is actually specified. The matrix  $\Sigma$  has the value  $CC'$ .

The coefficient matrices and values of  $b_0$ ,  $\underline{b}$ , SIG, and SIGMA for the various test cases are specified in the computer printouts of Appendix B. The values of the corresponding covariance matrices  $\Sigma$  are also specified in those printouts, as part of a larger matrix,  $\Sigma_A$ . The covariance matrix,  $\Sigma_A$ , specified in the printouts includes an additional row and column, which contains the variance and covariances of  $y$  with each  $x$ , i.e.,

$$\begin{aligned} \Sigma_A &= \text{var}(y, \underline{x}') \\ &= \text{var}(b_0 \underline{1} + \underline{x}' \underline{b} + \underline{e}, \underline{x}') \end{aligned}$$



$$= \begin{pmatrix} \underline{b}' \Sigma \underline{b} + \text{SIG}^2 & \underline{b}' \Sigma \\ \Sigma \underline{b} & \Sigma \end{pmatrix}$$

The normally distributed random numbers were generated by the algorithm

$$f = \text{SIGMA} \sum_{i=1}^{12} u_i$$

where  $u_1, u_2, \dots, u_{12}$ , are a sequence of independent uniformly distributed random numbers. The number  $u_i$  was generated by the multiplicative congruential method, specified for the 16-bit Radio Shack Model II microprocessor by the algorithm:

$$\begin{aligned} \text{IX}_{\text{tent}} &= \text{IX}_{\text{old}} (2^{16} + 3) = \text{IX}_{\text{old}} 259 \\ \text{IX}_{\text{new}} &= \begin{cases} \text{IX}_{\text{tent}} & \text{if } \text{IX}_{\text{tent}} \geq 0 \\ \text{IX}_{\text{tent}} + (2^{16}-1)+1 = \text{IX}_{\text{tent}} + 32767+1 & \text{if } \text{IX}_{\text{tent}} < 0 \end{cases} \end{aligned}$$

$$u_i = \text{IX}_{\text{new}} (2^{-15}) = \text{IX}_{\text{new}} (.30517578\text{E-}4)$$

$$\text{IX}_{\text{old}} = \text{IX}_{\text{new}}$$

The multiplicative congruential method for generating pseudo-random numbers is described in Reference 10.

The rationale for the specification of the test cases is as follows. For the case of a single independent variable ( $m = 1$ ), the four test cases involved four different values of the ratio SIG/SIGMA: .1, .5, 1.0, and 5. The case with SIG/SIGMA = .1 is easiest, since there is a substantial amount of variation in  $x$  and a small model error term.

For the cases with  $m$  greater than one, the value of SIG/SIGMA was set equal to .1, and the complexity of the variance matrix was varied. For the first subcase, was set equal to the identity matrix, i.e., the  $x$ 's were orthogonal (uncorrelated). For the second subcase, the coefficient matrix used to generate the  $x$ 's (and hence  $\Sigma$ ) was specified to correspond to a low-to-moderate degree of correlation among the  $x$ 's. For the third subcase, the coefficient matrix was specified to produce a moderate-to-high degree of correlation among the  $x$ 's. In the final subcase, a linear dependency was introduced among

the x's: the last x was set equal to the sum of the two preceding x's.

The values of the regression coefficient,  $b$ , were chosen to make the estimation difficult. The correlations among the x's were all positive, and so the regression coefficients were specified to be plus ones and minus ones. This caused convergence of the iterative algorithm to be slow, because each time an adjustment was made in one direction for a particular coefficient, a reverse adjustment would be required on the next iteration for coefficients having the opposite sign.

### 3. Synthesis of a Candidate Algorithm

In order to demonstrate the feasibility of the proposed concept of developing a fast estimation algorithm that could be used as an alternative to the classical least-squares algorithm, it was proposed to synthesize one or more alternative algorithms, and to examine their performance. The class of algorithms that was synthesized in this project are extensions of the Wald and Bartlett estimators, extended to the case of more than one independent variable.

The algorithm is iterative. At the k-th iteration, the residuals,  $e_j^{k-1}$ , from the preceding iteration are regressed on a particular independent variable,  $x_i$ , using either the Wald or Bartlett method:

$$e_j^{k-1} = b_0^k + b_1^k x_{ij} + e_j^k$$

where  $b_0^k$  and  $b_1^k$  are the Wald or Bartlett estimates. In the regression process of determining  $b_0^k$  and  $b_1^k$ , of course, the value of  $e_j^k$  is unknown. Once the values of  $b_0^k$  and  $b_1^k$  are determined, the value of  $e_j^k$  may be computed as follows:

$$e_j^k = e_j^{k-1} - b_0^k - b_1^k x_{ij}$$

The process cycles through all of the x's in order. That is, the first iteration regresses y on  $x_1$ . The second iteration regresses the residuals from the first regression on  $x_2$ . The third iteration regresses the residuals from the second regression on  $x_3$ , and so on. If there are m independent variables, then the (m+1)-st regression begins over again, and regresses the residuals from the m-th regression on  $x_1$ , and so on.

The preceding iterative procedure is conjectured to produce consistent parameter estimates of the model parameters

(i.e., estimates that converge in probability to the true parameter values as the sample size increases). Determination of whether this is true, or of conditions under which it is true, should be addressed in Phase II.

In the present study, several modifications of the preceding procedure were examined. For example, a "stepwise" procedure was considered, in which the regression at each iteration was performed on the independent variable,  $x_r$ , which resulted in the greatest reduction in the criterion:

$$b_r IQ_r$$

where  $IQ_r$  denotes the interquartile range of  $x_r$ , and  $b_r$  denotes the Wald regression coefficient of the residuals at that iteration on  $x_r$ . Another approach considered was to regress the residuals separately on all  $x$ 's, and then to adjust all  $m$  regression coefficients by a specified fraction ("stepsize") of the indicated adjustment. No modified algorithm was found, however, that performed faster than the one specified above.

Both the Wald and Bartlett procedures were applied to the test cases for which  $m = 1$ , but only the Wald procedure was applied to the other test cases. The iterative procedure described above will be referred to as an "iterative Wald-Bartlett" estimation method.

#### E. Comparison of the Performance of the Candidate Algorithm to the Performance of the Classical Least-Squares Algorithm

Each of the sixteen test cases was analyzed using both the classical least-squares (CLS) and the iterative Wald-Bartlett (IWB) algorithms. The performance of each case was determined, using the processing speed and the coefficient of determination as the measures of performance, or observing whether the method failed catastrophically. The closeness of the residual standard error to the parameter SIG was also noted. The results of these simulations are described in the next section of this report.

## IV. SIMULATION RESULTS

### A. Test Cases Involving a Single Independent Variable

The four test cases with  $m = 1$  independent variable were analyzed using both the Wald and Bartlett estimates, and the classical least-squares estimator. In general, all cases ran so fast that speed differences could not be reliably determined among the methods. All methods produced parameter estimates close to the true values, and all produced similar values for the residual standard error and the coefficient of determination. A summary of the results of the four cases with  $m = 1$ , and all of the other test cases as well, is presented in Figure 2.

### B. Test Cases Involving Three Independent Variables

For the test cases with  $m = 3$  independent variables, the data were analyzed with both the iterative Wald-Bartlett (IWB) and classical least-squares (CLS) methods. The results are presented in Figure 2. Both the CLS and IWB methods appear to be of comparable speed, in cases in which the CLS method does not fail. The CLS method fails if the correlations among the  $x$ 's are high, or if there is a linear dependency among the  $x$ 's. The IWB method succeeds in these cases, but does not converge very fast. The slow convergence is probably due to the fact that the regression coefficients were purposely specified to cause slow convergence.

### C. Test Cases Involving Six Explanatory Variables

The results for the cases involving six explanatory variables are as follows. For cases in which the CLS method does not fail, it is somewhat faster than the IWB method, for the difficult estimation problems represented by the test cases. For cases in which the correlations among the  $x$ 's is high, the CLS method fails. In these cases, the IWB method is slow to converge. Once again, the slow convergence is probably due to the "pathological" specification of the regression coefficients.

### D. Test Cases Involving Ten Independent Variables

Figure 2. Algorithm Performance on Sixteen Test Cases

Case No.	<u>Model Parameters</u>					<u>Performance Measures</u>				
	m	Corr	SIG SIGMA	SIG	CD	Est Meth	Time (sec)	RSE	CD	Fail
1	1	-	.1	1.0	.9901	CLS	40	1.12	.9864	
						WALD	35	1.12	.9862	
						BART	35	1.12	.9862	
2	1	-	.5	1.0	.8000	CLS	40	1.12	.7554	
						WALD	35	1.12	.7528	
						BART	35	1.12	.7528	
3	1	-	1.0	1.0	.5000	CLS	40	1.12	.4547	
						WALD	35	1.12	.4489	
						BART	35	1.12	.4489	
4	1	-	5.0	1.0	.0385	CLS	40	1.12	.0535	
						WALD	35	1.12	.0436	
						BART	35	1.12	.0435	
5	3	0	.1	1.0	.9967	CLS	65	1.15	.9951	
						IWB	70	1.15	.9950	
6	3	Low	.1	1.0	.9967	CLS	65	1.15	.9951	
						IWB	55	1.16	.9949	
7	3	High	.1	1.0	.9955	CLS	65	6.20	-	Fail
						IWB	200	1.20	.9928	
8	3	LDEP	.1	1.0	.9975	CLS	65	1.52	-	Fail
						IWB	70	1.15	.9961	
9	6	0	.1	1.0	.9983	CLS	105	.99	.9987	
						IWB	130	1.01	.9986	
10	6	Low	.1	1.0	.9985	CLS	105	.99	.9990	
						IWB	155	1.15	.9986	
11	6	High	.1	1.0	.9989	CLS	105	107.	-	Fail
						IWB	840	1.13	.9990	
12	6	LDEP	.1	1.0	.9986	CLS	105	8.68	-	Fail
						IWB	70	1.00	.9987	

Figure 2 (cont.). Algorithm Performance on Sixteen Test Cases

Case No.	<u>Model Parameters</u>				<u>Performance Measures</u>					
	m	Corr	<u>SIG</u> SIGMA	SIG	CD	Est Meth	Time (sec)	RSE	CD	Fail
13	10	0	.1	1.0	.9990	CLS	180	1.10	.9989	
						IWB	200	1.28	.9988	
14	10	Low	.1	1.0	.9988	CLS	180	1.10	.9988	
						IWB	300	1.17	.9984	
15	10	High	.1	1.0	.9997	CLS	180	610.	-	Fail
						IWB	1800	2.03	.9991	
16	10	LDEP	.1	1.0	.9991	CLS	-	-	-	Fail
						IWB	185	1.13	.9988	

Legend:

m: number of explanatory variables (x's)

Corr: degree of correlation in x's

CD: coefficient of determination

Est Meth: CLS: classical least-squares method

WALD: Wald's method

BART: Bartlett's method

IWB: iterative Wald-Bartlett method

Time: time to read data, process data, and print results

RSE: residual standard error

Fail: the CLS method failed to invert the correlation matrix, or the correlation matrix was singular, and the solution was incorrect

The results for the test cases involving ten independent variables were as follows. For cases in which the CLS method succeeded, it is faster than the IWB method for the test cases studied. For the cases in which the correlation among the x's is high, the CLS method fails, and the IYB method is slow. In the worst case (case 15), the process was terminated after 1800 seconds. At that point, the model had reached a solution close to the correct solution, but not correct -- the residual standard error was 2.03, compared to the true value of 1.0.

In summary, the IYB method appears to be comparable in speed to the CLS method for problems with a small number of explanatory variables or low correlations among the x's. For difficult problems (with high correlations among the x's), the CLS method fails. The IYB method converged in all of the difficult cases but one.

## V. CONCLUSIONS AND RECOMMENDATIONS

### A. Conclusions

This study has demonstrated the feasibility of developing an algorithm for estimating the parameters of a linear statistical model and making predictions based on the estimated model, that is comparable in speed to the classical least-squares method for problems of low to moderate difficulty, and is definitely more robust, in the sense that it is less subject to catastrophic failure. The feasibility was established by synthesizing an algorithm -- an extension of the estimation procedures of Wald and Bartlett -- which often outperformed the classical method.

The availability of a fast, robust estimation procedure would be beneficial to both military and nonmilitary applications. In a military context, there is a growing need for faster estimation procedures -- current procedures cannot accomplish tracking of large numbers of objects in real-time, or accomplish large-scale sensor exploitation in real-time. Also, embedded-processor estimation algorithms are non-interactive (i.e., must perform without human intervention), and are potentially subject to catastrophic failure if based on the classical least-squares procedure, if highly correlated data are entered into the data input stream.

In both military and commercial/industrial applications, there is a requirement for "fail-safe" estimation algorithms. The classical least-squares algorithm that is currently in use is not fail-safe. For problems involving moderate or large numbers of variables, computer roundoff errors can ruin the estimates, and the user may be totally unaware of the failure. The danger of this occurrence is particularly strong in microcomputers having short word lengths (e.g., 16 bit microprocessors), and is a problem even for 32-bit machines. Furthermore, many persons using statistical software packages are not aware that linear dependencies in the variables can cause the complete failure of the methods. Once again, roundoff errors may obscure the problem, so that the user has no reason to believe that the estimation algorithm failed.



While the present project has demonstrated the feasibility of developing fast, robust algorithms, it has not accomplished the ultimate goal of developing such an algorithm. The algorithm that was synthesized in this project is not considered to be a final solution to this problem -- it does not outperform the classical least-squares method in every case, and its theoretical properties (e.g., convergence, consistency of estimates) are unknown. Continuation of the effort to develop improved (fast, robust) estimators will require substantial additional effort. The present study suggests, however, that the potential for success of such an effort is high.

#### B. Recommendations

It is recommended that a Phase II study be conducted, oriented toward the goal of developing improved algorithms for estimation, prediction and control. It is further recommended that the Phase II effort should address the following tasks.

1. Extension of the results of the Phase I study, to include analysis of a wider range of test cases, and measurement of the full set of performance measures identified in this study.
2. Development of a broader class of algorithms, additional to the iterative Wald-Bartlett method developed in this Phase I effort.
3. Extension of the algorithm to consider estimation problems additional to the multiple linear regression problem considered in the Phase I effort. Consideration should be given to developing fast, robust methods for the full range of problems currently addressed by least-squares methods, such as multivariate analysis of variance and time series analysis procedures (currently done by Box-Jenkins, Kalman filter, and state-space methods).
4. Analyze the theoretical numerical and statistical properties of candidate algorithms, such as convergence conditions and consistency.

If successful, it is expected that there would be a substantial military and non-military demand for a statistical estimation computer program package based on the improved methods. With respect to military applications, such methods offer the potential for fast, fail-safe

processing of tracking and sensor exploitation data. With respect to non-military applications, the methods are much more "user-friendly" than the least-squares method, in that the user would be protected from catastrophic failures, and would not need to understand matrix algebra concepts such as linear dependencies, singularities, and numerical stability problems in matrix inversion, to be assured of successful application of the procedures. In view of the large number of persons involved in data analysis, and the growing use of microcomputers, the development of such methods is considered to be a very significant contribution to the field of data analysis.

The determination that the iterative Wald-Bartlett method avoided the catastrophic failure problem of the classical least-squares method was serendipitously discovered during the course of the Phase I investigation. Since this property of the algorithm is judged to be probably more important than speed in many applications, it is recommended that the title of the Phase II study be changed from "Fast Algorithms for Estimation, Prediction and Control," to "Improved Algorithms for Estimation, Prediction and Control."

Reduction of the danger of obtaining wrong answers from regression analyses represents an area of potentially great benefit, to a wide class of data analysts. The original concept of this study was to develop fast algorithms, primarily for real-time applications such as tracking, sensor exploitation, or industrial process control. Those applications, while important, concern relatively few data analysts. The discovery of the robustness of the iterative Wald-Bartlett algorithm, however, could have substantial impact for a wide class of data analysts. For example, a typical logistics application involves the determination of parametric cost estimating relationships. These relationships are estimated by linear regression analysis. Since the models developed are empirical in nature, they involve the analysis of a large number of cost-related variables. The presence of a linear dependency in the data, or the occurrence of a roundoff-error-caused matrix inversion failure caused by high correlation among some of the explanatory variables, could produce incorrect results. The availability of "fail-safe" algorithms would benefit this and many other similar applications. In view of the substantial amount of funds expended by the Office of Naval Research on parametric cost analysis and other data analysis, the benefits of improved estimation methods would be substantial.

## REFERENCES

1. Kailath, T., "A View of Three Decades of Linear Filtering Theory," IEEE Transactions on Information Theory, Vol. IT-20, No. 2, March 1974, pp. 146-181.
2. Aggarwal, A. K., "Numerical Methods in Least-Squares Parameter Estimation," The Journal of the Astronautical Sciences, Vol. XXXVIII, No. 2, April-June 1982, pp. 181-189.
3. Buzbee, B. L., N. Metropolis, and D. H. Sharp, "Supercomputing Frontiers: What the Future Holds," Defense Science 2002+, Vol. 3, No. 2, April 1984, pp. 35-40.
4. Tucker, Jonathan B., "Biochips: Can Molecules Compute?" High Technology, Vol. 4, No. 2, February 1984, pp. 36-47.
5. Wald, A., "The Fitting of Straight Lines if Both Variables are Subject to Error," Annals of Mathematical Statistics, Vol. 11, 1940, pp. 284-300.
6. Bartlett, M. S., "Fitting a Straight Line When Both Variables Are Subject to Error," Biometrics, Vol. 5, 1949, pp. 207-212.
7. Hooper, J. W. and H. Theil, "The Extension of Wald's Method of Fitting Straight Lines to Multiple Regression," Review of the International Statistical Institute, Vol. 26, 1958, pp. 37-47.
8. Cooley, W. W., and Lohnes, P. R., Multivariate Procedures for the Behavioral Sciences, Wiley, 1962.
9. System/360 Scientific Subroutine Package, Version III, Programmer's Manual, Program Number 360A-CM-03X, Fifth Edition, GH20-0205-4, International Business Machines Corporation, White Plains, New York, August 1970.
10. Newman, T. G., and P. L. Odell, The Generation of Random Variates, Charles Griffin & Company, London, 1971.

Appendix A  
Computer Program Source Code Listings

```
00100 C PROGRAM SIMULA
00200 C THIS PROGRAM GENERATES THE VECTOR RANDOM VARIABLE X=C*F
00300 C AND THE SCALAR RANDOM VARIABLE Y=B0+B*X+SIG*E, WHERE
00400 C F IS AN MX1 RANDOM VECTOR, X IS AN MX1 VECTOR,
00500 C C IS AN MXM MATRIX, E(F)=0, VAR(F)=I*SIGMA**2,
00600 C I IS AN MXM IDENTITY MATRIX
00700 C B0 IS A SCALAR PARAMETER, B IS AN MX1 PARAMETER VECTOR
00800 C E IS A SCALAR N(0,1) ERROR TERM
00900 C N OBSERVATIONS ARE GENERATED, AND WRITTEN TO FILE FNAME1
01000 C INPUT M,C(M,M), SIGMA, N, B0,B(N),SIG
01100 C OUTPUT THE N OBSERVATIONS (1 Y-COMPONENT AND M X-COMPONENTS)
01200 C DIMENSION C(11,11), CT(11,11),VARX(11,11), X(10),F(10)
01300 C DIMENSION SX(11),B(10)
01400 C DOUBLE PRECISION FNAME1(2)
01500 C DATA LUN1,LUN2,LUN6/1,2,6/
01600 C LUN1:TERMINAL, LUN2:PRINTER, LUN6:DATA FILE FOR OUTPUT
01700 C WRITE(LUN2,6)
01800 C 6 FORMAT(15H0PROGRAM SIMULA,/
01900 C 142H GENERATES MULTIVARIABLE UNIVARIATE SAMPLE)
02000 C WRITE(LUN1,9)
02100 C 9 FORMAT(53H ENTER FILE NAME, 1-16 ALPHA CHARS, LAST CHAR BLANK: )
02200 C READ(LUN1,8)FNAME1
02300 C 8 FORMAT(2A8)
02400 C WRITE(LUN2,7)FNAME1
02500 C 7 FORMAT(12H0FILE NAME: ,2A8)
02600 C CALL OPEN(LUN6,FNAME1,80)
02700 C MXDIM=10
02800 C 5 CONTINUE
02900 C WRITE(LUN1,10)
03000 C 10 FORMAT(20H0INPUT DIMENSION, M:)
03100 C READ(LUN1,20)M
03200 C 20 FORMAT(I2)
03300 C IF(M.GT.MXDIM)GO TO 5
03400 C WRITE(LUN2,31)M
03500 C 31 FORMAT(17H0DIMENSION (M) = ,I2)
03600 C WRITE(LUN2,56)
03700 C 56 FORMAT(26H0COEFFICIENT MATRIX (C)...)
03800 C WRITE(LUN1,42)
03900 C 42 FORMAT(51H0INPUT 1 TO USE IDENTITY COEFT MATRIX, 0 OTHERWISE:)
04000 C READ(LUN1,43)IOPT
04100 C 43 FORMAT(I1)
04200 C IF(IOPT.LE.0)GO TO 44
04300 C DO 46 I=1,M
04400 C DO 47 J=1,M
04500 C 47 C(I,J)=0.0
04600 C C(I,I)=1.0
04700 C 46 CONTINUE
04800 C WRITE(LUN2,48)
04900 C 48 FORMAT(37H COEFFT MATRIX IS IDENTITY MATRIX)
05000 C GO TO 45
05100 C 44 CONTINUE
05200 C DO 100 I=1,M
05300 C WRITE(LUN1,40),I,M
05400 C 40 FORMAT(10H INPUT ROW,I3,21H OF COEFFT MATRIX, C./
05500 C 15H M = ,I2,21H VALUES, 8 AT A TIME:/)
05600 C READ(LUN1,50)(C(I,J),J=1,M)
05700 C WRITE(LUN2,57)(C(I,J),J=1,M)
05800 C 57 FORMAT(1X,8F10.4)
```

```

05900      50 FORMAT(8F10.4)
06000      55 FORMAT(8F10.4)
06100     100 CONTINUE
06200      45 CONTINUE
06300      WRITE(LUN1,60)
06400      60 FORMAT(14H INPUT SIGMA: )
06500      READ(LUN1,70)SIGMA
06600      70 FORMAT(F10.4)
06700      WRITE(LUN2,81)SIGMA
06800      81 FORMAT(49H0SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = ,F10.4)
06900      WRITE(LUN1,710)
07000     710 FORMAT(23H INPUT B0 (INTERCEPT): )
07100      READ(LUN1,70)B0
07200      WRITE(LUN2,711)B0
07300     711 FORMAT(18H0B0 (INTERCEPT) = ,F10.4)
07400      WRITE(LUN1,712)M
07500     712 FORMAT(27H INPUT PARAMETER VECTOR, B./
07600      15H M = ,I2,21H VALUES, 8 AT A TIME:/)
07700      READ(LUN1,50)(B(I),I=1,M)
07800      WRITE(LUN2,727)
07900     727 FORMAT(24H0PARAMETER VECTOR (B)...)
08000      WRITE(LUN2,713)(B(I),I=1,M)
08100     713 FORMAT(1X,8F10.4)
08200      WRITE(LUN1,714)
08300     714 FORMAT(12H INPUT SIG: )
08400      READ(LUN1,50)SIG
08500      WRITE(LUN2,715)SIG
08600     715 FORMAT(29H0SIG (MODEL ERROR STD DEV) = ,F10.4)
08700      WRITE(LUN1,90)
08800      90 FORMAT(36H INPUT NO OF OBSNS, N, TO SIMULATE: )
08900      READ(LUN1,35)N
09000      35 FORMAT(I3)
09100      WRITE(LUN2,36)N
09200      36 FORMAT(14H0NO OF OBSNS =,I4)
09300      WRITE(LUN1,91)
09400      91 FORMAT(33H INPUT NO OF RANDOM NOS TO SKIP: )
09500      READ(LUN1,35)NSKIP
09600      WRITE(LUN2,92)NSKIP
09700      92 FORMAT(6H SKIP ,I3,31H RANDOM NOS PRIOR TO SIMULATION)
09800      DO 95 I=1,NSKIP
09900      CALL RANDN(TEMP)
10000     95 CONTINUE
10100      WRITE(LUN1,552)
10200     552 FORMAT(36H INPUT NO OF OBSERVATIONS TO PRINT: )
10300      READ(LUN1,35)NPRT
10400     C   COMPUTE VARIANCE MATRIX OF X
10500      CALL ATRAN(C,M,M,CT)
10600      CALL AB(C,CT,M,M,M,VARX)
10700      SIGS=SIGMA**2
10800      CALL ASCAL(VARX,SIGS,M,M,VARX)
10900     C   COMPUTE ROW AND COLUMN CORRESPONDING TO Y
11000      CALL AX(VARX,B,M,M,X)
11100      MP=M+1
11200      DO 325 I=1,M
11300      VARX(MP,I)=X(I)
11400      VARX(I,MP)=X(I)
11500     325 CONTINUE
11600      CALL CROSP(B,X,M,CP)
11700      VARX(MP,MP)=CP+SIG**2
11800      WRITE(LUN2,58)

```

```

1900      58 FORMAT(27H0VARIANCE MATRIX OF X, Y...)
12000     DO 330 I=1,MP
12100     SX(I)=SQRT(VARX(I,I))
12200     WRITE(LUN2,57)(VARX(I,J),J=1,MP)
12300     330 CONTINUE
12400     VY=VARX(MP,MP)
12500     VRES=SIG**2
12600     CD=1.-VRES/VY
12700     WRITE(LUN2,640)CD
12800     640 FORMAT(32H0COEFFICIENT OF DETERMINATION = ,F10.4)
12900     C COMPUTE CORRELATION MATRIX
13000     DO 610 I=1,MP
13100     DO 610 J=1,MP
13200     610 VARX(I,J)=VARX(I,J)/(SX(I)*SX(J))
13300     WRITE(LUN2,54)
13400     54 FORMAT(30H0CORRELATION MATRIX OF X, Y...)
13500     DO 630 I=1,MP
13600     WRITE(LUN2,57)(VARX(I,J),J=1,MP)
13700     630 CONTINUE
13800     WRITE(LUN2,550)
13900     550 FORMAT(16H0OBSERVATIONS.../
14000     124H DEPENDENT VARIABLE LAST)
14100     WRITE(LUN2,553)NPRT
14200     553 FORMAT(8H (FIRST ,I3,15H OBSNS PRINTED))
14300     KPRT=0
14400     DO 500 K=1,N
14500     DO 400 I=1,M
14600     CALL RANDN(TEMP)
14700     F(I)=SIGMA*TEMP
14800     400 CONTINUE
14900     CALL AX(C,F,M,M,X)
15000     C COMPUTE Y
15100     Y=B0
15200     DO 720 I=1,M
15300     720 Y=Y+B(I)*X(I)
15400     CALL RANDN(TEMP)
15500     Y=Y+SIG*TEMP
15600     WRITE(LUN6,55)(X(J),J=1,M),Y
15700     KPRT=KPRT+1
15800     IF(KPRT.GT.NPRT)GO TO 501
15900     WRITE(LUN2,551)K,(X(J),J=1,M),Y
16000     501 CONTINUE
16100     500 CONTINUE
16200     551 FORMAT(9H OBSN NO ,I4,2H: ,2(/1X,8F10.4))
16300     END

```

Program Unit Length=08ED (2285) Bytes  
Data Area Length=0B6C (2924) Bytes

Subroutines Referenced:

\$I3	\$I1	\$I0
\$QRT	\$INIT	\$W2
\$ND	\$R2	OPEN
\$M9	\$L1	\$T1
\$RANDN	ATRAN	AB
\$SCAL	\$EA	AX
\$CROSP	\$AB	\$DB
\$SB	\$NB	\$MB
\$EX		

Variables:

X	0001"	CT	01E5"	VARX	03C9"
B	05AD"	F	05D5"	SX	05FD"
UN2	0629"	FNAM1	0651"	LUN1	0661"
M	0663"	LUN6	0665"	MXDIM	06FE"
I	0719"	T:000002		IOPT	0790"
I	0796"	J	0798"	T:000000	079A"
I	010000	SIGMA	0846"	B0	08A9"
IG	0948"	N	099D"	NSKIP	09DF"
TEMP	0A11"	NPRT	0A3E"	SIGS	0A4C"
P	0A5C"	T:020000	0A5E"	T:030000	0A60"
P	0A62"	VY	0A8A"	VRES	0A8E"
CD	0A92"	T:000001	0AC1"	KPRT	0B3B"
K	0B3D"	Y	0B45"		

Labels:

SL	0006'	6L	0667"	9L	06AA"
L	06E4"	7L	06E9"	5L	005E'
10L	0700"	20L	071B"	31L	0720"
5L	0739"	42L	0758"	43L	0792"
4L	0188'	46L	0169'	47L	0105'
48L	079E"	45L	0250'	100L	0240'
49L	07C8"	50L	0823"	57L	0818"
5L	082B"	60L	0833"	70L	084A"
61L	0851"	710L	088D"	711L	08AD"
712L	08CA"	727L	090F"	713L	092C"
14L	0937"	715L	094C"	90L	0974"
5L	099F"	36L	09A3"	91L	09B9"
92L	09E1"	95L	0A15'	552L	0A15"
25L	04F0'	58L	0A6A"	330L	05D1'
40L	0A96"	610L	064F'	54L	0AC5"
630L	071F'	550L	0AE8"	553L	0B19"
500L	08C6'	400L	0788'	720L	07B6'
01L	08C6'	551L	0B49"		

```
16400 SUBROUTINE ATRAN(A,M,N,B)
16500 C THIS SUBROUTINE COMPUTES B=A-TRANPOSE
16600 C INPUT MATRIX A(M,N),M,N
16700 C OUTPUT MATRIX B
16800 C A AND B MUST BE STORED IN DIFFERENT LOCATIONS
16900 C DIMENSION A(M,N),B(N,M)
17000 DIMENSION A(11,11),B(11,11)
17100 DO 100 I=1,M
17200 DO 110 J=1,N
17300 B(J,I)=A(I,J)
17400 110 CONTINUE
17500 100 CONTINUE
17600 RETURN
17700 END
```

Program Unit Length=0090 (144) Bytes  
Data Area Length=0013 (19) Bytes

Subroutines Referenced:

\$AT \$M9 \$L1  
TI



Variables:

0001"	M	0003"	N	0005"
0007"	I	0009"	J	000B"
T:000000	000D"	T:010000	000F"	T:020000
				0011"

Labels:

000L 007B' 110L 0067'

```

17800      SUBROUTINE AB(A,B,L,M,N,C)
17900 C    THIS SUBROUTINE COMPUTES MATRIX PRODUCT C(L,N)=A(L,M)*B(M,N)
18000 C    INPUT MATRIX A(L,M),MATRIX B(M,N),L,M,N
18100 C    OUTPUT MATRIX C(L,N)
18200 C    DIMENSION A(L,M),B(M,N),C(L,N)
18300 C    DIMENSION A(11,11),B(11,11),C(11,11)
18400      DO 100 I=1,L
18500      DO 110 J=1,N
18600      TEMP=0.0
18700      DO 120 K=1,M
18800      TEMP=TEMP+A(I,K)*B(K,J)
18900 120 CONTINUE
19000      C(I,J)=TEMP
19100 110 CONTINUE
19200 100 CONTINUE
19300      RETURN
19400      END
    
```

Program Unit Length=00F5 (245) Bytes

Data Area Length=001D (29) Bytes

Subroutines Referenced:

AT	\$L1	\$T1
SM9	\$MB	\$AB

Variables:

0001"	B	0003"	L	0005"
0007"	N	0009"	C	000B"
000D"	J	000F"	TEMP	0011"
0015"	T:000000	0017"	T:010000	0019"
020000	001B"			

Labels:

000L 00DC' 110L 00C8' 120L 0085'

```

19500      SUBROUTINE ASCAL(A,SCAL,M,N,B)
19600 C    THIS SUBROUTINE MULTIPLIES MATRIX A TIMES SCALAR SCAL
19700 C    B=A*SCAL
19800 C    MATRICES A AND B MAY BE STORED IN THE SAME LOCATION
19900 C    INPUT MATRIX A(M,N), SCALARS SCAL, M,N
20000 C    OUTPUT MATRIX B(N,M)
20100 C    DIMENSION A(M,N),B(M,N)
20200 C    DIMENSION A(11,11),B(11,11)
20300      DO 100 I=1,M
20400      DO 110 J=1,N
20500      B(I,J)=SCAL*A(I,J)
    
```

```

0600 110 CONTINUE
20700 100 CONTINUE
20800 RETURN
20900 END

```

```

Program Unit Length=0081 (129) Bytes
Data Area Length=0015 (21) Bytes

```

Subroutines Referenced:

```

$AT          $M9          $L1
$MB          $T1

```

Variables:

```

A          0001"          SCAL          0003"          M          0005"
          0007"          B          0009"          I          000B"
          000D"          T:000000          000F"          T:010000          0011"
T:020000          0013"

```

Labels:

```

100L 006C'          110L 0058'

```

```

21000          SUBROUTINE AX(A,X,M,N,Y)
21100 C THIS SUBROUTINE MULTIPLIES MATRIX A(MXN) TIMES VECTOR X(NX1)
21200 C I.E., COMPUTES Y = A*X
21300 C INPUT MATRIX A(M,N), VECTOR X(N)
21400 C OUTPUT MATRIX Y(M)
21500 C X AND Y MUST BE IN DIFFERENT STORAGE LOCATIONS
21600 C DIMENSION A(M,N),X(1),Y(1)
21700 DIMENSION A(11,11),X(1),Y(1)
21800 DO 100 I=1,M
21900 TEMP=0.0
22000 DO 110 J=1,N
22100 TEMP=TEMP+A(I,J)*X(J)
22200 110 CONTINUE
22300 Y(I)=TEMP
22400 100 CONTINUE
22500 RETURN
22600 END

```

```

Program Unit Length=00B7 (183) Bytes
Data Area Length=0017 (23) Bytes

```

Subroutines Referenced:

```

$AT          $L1          $T1
$M9          $MB          $AB

```

Variables:

```

          0001"          X          0003"          M          0005"
          0007"          Y          0009"          I          000B"
TEMP          000D"          J          0011"          T:000000          0013"
T:010000          0015"

```

Labels:

```

100L 009E'          110L 006D'

```

```

22700      SUBROUTINE CROSP(X,Y,M,CP)
22800 C    THIS SUBROUTINE COMPUTES THE CROSSPRODUCT OF VECTORS X AND Y
22900 C    INPUT VECTORS X(M),Y(M), DIMENSION M
23000 C    OUTPUT CROSSPRODUCT CP
23100      DIMENSION X(1),Y(1)
23200      CP=0.0
23300      DO 100 I=1,M
23400      CP=CP+X(I)*Y(I)
23500      100 CONTINUE
23600      RETURN
23700      END

```

Program Unit Length=006B (107) Bytes  
Data Area Length=000F (15) Bytes

Subroutines Referenced:

\$AT	\$L1	\$T1
\$MB	\$AB	

Variables:

CP	0001"	Y	0003"	M	0005"	
T:010000	0007"	I	0009"	T:000000	000B"	
	000D"					

Labels:

00L 0052'

```

23800      SUBROUTINE RANDN(V)
23900 C    THIS SUBROUTINE GENERATES STANDARD NORMAL DEVIATES
24000 C    BY SUMMING 12 UNIFORMLY DISTRIBUTED RANDOM NUMBERS
24100 C    OUTPUT V IS THE RANDOM NUMBER
24200      A=0.0
24300      DO 50 I=1,12
24400      CALL RANDU(Y)
24500      A=A+Y
24600      50 CONTINUE
24700      V=A-6.
24800      RETURN
24900      END

```

Program Unit Length=0055 (85) Bytes  
Data Area Length=000D (13) Bytes

Subroutines Referenced:

\$L1	\$T1	RANDU
\$AB	\$SB	

Variables:

Y	0001"	A	0003"	I	0007"
	0009"				

Labels:

00L 002D'

```
25000      SUBROUTINE RANDU(YFL)
25100      C  THIS SUBROUTINE GENERATES UNIFORMLY DISTRIBUTED RANDOM NUMBERS
25200      C  BY THE MULTIPLICATIVE CONGRUENTIAL ("RESIDUE") METHOD
25300      C  OUTPUT YFL IS THE RANDOM NUMBER
25400      DATA IX/32769/
25500      IY=IX*259
25600      IF(IY)5,6,6
25700      5 IY=IY+32767+1
25800      6 YFL=IY
25900      YFL=YFL*.30517578E-4
26000      IX=IY
26100      RETURN
26200      END
```

Program Unit Length=004C (76) Bytes  
Data Area Length=0007 (7) Bytes

Subroutines Referenced:

```
LM9          $CA          $T1
$LL          $MB
```

Variables:

```
YFL          0001"          IX          0003"          IY          0005"
```

Labels:

```
5L          0019'          6L          0023'
```

Created: 26-Nov-80

```

00100 C PROGRAM MREG
00200 C PERFORMS MULTIPLE LINEAR REGRESSION ANALYSIS
00300 DIMENSION XBAR(11),STD(11),RX(121),R(66),B(11),D(11),T(10)
00400 DIMENSION LL(10),MM(10),SB(10),RY(10),ISAVE(11)
00500 DIMENSION X(1),ANS(11),NAME(40)
00600 DIMENSION RXI(10,10),RXJ(10,10),RPROD(10,10)
00700 C DOUBLE PRECISION RX
00800 DOUBLE PRECISION FNAM1(2)
00900 DATA LUN1,LUN2,LUN6/1,2,6/
01000 C LUN1:TERMINAL; LUN2:PRINTER; LUN6:DATA FILE FOR INPUT
01100 WRITE(LUN2,80)
01200 80 FORMAT(13H0PROGRAM MREG,/
01300 136H MULTIPLE LINEAR REGRESSION ANALYSIS)
01400 WRITE(LUN1,90)
01500 90 FORMAT(53H ENTER FILE NAME, 1-16 ALPHA CHARS, LAST CHAR BLANK: )
01600 READ(LUN1,95)FNAM1
01700 95 FORMAT(2A8)
01800 CALL OPEN(LUN6,FNAM1,80)
01900 WRITE(LUN2,209)FNAM1
02000 209 FORMAT(17H0DATA FILE NAME: ,2A8)
02100 WRITE(LUN1,201)
02200 201 FORMAT(46H ENTER NO OF OBS, NO OF VARS, INDEX OF DEP VAR/
02300 133H NO OF INDEP VARS IN REGRESSION: )
02400 READ(LUN1,18)NOBS,NVAR,IDEP,NIND
02500 18 FORMAT(16I5)
02600 WRITE(LUN2,16)
02700 16 FORMAT(40H0NO OF OBS, NO OF VARS, INDEX OF DEP VAR/
02800 149H IN REGRESSION, NO OF INDEP VARS IN REGRESSION...)
02900 IO=0
03000 WRITE(LUN2,13)NOBS,NVAR,IDEP,NIND
03100 13 FORMAT(1X,8I10)
03200 WRITE(LUN1,19)
03300 19 FORMAT(44H ENTER INDICES OF INDEP VARS, IN ASC ORDER: /)
03400 READ(LUN1,18)(ISAVE(I),I=1,NIND)
03500 WRITE(LUN2,215)
03600 215 FORMAT(39H INDICES OF INDEP VARS IN REGRESSION...)
03700 WRITE(LUN2,13)(ISAVE(I),I=1,NIND)
03800 WRITE(LUN1,221)
03900 221 FORMAT(33H ENTER MAX NO OF OBSNS TO PRINT: )
04000 READ(LUN1,222)NPRTM
04100 222 FORMAT(I4)
04200 NPRTM=MIN0(NPRTM,NOBS)
04300 NOPR=0
04400 WRITE(LUN1,29)
04500 29 FORMAT(45H ENTER 1 TO PERFORM REGRESSION, 0 TO SUPPRESS/)
04600 READ(LUN1,203)IREG
04700 203 FORMAT(I1)
04800 WRITE(LUN2,12)
04900 12 FORMAT(8H0DATA...)
05000 WRITE(LUN2,223)NPRTM
05100 223 FORMAT(8H (FIRST ,I4,7H OBSNS))
05200 CALL CORRE(NOBS,NVAR,IO,X,XBAR,STD,RX,R,D,B,T,NPRTM,NOPR)
05300 WRITE(LUN2,32)
05400 32 FORMAT(27H0MEANS FOR ALL VARIABLES...)
05500 WRITE(LUN2,20)(XBAR(I),I=1,NVAR)
05600 WRITE(LUN2,34)
05700 34 FORMAT(41H0STANDARD DEVIATIONS FOR ALL VARIABLES...)
05800 WRITE(LUN2,20)(STD(I),I=1,NVAR)

```

```

05900      WRITE(LUN2,28)
06000 28  FORMAT(22H0CORRELATION MATRIX...)
06100      DO 130 I=1,NVAR
06200          N1=I*(I-1)/2+1
06300          N2=N1+I-1
06400 130 WRITE(LUN2,20)(R(J),J=N1,N2)
06500      20  FORMAT(1X,8F10.4)
06600          IF(IREG)125,125,126
06700 126  CONTINUE
06800          CALL ORDER(NVAR,R,IDEF,NIND,ISAVE,RX,RY)
06900          DO 104 I=1,NIND
07000              K1=NIND*(I-1)
07100              DO 104 J=1,NIND
07200                  K=K1+J
07300 104  RXI(I,J)=RX(K)
07400          CALL MINV(RX,NIND,DD,LL,MM)
07500          WRITE(LUN2,213)DD
07600 213  FORMAT(15H0DETERMINANT = ,E10.4)
07700          IF(DD)211,211,212
07800 211  WRITE(LUN2,214)
07900 214  FORMAT(49H0ZERO DETERMINANT, REGRESSION CANNOT BE PERFORMED)
08000          GO TO 100
08100 212  CONTINUE
08200          DO 108 I=1,NIND
08300              K1=NIND*(I-1)
08400              DO 108 J=1,NIND
08500                  K=K1+J
08600 108  RXJ(I,J)=RX(K)
08700          WRITE(LUN2,25)
08800      25  FORMAT(61H0INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSI
08900          1ON...)
09000          DO 105 I=1,NIND
09100              N1=NIND*(I-1)+1
09200              N2=N1+I-1
09300 105  WRITE(LUN2,20)(RX(J),J=N1,N2)
09400          DO 107 I=1,NIND
09500              DO 107 J=1,I
09600                  RP=0
09700                  DO 106 K=1,NIND
09800 106  RP=RP+RXI(I,K)*RXJ(K,J)
09900 107  RPROD(I,J)=RP
10000          WRITE(LUN2,24)
10100      24  FORMAT(33H0PRODUCT OF MATRIX AND INVERSE...,/
10200          152H SHOULD BE IDENTITY MATRIX. IF NOT, ROUND OFF ERRORS/
10300          252H HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.)
10400          DO 103 I=1,NIND
10500              N1=1
10600              N2=I
10700 103  WRITE(LUN2,20)(RPROD(I,J),J=N1,N2)
10800          CALL MULTR(NOBS,NIND,XBAR,STD,D,RX,RY,ISAVE,B,SB,T,ANS)
10900          DO 102 I=1,NIND
11000              N1=NIND*(I-1)+1
11100              N2=N1+I-1
11200              JJ=0
11300              DO 102 J=N1,N2
11400                  JJ=JJ+1
11500 102  RX(J)=RX(J)/(STD(I)*STD(JJ)*(NOBS-1))
11600          WRITE(LUN2,26)
11700      26  FORMAT(67H0INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN RE
11800          1GRESSION...)

```

```

1900      DO 101 I=1,NIND
12000     N1=NIND*(I-1)+1
12100     N2=N1+I-1
12200 101 WRITE(LUN2,33)(RX(J),J=N1,N2)
12300     33 FORMAT(1X,8(2X,E10.4))
12400     WRITE(LUN2,35)IDEP
12500     35 FORMAT(34H0INDEX OF DEP VAR IN REGRESSION = ,I10)
12600     WRITE(LUN2,36)
12700     36 FORMAT(39H INDICES OF INDEP VARS IN REGRESSION...)
12800     WRITE(LUN2,21)(ISAVE(I),I=1,NIND)
12900     21 FORMAT(1X,8I10)
13000     WRITE(LUN2,38)
13100     38 FORMAT(27H REGRESSION COEFFICIENTS...)
13200     WRITE(LUN2,20)(B(I),I=1,NIND)
13300     WRITE(LUN2,40)
13400     40 FORMAT(12H T-VALUES...)
13500     WRITE(LUN2,20)(T(I),I=1,NIND)
13600     DO 37 I=1,NIND
13700     37 T(I)=B(I)/T(I)
13800     WRITE(LUN2,41)
13900     41 FORMAT(25H STD DEVS OF REG COEFS...)
14000     WRITE(LUN2,20)(T(I),I=1,NIND)
14100     WRITE(LUN2,42)ANS(1)
14200     42 FORMAT(12H INTERCEPT =,F10.4)
14300     WRITE(LUN2,47)ANS(3)
14400     47 FORMAT(27H RESIDUAL STANDARD ERROR = ,F10.4)
14500     WRITE(LUN2,43)ANS(2)
14600     43 FORMAT(42H SAMPLE MULTIPLE CORRELATION COEFFICIENT =,F10.4)
14700     WRITE(LUN2,45)ANS(11)
14800     45 FORMAT(38H SAMPLE COEFFICIENT OF DETERMINATION =,F10.4)
14900     WRITE(LUN2,48)ANS(4)
15000     48 FORMAT(50H SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR =,
15100     1F14.4)
15200     WRITE(LUN2,49)ANS(5)
15300     49 FORMAT(42H DEGREES OF FREEDOM ASSOCIATED WITH SSAR =,F10.4)
15400     WRITE(LUN2,50)ANS(6)
15500     50 FORMAT(22H MEAN SQUARE OF SSAR =,F14.4)
15600     WRITE(LUN2,51)ANS(7)
15700     51 FORMAT(53H SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SDDR =,
15800     1F14.4)
15900     WRITE(LUN2,52)ANS(8)
16000     52 FORMAT(42H DEGREES OF FREEDOM ASSOCIATED WITH SDDR =,F10.4)
16100     WRITE(LUN2,53)ANS(9)
16200     53 FORMAT(22H MEAN SQUARE OF SDDR =,F14.4)
16300     WRITE(LUN2,54)ANS(10)
16400     54 FORMAT(10H F-VALUE =,F14.4)
16500 125 CONTINUE
16600 100 CONTINUE
16700     END

```

Program Unit Length=094E (2382) Bytes  
Data Area Length=10D5 (4309) Bytes

Subroutines Referenced:

MINO	\$I3	\$I1
PIO	\$INIT	\$W2
IND	\$R2	OPEN
CORRE	\$M9	\$D9
ORDER	\$L1	\$T1

INV		\$CA		\$MB	
\$AB		MULTR		\$DB	
EX					

Variables:

BAR	0001"	STD	002D"	RX	0059"
R	023D"	B	0345"	D	0371"
T	039D"	LL	03C5"	MM	03D9"
B	03ED"	RY	0415"	ISAVE	043D"
	0453"	ANS	0457"	NAME	0483"
RXI	04D3"	RXJ	0663"	RPROD	07F3"
NAM1	0983"	LUN1	0993"	LUN2	0995"
LUN6	0997"	NOBS	0A85"	NVAR	0A87"
I	0A89"	NIND	0A8B"	IO	0AFB"
I	0B3E"	T:000000	0B40"	NPRTM	0B94"
OPR	0B9A"	IREG	0BCF"	N1	0C79"
2	0C7B"	J	0C7D"	K1	0C94"
K	0C96"	T:010000	0C98"	DD	0C9A"
P	0D36"	T:020000	0D3A"	JJ	0DE7"
T:000001	0DE9"				

Labels:

55L	0006'	80L	0999"	90L	09D4"
95L	0A0E"	209L	0A13"	201L	0A2D"
8L	0A93"	16L	0A99"	13L	0B03"
9L	0B0C"	215L	0B42"	221L	0B6E"
222L	0B96"	29L	0B9C"	203L	0BD1"
2L	0BD5"	223L	0BE1"	32L	0C10"
DL	0C7F"	34L	0C30"	28L	0C5E"
130L	027A'	125L	093F'	126L	02D2'
104L	0303'	213L	0CA4"	211L	0381'
12L	0390'	214L	0CBE"	100L	093F'
108L	03B5'	25L	0CF4"	105L	042F'
107L	04F5'	106L	0497'	24L	0D3C"
03L	055D'	102L	05F6'	26L	0DED"
101L	068F'	33L	0E35"	35L	0E45"
36L	0E70"	21L	0E9C"	38L	0EA5"
0L	0EC5"	37L	07C7'	41L	0ED6"
2L	0EF4"	47L	0F0B"	43L	0F31"
45L	0F66"	48L	0F97"	49L	0FD4"
50L	1009"	51L	102A"	52L	106A"
3L	109F"	54L	10C0"		

```

16800      SUBROUTINE DATA(M,D,NPRTM,NOPR)
16900      DIMENSION D(1)
17000      LUN2=2
17100      LUN6=6
17200      READ(LUN6,10)(D(I),I=1,M)
17300      10  FORMAT(8F10.4)
17400      NOPR=NOPR+1
17500      IF(NOPR.GT.NPRTM)GO TO 12
17600      WRITE(LUN2,11)(D(I),I=1,M)
17700      12  CONTINUE
17800      11  FORMAT(1X,8F10.4)
17900      RETURN
18000      END

```

Program Unit Length=00D5 (213) Bytes



Data Area Length=0025 (37) Bytes

Subroutines Referenced:

\$I1	\$AT	\$R2
\$ND	\$W2	

Variables:

0001"	D	0003"	NPRTM	0005"
0007"	LUN2	0009"	LUN6	000B"
000D"	T:000000	000F"	T:000002	0019"

Labels:

10L	0011"	12L	00D0'	11L	001A"
-----	-------	-----	-------	-----	-------

```

18100 SUBROUTINE ORDER(M,R,NDEP,K,ISAVE,RX,RY)
18200 DIMENSION R(1),ISAVE(1),RX(1),RY(1)
18300 C DOUBLE PRECISION RX
18400 MM=0
18500 DO 130 J=1,K
18600 L2=ISAVE(J)
18700 IF(NDEP-L2)122,123,123
18800 122 L=NDEP+(L2*L2-L2)/2
18900 GO TO 125
19000 123 L=L2+(NDEP*NDEP-NDEP)/2
19100 125 RY(J)=R(L)
19200 DO 130 I=1,K
19300 L1=ISAVE(I)
19400 IF(L1-L2)127,128,128
19500 127 L=L1+(L2*L2-L2)/2
19600 GO TO 129
19700 128 L=L2+(L1*L1-L1)/2
19800 129 MM=MM+1
19900 130 RX(MM)=R(L)
20000 ISAVE(K+1)=NDEP
20100 RETURN
20200 END

```

Program Unit Length=01AC (428) Bytes  
Data Area Length=001D (29) Bytes

Subroutines Referenced:

\$AT	\$M9	\$D9
\$L1	\$T1	

Variables:

0001"	R	0003"	NDEP	0005"
0007"	ISAVE	0009"	RX	000B"
000D"	MM	000F"	J	0011"
0013"	T:000000	0015"	L	0017"
0019"	L1	001B"		

Labels:

130L	013D'	122L	004A'	123L	0072'
125L	009B'	127L	00F1'	128L	0115'

```

20300      SUBROUTINE ARRAY(MODE,I,J,N,M,S,D)
20400      DIMENSION S(1),D(1)
20500      NI=N-I
20600      IF(MODE-1)100,100,120
20700      100 IJ=I*J+1
20800          NM=N*J+1
20900          DO 110 K=1,J
21000              NM=NM-NI
21100          DO 110 L=1,I
21200              IJ=IJ-1
21300              NM=NM-1
21400      110 D(NM)=S(IJ)
21500          GO TO 140
21600      120 IJ=0
21700          NM=0
21800          DO 130 K=1,J
21900          DO 125 K=1,I
22000              IJ=IJ+1
22100              NM=NM+1
22200      125 S(IJ)=D(NM)
22300      130 NM=NM+NI
22400      140 RETURN
22500      END
    
```

Program Unit Length=0168 (360) Bytes  
 Data Area Length=001B (27) Bytes

Subroutines Referenced:

\$AT	\$M9	\$L1
\$F1		

Variables:

DDE	0001"	I	0003"	J	0005"	
	0007"	M	0009"	S	000B"	
D	000D"	NI	000F"	T:000000		0011"
	0013"	NM	0015"	K	0017"	
	0019"					

Labels:

100L	003D'	120L	00E6'	110L	0093'
140L	0167'	130L	0148'	125L	010C'

```

22600      SUBROUTINE MULTR(N,K,XBAR,STD,D,RX,RY,ISAVE,B,SB,T,ANS)
22700      DIMENSION XBAR(1),STD(1),D(1),RX(1),RY(1),ISAVE(1),B(1),SB(1)
22800      DIMENSION T(1),ANS(1)
22900      C DOUBLE PRECISION RX
23000          MM=K+1
23100          DO 100 J=1,K
23200      100 B(J)=0.0
23300          DO 110 J=1,K
23400              Ll=K*(J-1)
23500              DO 110 I=1,K
23600                  L=Ll+I
23700      110 B(J)=B(J)+RY(I)*RX(L)
23800          RM=0.0
    
```

```

3900      BO=0.0
24000     L1=ISAVE(MM)
24100     DO 120 I=1,K
24200     RM=RM+B(I)*RY(I)
24300     L=ISAVE(I)
24400     B(I)=B(I)*STD(L1)/STD(L)
24500     120 BO=BO+B(I)*XBAR(L)
24600     BO=XBAR(L1)-BO
24700     SSAR=RM*D(L1)
24800     SSSDR=D(L1)-SSAR
24900     FN=N-K-1
25000     SY=SSDR/FN
25100     DO 130 J=1,K
25200     L1=K*(J-1)+J
25300     L=ISAVE(J)
25400     C 125 SB(J)=DSQRT(DABS((RX(L1)/D(L))*SY))
25500     125 SB(J)=SQRT(ABS((RX(L1)/D(L))*SY))
25600     130 T(J)=B(J)/SB(J)
25700     135 SY=SQRT(ABS(SY))
25800     FK=K
25900     SSARM=SSAR/FK
26000     SSSDRM=SSDR/FN
26100     F=SSARM/SSDRM
26200     ANS(11)=RM
26300     122 RM=SQRT(ABS(RM))
26400     ANS(1)=BO
26500     ANS(2)=RM
26600     ANS(3)=SY
26700     ANS(4)=SSAR
26800     ANS(5)=FK
26900     ANS(6)=SSARM
27000     ANS(7)=SSDR
27100     ANS(8)=FN
27200     ANS(9)=SSDRM
27300     ANS(10)=F
27400     RETURN
27500     END

```

Program Unit Length=04B0 (1200) Bytes  
Data Area Length=0065 (101) Bytes

Subroutines Referenced:

DSQRT	ABS	\$AT
\$L1	\$T1	\$M9
\$MB	\$AB	\$DB
\$SB	\$CA	

Variables:

	0001"	K	0003"	XBAR	0005"	
STD	0007"	D	0009"	RX	000B"	
RY	000D"	ISAVE	000F"	B	0011"	
B	0013"	T	0015"	ANS	0017"	
MM	0019"	J	001B"	T:000000		001D"
L1	001F"	I	0021"	L	0023"	
X:010000	0025"	RM	0027"	BO	002B"	
X:020000	002F"	T:030000		SSAR	0033"	
SSDR	0037"	FN	003B"	SY	003F"	
X:000001	0043"	T:010001	0047"	FK	004B"	

SARM 004F"                      SSDRM 0053"                      F 0057"  
 T:040000                      005B"                      T:050000                      005D"                      T:060000                      005F"  
 070000                      0061"                      T:080000                      0063"

Labels:

0L	0020'	110L	007A'	120L	0194'
130L	02F3'	125L	0290'	135L	0343'
122L	03B8'				

```

27600      SUBROUTINE MINV(A,N,DD,L,M)
27700      DIMENSION A(1),L(1),M(1)
27800      C    DOUBLE PRECISION A,D,BIGA,HOLD
27900          D=1.0
28000          NK=-N
28100          DO 80 K=1,N
28200            NK=NK+N
28300            L(K)=K
28400            M(K)=K
28500            KK=NK+K
28600            BIGA=A(KK)
28700            DO 20 J=K,N
28800              IZ=N*(J-1)
28900              DO 20 I=K,N
29000                IJ=IZ+1
29100                C 10 IF(DABS(BIGA)-DABS(A(IJ)))15,20,20
29200                  10 IF(ABS(BIGA)-ABS(A(IJ)))15,20,20
29300                  15 ABIGA=A(IJ)
29400                    L(K)=I
29500                    M(K)=J
29600                20 CONTINUE
29700                  J=L(K)
29800                  IF(J-K)35,35,25
29900                25 KI=K-N
30000                  DO 30 I=1,N
30100                    KI=KI+N
30200                    HOLD=-A(KI)
30300                    JI=KI-K+J
30400                    A(KI)=A(JI)
30500                30 A(JI)=HOLD
30600                35 I=M(K)
30700                  IF(I-K)45,45,38
30800                38 JP=N*(I-1)
30900                  DO 40 J=1,N
31000                    JK=NK+J
31100                    JI=JP+J
31200                    HOLD=-A(JK)
31300                    A(JK)=A(JI)
31400                40 A(JI)=HOLD
31500                45 IF(BIGA)48,46,48
31600                46 D=0.0
31700                  GO TO 150
31800                48 DO 55 I=1,N
31900                  IF(I-K)50,55,50
32000                50 IK=NK+I
32100                  A(IK)=A(IK)/(-BIGA)
32200                55 CONTINUE
32300                  DO 65 I=1,N
32400                    IK=NK+I
32500                    HOLD=A(IK)
  
```

```

600      IJ=I-N
32700    DO 65 J=1,N
32800    IJ=IJ+N
32900    IF(I-K)60,65,60
33000    60 IF(J-K)62,65,62
33100    62 KJ=IJ-I+K
33200    A(IJ)=HOLD*A(KJ)+A(IJ)
33300    65 CONTINUE
33400    KJ=K-N
33500    DO 75 J=1,N
33600    KJ=KJ+N
33700    IF(J-K)70,75,70
33800    70 A(KJ)=A(KJ)/BIGA
33900    75 CONTINUE
34000    D=D*BIGA
34100    A(KK)=1.0/BIGA
34200    80 CONTINUE
34300    K=N
34400    100 K=(K-1)
34500    IF(K)150,150,105
34600    105 I=L(K)
34700    IF(I-K)120,120,108
34800    108 JQ=N*(K-1)
34900    JR=N*(I-1)
35000    DO 110 J=1,N
35100    JK=JQ+J
35200    HOLD=A(JK)
35300    JI=JR+J
35400    A(JK)=-A(JI)
35500    110 A(IJ)=HOLD
35600    120 J=M(K)
35700    IF(J-K)100,100,125
35800    125 KI=K-N
35900    DO 130 I=1,N
36000    KI=KI+N
36100    HOLD=A(KI)
36200    JI=KI-K+J
36300    A(KI)=-A(JI)
36400    130 A(JI)=HOLD
36500    GO TO 100
36600    150 CONTINUE
36700    DD=D
36800    RETURN
36900    END

```

Program Unit Length=06C3 (1731) Bytes

Data Area Length=0047 (71) Bytes

Subroutines Referenced:

SBS	\$AT	\$L1
\$T1	\$M9	\$SB
\$NB	\$DB	\$MB
\$AB		

Variables:

	0001"	N	0003"	DD	0005"
L	0007"	M	0009"	D	000B"
K	000F"	K	0011"	KK	0013"

ABIGA	0015"	T:000000	0019"	T:010000	001B"
T:020000	001D"	J	001F"	IZ	0021"
	0023"	IJ	0025"	T:000001	0027"
010001	002B"	ABIGA	002F"	KI	0033"
HOLD	0035"	JI	0039"	JP	003B"
IK	003D"	IK	003F"	KJ	0041"
	0043"	JR	0045"		

Labels:

00L	04EB'	20L	012D'	10L	00B4'
15L	00E6'	35L	021C'	25L	017F'
0L	01EB'	45L	02D9'	38L	0246'
0L	02A8'	48L	02F1'	46L	02E2'
150L	06AE'	55L	033B'	50L	0310'
65L	0418'	60L	03B6'	62L	03CF'
5L	04A2'	70L	0482'	100L	0509'
05L	051B'	120L	05E4'	108L	0545'
110L	05B3'	125L	060E'	130L	067A'

```

7000      SUBROUTINE CORRE(N,M,IO,X,XBAR,STD,RX,R,B,D,T,NPRTM,NOPR)
37100     DIMENSION X(1),XBAR(1),STD(1),RX(1),R(1),B(1),D(1),T(1)
37200     C      DOUBLE PRECISION RX
37300     DO 100 J=1,M
37400       B(J)=0.0
37500     100   T(J)=0.0
37600       K=(M*M+M)/2
37700       DO 102 I=1,K
37800     102   R(I)=0.0
37900       FN=N
38000       L=0
38100       IF(IO)105,127,105
38200     105   DO 108 J=1,M
38300         DO 107 I=1,N
38400           L=L+1
38500     107   T(J)=T(J)+X(L)
38600           XBAR(J)=T(J)
38700     108   T(J)=T(J)/FN
38800         DO 115 I=1,N
38900           JK=0
39000           L=I-N
39100           DO 110 J=1,M
39200             L=L+N
39300             D(J)=X(L)-T(J)
39400     110   B(J)=B(J)+D(J)
39500           DO 115 J=1,M
39600             DO 115 K=1,J
39700               JK=JK+1
39800     115   R(JK)=R(JK)+D(J)*D(K)
39900           GO TO 205
40000     127   IF(N-M)130,130,135
40100     130   KK=N
40200           GO TO 137
40300     135   KK=M
40400     137   DO 140 I=1,KK
40500           CALL DATA (M,D,NPRTM,NOPR)
40600           DO 140 J=1,M
40700             T(J)=T(J)+D(J)
40800             L=L+1
40900     140   RX(L)=D(J)

```

```

000      FKK=KK
41100    DO 150 J=1,M
41200    XBAR(J)=T(J)
41300    150 T(J)=T(J)/FKK
41400    L=0
41500    DO 180 I=1,KK
41600    JK=0
41700    DO 170 J=1,M
41800    L=L+1
41900    170 D(J)=RX(L)-T(J)
42000    DO 180 J=1,M
42100    B(J)=B(J)+D(J)
42200    DO 180 K=1,J
42300    JK=JK+1
42400    180 R(JK)=R(JK)+D(J)*D(K)
42500    IF(N-KK)205,205,185
42600    185 KK=N-KK
42700    DO 200 I=1,KK
42800    JK=0
42900    CALL DATA (M,D,NPRTM,NOPR)
43000    DO 190 J=1,M
43100    XBAR(J)=XBAR(J)+D(J)
43200    D(J)=D(J)-T(J)
43300    190 B(J)=B(J)+D(J)
43400    DO 200 J=1,M
43500    DO 200 K=1,J
43600    JK=JK+1
43700    200 R(JK)=R(JK)+D(J)*D(K)
43800    205 JK=0
43900    DO 210 J=1,M
44000    XBAR(J)=XBAR(J)/FN
44100    DO 210 K=1,J
44200    JK=JK+1
44300    210 R(JK)=R(JK)-B(J)*B(K)/FN
44400    JK=0
44500    DO 220 J=1,M
44600    JK=JK+J
44700    220 STD(J)=SQRT(ABS(R(JK)))
44800    DO 230 J=1,M
44900    DO 230 K=J,M
45000    JK=J+(K*K-K)/2
45100    L=M*(J-1)+K
45200    RX(L)=R(JK)
45300    L=M*(K-1)+J
45400    RX(L)=R(JK)
45500    IF(STD(J)*STD(K))225,222,225
45600    222 R(JK)=0.0
45700    GO TO 230
45800    225 R(JK)=R(JK)/(STD(J)*STD(K))
45900    230 CONTINUE
46000    FN=SQRT(FN-1.0)
46100    DO 240 J=1,M
46200    240 STD(J)=STD(J)/FN
46300    L=-M
46400    DO 250 I=1,M
46500    L=L+M+1
46600    250 B(I)=RX(L)
46700    RETURN
46800    END

```

Program Unit Length=0A03 (2563) Bytes  
Data Area Length=0043 (67) Bytes

Subroutines Referenced:

SQRT	ABS	\$AT
L1	\$T1	\$M9
D9	\$CA	\$AB
\$DB	\$SB	\$NB
MB	DATA	

Variables:

0001"	M	0003"	IO	0005"	
0007"	XBAR	0009"	STD	000B"	
000D"	R	000F"	B	0011"	
0013"	T	0015"	NPRTM	0017"	
0019"	J	001B"	T:000000		001D"
001F"	I	0021"	FN	0023"	
0027"	T:010000		JK	002B"	
0020000	002D"	KK	002F"	FKK	0035"
T:030000	003D"	T:000001			
					0029"
					003F"

Labels:

100L	0032'	102L	0088'	105L	00DB'
127L	02CF'	108L	0155'	107L	00EE'
115L	024F'	110L	01FD'	205L	06C4'
130L	02F2'	135L	02FF'	137L	0309'
140L	0360'	150L	03E3'	180L	04C7'
170L	0436'	185L	055F'	200L	064B'
190L	05F9'	210L	06FD'	220L	0786'
230L	0918'	225L	08C7'	222L	08A7'
240L	0964'	250L	09BE'		



```
00100 C PROGRAM ITEREG
00200 C ESTIMATES REGRESSION COEFFTS BY AN ITERATIVE "WALD/BARTLETT"
00300 C METHOD.
00400 C INPUT DATA READ FROM FILE.
00500 DIMENSION Y(100),XX(100,11),X(100),XORD(100),INDX(100)
00600 DIMENSION IRANK(100),XBAR(11),IIND(10),IX(100,10),COEF(11)
00700 DIMENSION RIQ(11),RMED(11)
00800 DOUBLE PRECISION FNAM1(2)
00900 DATA LUN1,LUN2,LUN6/1,2,6/
01000 C LUN1:TERMINAL; LUN2:PRINTER; LUN6:DATA FILE FOR OUTPUT
01100 WRITE(LUN2,80)
01200 80 FORMAT(15H0PROGRAM ITEREG,/
01300 157H ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD)
01400 WRITE(LUN1,90)
01500 90 FORMAT(53H ENTER FILE NAME, 1-16 ALPHA CHARS, LAST CHAR BLANK: )
01600 READ(LUN1,95)FNAM1
01700 95 FORMAT(2A8)
01800 CALL OPEN(LUN6,FNAM1,80)
01900 WRITE(LUN2,209)FNAM1
02000 209 FORMAT(17H0DATA FILE NAME: ,2A8)
02100 WRITE(LUN1,201)
02200 201 FORMAT(46H ENTER NO OF OBS, NO OF VARS, INDEX OF DEP VAR/
02300 133H NO OF INDEP VARS IN REGRESSION: )
02400 READ(LUN1,18)NOBS,NVAR,IDEP,NIND
02500 18 FORMAT(16I5)
02600 WRITE(LUN2,16)
02700 16 FORMAT(40H0NO OF OBS, NO OF VARS, INDEX OF DEP VAR/
02800 149H IN REGRESSION, NO OF INDEP VARS IN REGRESSION...)
02900 IO=0
03000 WRITE(LUN2,13)NOBS,NVAR,IDEP,NIND
03100 13 FORMAT(1X,8I10)
03200 WRITE(LUN1,19)
03300 19 FORMAT(44H ENTER INDICES OF INDEP VARS, IN ASC ORDER: /)
03400 READ(LUN1,18)(IIND(I),I=1,NIND)
03500 WRITE(LUN2,215)
03600 215 FORMAT(39H INDICES OF INDEP VARS IN REGRESSION...)
03700 WRITE(LUN2,13)(IIND(I),I=1,NIND)
03800 WRITE(LUN1,221)
03900 221 FORMAT(33H ENTER MAX NO OF OBSNS TO PRINT: )
04000 READ(LUN1,222)NPRTM
04100 222 FORMAT(I4)
04200 NPRTM=MIN0(NPRTM,NOBS)
04300 NOPR=0
04400 WRITE(LUN1,230)
04500 230 FORMAT(49H INPUT 0 TO USE WALD EST, 1 TO USE BARTLETT EST: )
04600 READ(LUN1,231)IWB
04700 231 FORMAT(I1)
04800 IF(IWB.GT.0)GO TO 232
04900 WRITE(LUN2,233)
05000 233 FORMAT(56H0REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD
05100 1)
05200 GO TO 234
05300 232 CONTINUE
05400 WRITE(LUN2,235)
05500 235 FORMAT(60H0REGRESSION METHOD USED AT EACH ITERATION: BARTLETT'S ME
05600 1THOD)
05700 234 CONTINUE
05800 WRITE(LUN1,1239)
```

```

0900      1239  FORMAT(36H ENTER NO OF ITERATIONS TO PERFORM: )
06000    READ(LUN1,1250)NITER
05100    1250  FORMAT(I3)
05200    WRITE(LUN1,1264)
06300    1264  FORMAT(44H PRINT RESULTS EVERY N ITERATIONS; INPUT N: )
05400    READ(LUN1,1250)NPRT
05500    KPRT=0
06600    WRITE(LUN2,1262)NPRT
06700    1262  FORMAT(19H0PRINT AFTER EVERY ,I3,11H ITERATIONS)
05800    WRITE(LUN2,12)
05900    12   FORMAT(8H0DATA...)
07000    WRITE(LUN2,223)NPRTM
07100    223  FORMAT(8H (FIRST ,I4,7H OBSNS))
07200    C   READ DATA
07300    NINDP=NIND+1
07400    DO 1110 J=1,NOBS
07500    READ(LUN6,10)(X(I),I=1,NVAR)
07600    10   FORMAT(8F10.4)
07700    NOPR=NOPR+1
07800    IF(NOPR.GT.NPRTM)GO TO 224
07900    WRITE(LUN2,11)(X(I),I=1,NVAR)
08000    224  CONTINUE
08100    11   FORMAT(1X,8F10.4)
08200    Y(J)=X(IDEF)
08300    DO 1112 I=1,NIND
08400    INDEX=IIND(I)
08500    XX(J,I)=X(INDEX)
08600    1112 CONTINUE
08700    XX(J,NINDP)=X(IDEF)
08800    1110 CONTINUE
08900    C   COMPUTE MEANS
09000    DO 1120 I=1,NINDP
09100    COEF(I)=0.0
09200    XBAR(I)=0.0
09300    DO 1130 J=1,NOBS
09400    XBAR(I)=XBAR(I)+XX(J,I)
09500    1130 CONTINUE
09600    XBAR(I)=XBAR(I)/NOBS
09700    1120 CONTINUE
09800    WRITE(LUN2,32)
09900    32  FORMAT(35H0MEANS FOR VARIABLES IN REGRESSION,/
10000    127H DEPENDENT VARIABLE LAST...)
10100    WRITE(LUN2,20)(XBAR(I),I=1,NINDP)
10200    20  FORMAT(1X,8F10.4)
10300    COEF(NINDP)=XBAR(NINDP)
10400    C   ORDER THE X'S
10500    DO 1160 I=1,NIND
10600    DO 1170 J=1,NOBS
10700    X(J)=XX(J,I)
10800    1170 CONTINUE
10900    CALL RANK(X,NOBS,XORD,INDX,IRANK)
11000    DO 1180 J=1,NOBS
11100    XX(J,I)=XORD(J)
11200    IX(J,I)=INDX(J)
11300    1180 CONTINUE
11400    C   DETERMINE MEDIAN, INTERQUARTILE RANGE
11500    IQ1=.25*NOBS
11600    IQM=.5*NOBS
11700    IQ2=.75*NOBS
11800    RMED(I)=XORD(IQM)

```

```

900      RIQ(I)=XORD(IQ2)-XORD(IQ1)
1200    1160 CONTINUE
1210    CALL RANK(Y,NOBS,XORD,INDX,IRANK)
1220    RMED(NINDP)=XORD(IQM)
1230    RIQ(NINDP)=XORD(IQ2)-XORD(IQ1)
1240    WRITE(LUN2,22)
1250    22 FORMAT(11H0MEDIANS...)
1260    WRITE(LUN2,20)(RMED(I),I=1,NINDP)
1270    WRITE(LUN2,21)
1280    21 FORMAT(24H0INTERQUARTILE RANGES...)
1290    WRITE(LUN2,20)(RIQ(I),I=1,NINDP)
1300    WRITE(LUN2,1235)
1310    1235 FORMAT(38H0REGRESSION COEFFICIENTS, INTERCEPT...)
1320    IF(NITER.LE.0)GO TO 126
1330    125 CONTINUE
1340    DO 1300 ITER=1,NITER
1350  C COMPUTE REGRESSION COEFFT ADJUSTMENTS
1360    DO 1200 I=1,NIND
1370    DO 1210 J=1,NOBS
1380    X(J)=XX(J,I)
1390    INDX(J)=IX(J,I)
1400    1210 CONTINUE
1410    IF(IWB.GT.0)GO TO 1211
1420    CALL WALD(Y,X,NOBS,INDX,A)
1430    GO TO 1212
1440    1211 CONTINUE
1450    CALL BART(Y,X,NOBS,INDX,A)
1460    1212 CONTINUE
1470  C ADJUST REGRESSION COEFFT
1480    COEF(I)=COEF(I)+A
1490  C ADJUST Y
1500    DO 1220 J=1,NOBS
1510    IY=IX(J,I)
1520    Y(IY)=Y(IY)-A*XX(J,I)
1530    1220 CONTINUE
1540  C ADJUST CONSTANT TERM
1550    COEF(NINDP)=COEF(NINDP)-A*XBAR(I)
1560    1200 CONTINUE
1570  C PRINT RESULTS
1580    KPRT=KPRT+1
1590    IF(KPRT.LT.NPRT)GO TO 1270
1600    KPRT=0
1610    WRITE(LUN2,1230)(COEF(I),I=1,NINDP)
1620    1230 FORMAT(1X,8F10.4)
1630    1270 CONTINUE
1640    1300 CONTINUE
1650    WRITE(LUN1,1240)
1660    1240 FORMAT(41H ENTER NO OF ADDL ITERATIONS TO PERFORM: )
1670    READ(LUN1,1250)NITER
1680    IF(NITER.LE.0)GO TO 126
1690    WRITE(LUN1,1264)
1700    READ(LUN1,1250)NPRT
1710    WRITE(LUN2,1262)NPRT
1720    GO TO 125
1730    126 CONTINUE
1740  C ESTIMATE RESIDUAL VARIANCE
1750  C PUT THE X'S BACK IN ORIGINAL ORDER
1760    DO 1500 I=1,NIND
1770    DO 1510 J=1,NOBS
1780    X(J)=XX(J,I)

```

```

1900      1510 CONTINUE
18000      DO 1520 J=1,NOBS
18100      JORIG=IX(J,I)
18200      XX(JORIG,I)=X(J)
18300      1520 CONTINUE
18400      1500 CONTINUE
18500      YINT=COEF(NINDP)
18600      RV=0.0
18700      DO 1400 J=1,NOBS
18800      YEST=YINT
18900      DO 1410 I=1,NIND
19000      YEST=YEST+COEF(I)*XX(J,I)
19100      1410 CONTINUE
19200      RV=RV+(XX(J,NINDP)-YEST)**2
19300      1400 CONTINUE
19400      RV=RV/(NOBS-NIND-1)
19500      RSD=SQRT(RV)
19600      WRITE(LUN2,1420)RV,RSD
19700      1420 FORMAT(21H0RESIDUAL VARIANCE = ,F14.4/
19800      127H RESIDUAL STANDARD ERROR = ,F14.4)
19900      C COMPUTE COEFFT OF DETERMINATION
20000      VY=0.0
20100      DO 1430 J=1,NOBS
20200      VY=VY+XX(J,NINDP)**2
20300      1430 CONTINUE
20400      VY=(VY-NOBS*XBAR(NINDP)**2)/(NOBS-1)
20500      CD=(VY-RV)/VY
20600      WRITE(LUN2,1435)CD
20700      1435 FORMAT(32H COEFFICIENT OF DETERMINATION = ,F10.4)
20800      END

```

Program Unit Length=0C0E (3086) Bytes  
Data Area Length=2555 (9557) Bytes

Subroutines Referenced:

INO	\$I3	\$I1
IO	SQRT	\$INIT
\$W2	\$ND	\$R2
OPEN	\$L1	\$T1
M9	\$AB	\$DA
RANK	\$MA	\$CH
SSB	\$NB	WALD
PART	\$MB	\$EA
\$DB	\$EX	

Variables:

Y	0001"	XX	0191"	X	12C1"
WORD	1451"	INDX	15E1"	IRANK	16A9"
BAR	1771"	IIND	179D"	IX	17B1"
COEF	1F81"	RIQ	1FAD"	RMED	1FD9"
ENAM1	2005"	LUN1	2015"	LUN2	2017"
UN6	2019"	NOBS	211E"	NVAR	2120"
IDEP	2122"	NIND	2124"	IO	2194"
I	21D7"	T:000000	21D9"	NPRTM	222D"
OPR	2233"	IWB	226B"	T:000002	2271"
ITER	2319"	NPRT	2350"	KPRT	2352"
NINDP	23A3"	J	23A5"	INDEX	23BA"
T:010000	23BC"	T:020000	23BE"	T:030000	2418"

Q1	241A"	IQM	241C"	IQ2	241E"
ITER	247E"	A	2480"	IY	2490"
RORIG	24CB"	YINT	24CD"	RV	24D1"
TEST	24D5"	RSD	24D9"	VY	2522"
CD	2526"				

Labels:

SSL	0006'	80L	201B"	90L	206D"
SL	20A7"	209L	20AC"	201L	20C6"
BL	212C"	16L	2132"	13L	219C"
19L	21A5"	215L	21DB"	221L	2207"
22L	222F"	230L	2235"	231L	226D"
32L	01B1'	233L	2272"	234L	01BD'
235L	22AF"	1239L	22F0"	1250L	231B"
1264L	231F"	1262L	2354"	12L	237E"
23L	238A"	1110L	0389'	10L	23A7"
224L	02E6'	11L	23AF"	1112L	0349'
1120L	0437'	1130L	040C'	32L	23C0"
2L	2407"	1160L	05FA'	1170L	04E3'
180L	0562'	22L	2426"	21L	2436"
1235L	2453"	126L	096B'	125L	0714'
300L	08E7'	1200L	0877'	1210L	077D'
211L	07B0'	1212L	07BC'	1220L	083A'
1270L	08E7'	1230L	2492"	1240L	249D"
1500L	0A16'	1510L	09A2'	1520L	0A06'
100L	0ADE'	1410L	0A98'	1420L	24DD"
1430L	0B7C'	1435L	252A"		

```

0900      SUBROUTINE RANK(X,N,XORD,INDX,IRANK)
0000      C THIS SUBROUTINE DETERMINES THE RANKS OF THE COMPONENTS OF THE
21100     C VECTOR X, AND ALSO ORDERS THESE COMPONENTS (ASCENDING ORDER).
01200     C INPUT: VECTOR X(N), DIMENSION N
01300     C OUTPUT: RANKS OF ORIGINAL X'S ARE STORED IN VECTOR IRANK
21400     C ORDERED X'S ARE STORED IN XORD
21500     C ORIGINAL INDICES OF ORDERED X'S ARE STORED IN INDX
01600     DIMENSION X(1),XORD(1),IRANK(1),INDX(1)
21700     DO 50 I=1,N
21800     INDX(I)=I
01900     50 XORD(I)=X(I)
22000     NM1=N-1
22100     DO 100 I=1,NM1
02200     IP1=I+1
02300     DO 200 J=IP1,N
22400     IF(XORD(I).LE.XORD(J))GO TO 200
22500     TEMP=XORD(I)
02600     XORD(I)=XORD(J)
22700     XORD(J)=TEMP
22800     ITEMP=INDX(I)
02900     INDX(I)=INDX(J)
03000     INDX(J)=ITEMP
23100     200 CONTINUE
03200     100 CONTINUE
03300     DO 300 I=1,N
23400     I1=INDX(I)
23500     IRANK(I1)=I
03600     300 CONTINUE
23700     RETURN
23800     END

```

Program Unit Length=01AE (430) Bytes  
Data Area Length=0028 (40) Bytes

Subroutines Referenced:

SAT		\$L1		\$T1		
SB						
Variables:						
	0001"	N	0003"	XORD	0005"	
INDX	0007"	IRANK	0009"	I	000B"	
	000000	000D"	T:010000	000F"	NM1	0011"
	0013"	J	0015"	T:000002		0017"
TEMP	0018"	ITEMP	001C"	T:020000		001E"
T:030000		0020"	T:040000	0022"	T:050000	0024"
	0026"					

Labels:

0L	002D'	100L	015C'	200L	0148'
300L	0199'				

```
3900          SUBROUTINE WALD(Y,X,N,INDX,A)
24000 C THIS SUBROUTINE COMPUTES THE WALD ESTIMATE A IN THE REGRESSION
24100 C EQUATION Y=A*X.
24200 C INPUT VECTOR Y(DEP VAR), VECTOR X(INDEP VAR), VECTOR INDX
24300 C (ORIGINAL INDICES OF THE ORDERED X'S, BEFORE ORDERING), DIMENSION N
24400 C OUTPUT REGRESSION COEFFT A
24500 C X-DATA ARE ORDERED (ASCENDING ORDER)
24600 C INDX CONTAINS INDICES OF Y CORRESP TO EACH X
24700          DIMENSION Y(1),X(1),INDX(1)
24800 C XBAR1 = MEAN OF X'S BELOW MEDIAN
24900 C XBAR2 = MEAN OF X'S ABOVE MEDIAN
25000          XBAR1=0.0
25100          YBAR1=0.0
25200          XBAR2=0.0
25300          YBAR2=0.0
25400          N2=N/2
25500          DO 100 I=1,N2
25600          XBAR1=XBAR1+X(I)
25700          IY=INDX(I)
25800          YBAR1=YBAR1+Y(IY)
25900          100 CONTINUE
26000 C LEAVE OUT MIDDLE OBSN IF N IS ODD
26100          N3=N-N2+1
26200          DO 200 I=N3,N
26300          XBAR2=XBAR2+X(I)
26400          IY=INDX(I)
26500          YBAR2=YBAR2+Y(IY)
26600          200 CONTINUE
26700          A=(YBAR2-YBAR1)/(XBAR2-XBAR1)
26800          RETURN
26900          END
```

Program Unit Length=0157 (343) Bytes  
Data Area Length=0027 (39) Bytes

Subroutines Referenced:

AT	\$L1	\$T1
\$D9	\$AB	\$SB
\$DB		

Variables:

	0001"	X	0003"	N	0005"
INDX	0007"	A	0009"	XBAR1	000B"
YBAR1	000F"	XBAR2	0013"	YBAR2	0017"
N2	001B"	I	001D"	IY	001F"
N3	0021"	T:000001	0023"		

Labels:

100L	009F'	200L	0114'
------	-------	------	-------

```

27000      SUBROUTINE BART(Y,X,N,INDX,A)
27100      C THIS SUBROUTINE COMPUTES THE BARTLETT ESTIMATE A IN THE REGRESSION
27200      C EQUATION Y=A*X.
27300      C THE X-DATA ARE ORDERED (ASCENDING ORDER)
27400      C INDX CONTAINS INDICES OF Y CORRESPONDING TO EACH X
27500      C INPUT VECTOR Y(DEP VAR), VECTOR X(INDEP VAR), DIMENSION N
27600      C OUTPUT REGRESSION COEFFT A
27700      DIMENSION Y(1),X(1),INDX(1)
27800      N3=N/3
27900      XBAR1=0.0
28000      YBAR1=0.0
28100      XBAR2=0.0
28200      YBAR2=0.0
28300      DO 100 I=1,N3
28400      XBAR1=XBAR1+X(I)
28500      IY=INDX(I)
28600      YBAR1=YBAR1+Y(IY)
28700      100 CONTINUE
28800      N2=N-N3+1
28900      DO 200 I=N2,N
29000      XBAR2=XBAR2+X(I)
29100      IY=INDX(I)
29200      YBAR2=YBAR2+Y(IY)
29300      200 CONTINUE
29400      A=(YBAR2-YBAR1)/(XBAR2-XBAR1)
29500      RETURN
29600      END

```

Program Unit Length=0157 (343) Bytes  
Data Area Length=0027 (39) Bytes

Subroutines Referenced:

AT	\$D9	\$L1
T1	\$AB	\$SB
\$DB		

Variables:

Y	0001"	X	0003"	N	0005"
INDX	0007"	A	0009"	N3	000B"
YBAR1	000D"	YBAR1	0011"	XBAR2	0015"
YBAR2	0019"	I	001D"	IY	001F"
N2	0021"	T:000001	0023"		

Labels:

00L 009F' 200L 0114'



Appendix B  
Computer Program Output

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 1

COEFFICIENT MATRIX (C)...  
COEFFT MATRIX IS IDENTITY MATRIX

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

B0 (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...  
1.0000

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100  
SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...  
100.0000 100.0000  
100.0000 101.0000

COEFFICIENT OF DETERMINATION = .9901

CORRELATION MATRIX OF X, Y...  
1.0000 .9950  
.9950 1.0000

OBSERVATIONS...  
DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)  
OBSN NO 1:  
-18.3276 -18.8362  
OBSN NO 2:  
-13.7183 -13.1409  
OBSN NO 3:  
13.3911 12.3044

PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...  
100 2 2 1  
INDICES OF INDEP VARS IN REGRESSION...  
1

DATA...  
(FIRST 3 OBSNS)  
-18.3276 -18.8362  
-13.7183 -13.1409  
13.3911 12.3044

MEANS FOR ALL VARIABLES...  
1.2364 2.2018

STANDARD DEVIATIONS FOR ALL VARIABLES...  
9.4218 9.5641

CORRELATION MATRIX...  
1.0000  
.9932 1.0000

DETERMINANT = .1000E+01

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...  
1.0000

PRODUCT OF MATRIX AND INVERSE...  
SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.  
1.0000

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...  
.1138E-03

INDEX OF DEP VAR IN REGRESSION = 2  
INDICES OF INDEP VARS IN REGRESSION...  
1

REGRESSION COEFFICIENTS...

1.0082

T-VALUES...

84.2544

STD DEVS OF REG COEFS...

.0120

INTERCEPT = .9553

RESIDUAL STANDARD ERROR = 1.1217

SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9932

SAMPLE COEFFICIENT OF DETERMINATION = .9864

SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 8932.3516

DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 1.0000

MEAN SQUARE OF SSAR = 8932.3516

SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = 123.3125

DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 98.0000

MEAN SQUARE OF SDR =  
F-VALUE =

1.2583

7098.7979

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...  
100 2 2 1  
INDICES OF INDEP VARS IN REGRESSION...  
1

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)  
-18.3276 -18.8362  
-13.7183 -13.1409  
13.3911 12.3044

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

1.2364 2.2018

MEDIANS...

.9692 1.7888

INTERQUARTILE RANGES...

12.9687 13.4609

REGRESSION COEFFICIENTS, INTERCEPT...

1.0062 .9577  
1.0062 .9577  
1.0062 .9577  
1.0062 .9577  
1.0062 .9577  
1.0062 .9577  
1.0062 .9577  
1.0062 .9577  
1.0062 .9577  
1.0062 .9577

RESIDUAL VARIANCE = 1.2587  
RESIDUAL STANDARD ERROR = 1.1219  
COEFFICIENT OF DETERMINATION = .9862

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...  
100            2            2            1  
INDICES OF INDEP VARS IN REGRESSION...  
1

REGRESSION METHOD USED AT EACH ITERATION: BARTLETT'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

FIRST 3 OBSNS)  
-18.3276   -18.8362  
-13.7183   -13.1409  
13.3911   12.3044

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...  
1.2364      2.2018

MEDIANS...

.9692      1.7888

INTERQUARTILE RANGES...

12.9687    13.4609

REGRESSION COEFFICIENTS, INTERCEPT...

1.0103      .9527  
1.0103      .9527  
1.0103      .9527  
1.0103      .9527  
1.0103      .9527  
1.0103      .9527  
1.0103      .9527  
1.0103      .9527  
1.0103      .9527  
1.0103      .9527

RESIDUAL VARIANCE =            1.2587  
RESIDUAL STANDARD ERROR =      1.1219  
COEFFICIENT OF DETERMINATION =    .9862

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 1

COEFFICIENT MATRIX (C)...

COEFFT MATRIX IS IDENTITY MATRIX

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 2.0000

μ (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

4.0000	4.0000
4.0000	5.0000

COEFFICIENT OF DETERMINATION = .8000

CORRELATION MATRIX OF X, Y...

1.0000	.8944
.8944	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:
-3.6655	-4.1741
OBSN NO	2:
-2.7437	-2.1663
OBSN NO	3:
2.6782	1.5916

PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100 2 2 1

INDICES OF INDEP VARS IN REGRESSION...

1

DATA...

(FIRST 3 OBSNS)

-3.6655 -4.1741  
-2.7437 -2.1663  
2.6782 1.5916

MEANS FOR ALL VARIABLES...

.2473 1.2127

STANDARD DEVIATIONS FOR ALL VARIABLES...

1.8844 2.2566

CORRELATION MATRIX...

1.0000  
.8691 1.0000

DETERMINANT = .1000E+01

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.0000

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.2845E-02

INDEX OF DEP VAR IN REGRESSION = 2

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION COEFFICIENTS...

1.0408

F-VALUES...

17.3963

STD DEVS OF REG COEFS...

.0598

INTERCEPT = .9553

RESIDUAL STANDARD ERROR = 1.1217

SAMPLE MULTIPLE CORRELATION COEFFICIENT = .8691

SAMPLE COEFFICIENT OF DETERMINATION = .7554

SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 380.8064

DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 1.0000

MEAN SQUARE OF SSAR = 380.8064

SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = 123.3153

DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 98.0000



MEAN SQUARE OF SSDR = 1.2583  
F-VALUE = 302.6309

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            2            2            1

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

FIRST 3 OBSNS)

-3.6655	-4.1741
-2.7437	-2.1663
2.6782	1.5916

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.2473      1.2127

MEDIANS...

.1938      1.1462

INTERQUARTILE RANGES...

2.5937      3.2969

REGRESSION COEFFICIENTS, INTERCEPT...

1.0310	.9577
1.0310	.9577
1.0310	.9577
1.0310	.9577
1.0310	.9577
1.0310	.9577
1.0310	.9577
1.0310	.9577
1.0310	.9577
1.0310	.9577

RESIDUAL VARIANCE = 1.2587

RESIDUAL STANDARD ERROR = 1.1219

COEFFICIENT OF DETERMINATION = .7528

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            2            2            1

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION METHOD USED AT EACH ITERATION: BARTLETT'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-3.6655	-4.1741
-2.7437	-2.1663
2.6782	1.5916

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.2473	1.2127
-------	--------

MEDIANS...

.1938	1.1462
-------	--------

INTERQUARTILE RANGES...

2.5937	3.2969
--------	--------

REGRESSION COEFFICIENTS, INTERCEPT...

1.0513	.9527
1.0513	.9527
1.0513	.9527
1.0513	.9527
1.0513	.9527
1.0513	.9527
1.0513	.9527
1.0513	.9527
1.0513	.9527
1.0513	.9527

RESIDUAL VARIANCE = 1.2587

RESIDUAL STANDARD ERROR = 1.1219

COEFFICIENT OF DETERMINATION = .7528

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 1

COEFFICIENT MATRIX (C)...

COEFFT MATRIX IS IDENTITY MATRIX

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 1.0000

INTERCEPT = 1.0000

PARAMETER VECTOR (B)...

1.0000

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

1.0000 1.0000

1.0000 2.0000

COEFFICIENT OF DETERMINATION = .5000

CORRELATION MATRIX OF X, Y...

1.0000 .7071

.7071 1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST

(FIRST 3 OBSNS PRINTED)

OBSN NO 1:

-1.8328 -2.3413

OBSN NO 2:

-1.3718 -.7944

OBSN NO 3:

1.3391 .2524

PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            2            2            1

INDICES OF INDEP VARS IN REGRESSION...

1

DATA...

(FIRST        3 OBSNS)

-1.8328      -2.3413

-1.3718      -.7944

1.3391        .2524

MEANS FOR ALL VARIABLES...

.1236        1.0890

STANDARD DEVIATIONS FOR ALL VARIABLES...

.9422        1.5113

CORRELATION MATRIX...

1.0000

.6743        1.0000

DETERMINANT = .1000E+01

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.0000

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS

HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1138E-01

INDEX OF DEP VAR IN REGRESSION =                    2

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION COEFFICIENTS...

1.0816

T-VALUES...

9.0392

STD DEVS OF REG COEFS...

.1197

INTERCEPT = .9553

RESIDUAL STANDARD ERROR = 1.1217

SAMPLE MULTIPLE CORRELATION COEFFICIENT = .6743

SAMPLE COEFFICIENT OF DETERMINATION = .4547

SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 102.8119

DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 1.0000

MEAN SQUARE OF SSAR = 102.8119

SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = 123.3141

DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 98.0000

MEAN SQUARE OF SDR =  
F-VALUE =

1.2583

81.7065

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            2            2            1

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

FIRST 3 OBSNS)  
-1.8328    -2.3413  
-1.3718    -.7944  
1.3391     .2524

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.1236      1.0890

MEDIANS...

.0969      .9399

INTERQUARTILE RANGES...

1.2969      1.9375

REGRESSION COEFFICIENTS, INTERCEPT...

1.0619      .9577  
1.0619      .9577  
1.0619      .9577  
1.0619      .9577  
1.0619      .9577  
1.0619      .9577  
1.0619      .9577  
1.0619      .9577  
1.0619      .9577

RESIDUAL VARIANCE = 1.2587

RESIDUAL STANDARD ERROR = 1.1219

COEFFICIENT OF DETERMINATION = .4489

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            2            2            1

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION METHOD USED AT EACH ITERATION: BARTLETT'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

FIRST 3 OBSNS)  
-1.8328    -2.3413  
-1.3718    -.7944  
1.3391     .2524

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.1236      1.0890

MEDIANS...

.0969      .9399

INTERQUARTILE RANGES...

1.2969      1.9375

REGRESSION COEFFICIENTS, INTERCEPT...

1.1025      .9527  
1.1025      .9527  
1.1025      .9527  
1.1025      .9527  
1.1025      .9527  
1.1025      .9527  
1.1025      .9527  
1.1025      .9527  
1.1025      .9527  
1.1025      .9527

RESIDUAL VARIANCE = 1.2587

RESIDUAL STANDARD ERROR = 1.1219

COEFFICIENT OF DETERMINATION = .4489



PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 1

COEFFICIENT MATRIX (C)...

COEFFT MATRIX IS IDENTITY MATRIX

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = .2000

Q (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000

SIG (MODEL ERROR STD DEV) = 1.0000

N OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

.0400	.0400
.0400	1.0400

COEFFICIENT OF DETERMINATION = .0385

CORRELATION MATRIX OF X, Y...

1.0000	.1961
.1961	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:		
		-.3666	-.8751
OBSN NO	2:		
		-.2744	.3030
OBSN NO	3:		
		.2678	-.8188

PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILE

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100 2 2 1

INDICES OF INDEP VARS IN REGRESSION...

1

DATA...

(FIRST 3 OBSNS)

-.3666 -.8751  
-.2744 .3030  
.2678 -.8188

MEANS FOR ALL VARIABLES...

.0247 .9901

STANDARD DEVIATIONS FOR ALL VARIABLES...

.1884 1.1472

CORRELATION MATRIX...

1.0000  
.2313 1.0000

DETERMINANT = .1000E+01

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.0000

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUND OFF ERRORS

HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.2845E+00

INDEX OF DEP VAR IN REGRESSION = 2

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION COEFFICIENTS...

1.4080

T-VALUES...

2.3535

STD DEVS OF REG COEFS...

.5983

INTERCEPT = .9553

RESIDUAL STANDARD ERROR = 1.1217

SAMPLE MULTIPLE CORRELATION COEFFICIENT = .2313

SAMPLE COEFFICIENT OF DETERMINATION = .0535

SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 6.9698

DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 1.0000

MEAN SQUARE OF SSAR = 6.9698

SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSSR = 123.3141

DEGREES OF FREEDOM ASSOCIATED WITH SSSR = 98.0000

AD-A167 601

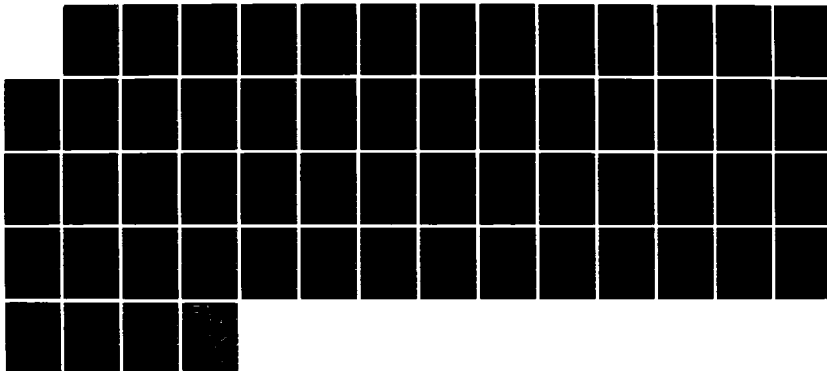
IMPROVED ALGORITHMS FOR ESTIMATION PREDICTION AND  
CONTROL(U) VISTA RESEARCH CORP TUSCON AZ J G CALDWELL  
26 JAN 86 N00014-85-C-0814

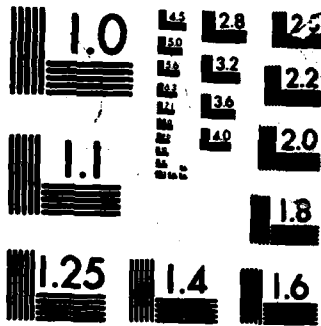
2/2

UNCLASSIFIED

F/G 9/2

NL





MICROCOPY

CHART

MEAN SQUARE OF SDR = 1.2583  
F-VALUE = 5.5391

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            2            2            1

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-.3666	-.8751
-.2744	.3030
.2678	-.8188

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.0247            .9901

MEDIANS...

.0194            .9687

INTERQUARTILE RANGES...

.2594            1.5406

REGRESSION COEFFICIENTS, INTERCEPT...

1.3097	.9577
1.3097	.9577
1.3097	.9577
1.3097	.9577
1.3097	.9577
1.3097	.9577
1.3097	.9577
1.3097	.9577
1.3097	.9577
1.3097	.9577

RESIDUAL VARIANCE = 1.2587

RESIDUAL STANDARD ERROR = 1.1219

COEFFICIENT OF DETERMINATION = .0436

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            2            2            1

INDICES OF INDEP VARS IN REGRESSION...

1

REGRESSION METHOD USED AT EACH ITERATION: BARTLETT'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

FIRST 3 OBSNS)  
-.3666      -.8751  
-.2744      .3030  
.2678      -.8188

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.0247      .9901

EDIANS...

.0194      .9687

INTERQUARTILE RANGES...

.2594      1.5406

REGRESSION COEFFICIENTS, INTERCEPT...

1.5126      .9527  
1.5126      .9527  
1.5126      .9527  
1.5126      .9527  
1.5126      .9527  
1.5126      .9527  
1.5126      .9527  
1.5126      .9527  
1.5126      .9527  
1.5126      .9527

RESIDUAL VARIANCE = 1.2587

RESIDUAL STANDARD ERROR = 1.1219

COEFFICIENT OF DETERMINATION = .0435

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 3

COEFFICIENT MATRIX (C)...

COEFFT MATRIX IS IDENTITY MATRIX

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

μ (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000 -1.0000 1.0000

SIG (MODEL ERROR STD DEV) = 1.0000

N OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	0.0000	0.0000	100.0000
0.0000	100.0000	0.0000	-100.0000
0.0000	0.0000	100.0000	100.0000
100.0000	-100.0000	100.0000	301.0000

COEFFICIENT OF DETERMINATION = .9967

CORRELATION MATRIX OF X, Y...

1.0000	0.0000	0.0000	.5764
0.0000	1.0000	0.0000	-.5764
0.0000	0.0000	1.0000	.5764
.5764	-.5764	.5764	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST

(FIRST 3 OBSNS PRINTED)

OBSN NO	1:			
	-18.3276	-15.0855	-13.7183	-16.3831
OBSN NO	2:			
	13.3911	-20.8667	-6.9995	28.7576
OBSN NO	3:			
	-14.8901	3.3520	-20.2808	-37.1018



PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFIL1

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	4	4	3
INDICES OF INDEP VARS IN REGRESSION...			
1	2	3	

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-16.3831
13.3911	-20.8667	-6.9995	28.7576
-14.8901	3.3520	-20.2808	-37.1018

MEANS FOR ALL VARIABLES...

.0505	-1.4574	.7599	3.1979
-------	---------	-------	--------

STANDARD DEVIATIONS FOR ALL VARIABLES...

9.2663	10.3026	8.5514	16.2291
--------	---------	--------	---------

CORRELATION MATRIX...

1.0000			
.1866	1.0000		
.1537	-.0497	1.0000	
.5326	-.5513	.6490	1.0000

DETERMINANT = .9362E+00

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.0655			
-.2075	1.0429		
-.1741	.0837	1.0309	

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000			
.0000	1.0000		
-.0000	-.000*	1.0000	

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1253E-03			
-.2195E-04	.9924E-04		
-.2219E-04	.9602E-05	.1424E-03	

INDEX OF DEP VAR IN REGRESSION = 4

INDICES OF INDEP VARS IN REGRESSION...

1	2	3	
REGRESSION COEFFICIENTS...			
.9964	-.9942	1.0063	

T-VALUES...  
77.3142 -86.6936 73.2542

STD DEVS OF REG COEFS...  
.0129 .0115 .0137

INTERCEPT = .9340

RESIDUAL STANDARD ERROR = 1.1511  
SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9976  
SAMPLE COEFFICIENT OF DETERMINATION = .9951  
SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 25947.7559  
DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 3.0000  
MEAN SQUARE OF SSAR = 8649.2520  
SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = 127.2129  
DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 96.0000  
MEAN SQUARE OF SSDR = 1.3251  
F-VALUE = 6527.0757

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	4	4	3
INDICES OF INDEP VARS IN REGRESSION...			
1	2	3	

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-16.3831
13.3911	-20.8667	-6.9995	28.7576
-14.8901	3.3520	-20.2808	-37.1018

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.0505	-1.4574	.7599	3.1979
-------	---------	-------	--------

MEDIANS...

-.0464	-1.8042	.5005	1.8826
--------	---------	-------	--------

INTERQUARTILE RANGES...

13.2813	14.8438	12.0312	23.7656
---------	---------	---------	---------

REGRESSION COEFFICIENTS, INTERCEPT...

.8977	-.9887	1.0277	.9307
.9905	-.9949	1.0086	.9316
.9952	-.9959	1.0077	.9306
.9956	-.9959	1.0076	.9305
.9956	-.9959	1.0076	.9305
.9956	-.9959	1.0076	.9305
.9956	-.9959	1.0076	.9305
.9956	-.9959	1.0076	.9305
.9956	-.9959	1.0076	.9305
.9956	-.9959	1.0076	.9305

RESIDUAL VARIANCE = 1.3257

RESIDUAL STANDARD ERROR = 1.1514

COEFFICIENT OF DETERMINATION = .9950

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 3

COEFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000
.1000	1.0000	.1000
.5000	.5000	1.0000

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

B0 (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000	-1.0000	1.0000
--------	---------	--------

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	10.0000	50.0000	140.0000
10.0000	102.0000	65.0000	-27.0000
50.0000	65.0000	150.0000	135.0000
140.0000	-27.0000	135.0000	303.0000

COEFFICIENT OF DETERMINATION = .9967

CORRELATION MATRIX OF X, Y...

1.0000	.0990	.4082	.8043
.0990	1.0000	.5255	-.1536
.4082	.5255	1.0000	.6332
.8043	-.1536	.6332	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:			
	-18.3276	-18.2900	-30.4248	-29.8850
OBSN NO	2:			
	13.3911	-20.2275	-10.7373	24.3806
OBSN NO	3:			
	-14.8901	-.1650	-26.0498	-39.3538

PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	4	4	3
INDICES OF INDEP VARS IN REGRESSION...			
1	2	3	

DATA...  
(FIRST 3 OBSNS)

-18.3276	-18.2900	-30.4248	-29.8850
13.3911	-20.2275	-10.7373	24.3806
-14.8901	-.1650	-26.0498	-39.3538

MEANS FOR ALL VARIABLES...

.0505	-1.3763	.0565	2.4134
-------	---------	-------	--------

STANDARD DEVIATIONS FOR ALL VARIABLES...

9.2663	10.5197	11.7402	16.2744
--------	---------	---------	---------

CORRELATION MATRIX...

1.0000			
.2833	1.0000		
.5885	.5806	1.0000	
.8099	-.0630	.6842	1.0000

DETERMINANT = .4300E+00

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.5418			
.1356	1.5204		
-.9860	-.9625	2.1390	

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000			
.0000	1.0000		
.0000	.000*	1.0000	

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1814E-03			
.1405E-04	.1388E-03		
-.9155E-04	-.7872E-04	.1568E-03	

INDEX OF DEP VAR IN REGRESSION = 4

INDICES OF INDEP VARS IN REGRESSION...

1	2	3
---	---	---

REGRESSION COEFFICIENTS...

.9931	-.9972	1.0060
-------	--------	--------

T-VALUES...

64.0471	-73.5204	69.7856
---------	----------	---------

STD DEVS OF REG COEFS...

.0155	.0136	.0144
-------	-------	-------

INTERCEPT = .9341

RESIDUAL STANDARD ERROR = 1.1514  
SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9976  
SAMPLE COEFFICIENT OF DETERMINATION = .9951  
SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 26093.6445  
DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 3.0000  
MEAN SQUARE OF SSAR = 8697.8818  
SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = 127.2637  
DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 96.0000  
MEAN SQUARE OF SSDR = 1.3257  
T-VALUE = 6561.1553

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	4	4	3
INDICES OF INDEP VARS IN REGRESSION...			
1	2	3	

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-18.3276	-18.2900	-30.4248	-29.8850
13.3911	-20.2275	-10.7373	24.3806
-14.8901	-.1650	-26.0498	-39.3538

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.0505	-1.3763	.0565	2.4134
-------	---------	-------	--------

MEDIANS...

-.0464	-1.3525	.2002	2.7556
--------	---------	-------	--------

INTERQUARTILE RANGES...

13.2813	15.3125	13.7500	25.5781
---------	---------	---------	---------

REGRESSION COEFFICIENTS, INTERCEPT...

1.4114	-.4568	.5350	1.6833
1.1608	-.7393	.7908	1.2927
1.0651	-.8805	.9055	1.0967
1.076	-.9453	.9553	1.0067
1.0126	-.9737	.9765	.9671
1.0066	-.9859	.9854	.9501
1.0041	-.9910	.9891	.9430
1.0031	-.9932	.9907	.9400
1.0027	-.9941	.9913	.9387
1.0025	-.9945	.9916	.9382

RESIDUAL VARIANCE = 1.3410

RESIDUAL STANDARD ERROR = 1.1580

COEFFICIENT OF DETERMINATION = .9949

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFIL1

DIMENSION (M) = 3

COEFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000
.5000	1.0000	.5000
.9000	.9000	1.0000

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

B0 (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000	-1.0000	1.0000
--------	---------	--------

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

KIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	50.0000	90.0000	140.0000
50.0000	150.0000	185.0000	85.0000
90.0000	185.0000	262.0000	167.0000
140.0000	85.0000	167.0000	223.0000

COEFFICIENT OF DETERMINATION = .9955

CORRELATION MATRIX OF X, Y...

1.0000	.4082	.5560	.9375
.4082	1.0000	.9332	.4648
.5560	.9332	1.0000	.6909
.9375	.4648	.6909	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:			
	-18.3276	-31.1084	-43.7901	-30.4319
OBSN NO	2:			
	13.3911	-17.6709	-13.7276	18.8337
OBSN NO	3:			
	-14.8901	-14.2334	-30.6551	-29.9006



PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFIL1

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            4            4            3

INDICES OF INDEP VARS IN REGRESSION...

1            2            3

DATA...

(FIRST 3 OBSNS)

-18.3276	-31.1084	-43.7901	-30.4319
13.3911	-17.6709	-13.7276	18.8337
-14.8901	-14.2334	-30.6651	-29.9006

MEANS FOR ALL VARIABLES...

.0505    -1.0521    -.5064    1.5265

STANDARD DEVIATIONS FOR ALL VARIABLES...

9.2663    12.8614    16.4787    14.1898

CORRELATION MATRIX...

1.0000			
.5608	1.0000		
.6908	.9437	1.0000	
.9460	.5593	.7603	1.0000

DETERMINANT = .1605E+00

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.6962			
-.0010	4.2709		
.0010	-3.4656	3.4656	

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.6964			
.9512	1.0000		
1.1719	.5640	-.5004	

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1995E-03			
-.8687E-07	.2608E-03		
.6780E-07	-.1652E-03	.1289E-03	

INDEX OF DEP VAR IN REGRESSION = 4

INDICES OF INDEP VARS IN REGRESSION...

1            2            3

REGRESSION COEFFICIENTS...

1.2840    -.2726    .6007

T-VALUES...

14.6515    -2.7213    8.5276

STD DEVS OF REG COEFS...

.0876    .1002    .0704

INTERCEPT = 1.4790

RESIDUAL STANDARD ERROR = 6.2037  
SAMPLE MULTIPLE CORRELATION COEFFICIENT = 1.0887  
SAMPLE COEFFICIENT OF DETERMINATION = 1.1853  
SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 23628.3906  
DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 3.0000  
MEAN SQUARE OF SSAR = 7876.1304  
SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSTR = -3694.6973  
DEGREES OF FREEDOM ASSOCIATED WITH SSTR = 96.0000  
MEAN SQUARE OF SSTR = -38.4864  
T-VALUE = -204.6469

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100            4            4            3

INDICES OF INDEP VARS IN REGRESSION...

1            2            3

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 10 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-18.3276	-31.1084	-43.7901	-30.4319
13.3911	-17.6709	-13.7276	18.8337
-14.8901	-14.2334	-30.6651	-29.9006

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.0505    -1.0521    -.5064    1.5265

MEDIANS...

-.0464    -.7959    .4599    2.5525

INTERQUARTILE RANGES...

13.2813    19.3750    22.0625    22.4375

REGRESSION COEFFICIENTS, INTERCEPT...

1.1720	-.4336	.5080	1.2684
1.0705	-.7462	.7845	1.0846
1.0178	-.9088	.9282	.9890
.9903	-.9933	1.0030	.9392
.9761	-1.0373	1.0418	.9134
.9686	-1.0602	1.0621	.9000
.9648	-1.0720	1.0726	.8930
.9628	-1.0782	1.0780	.8893
.9617	-1.0814	1.0809	.8874
.9612	-1.0831	1.0823	.8865

RESIDUAL VARIANCE = 1.4433

RESIDUAL STANDARD ERROR = 1.2014

COEFFICIENT OF DETERMINATION = .9928

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 3

COEFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000
0.0000	1.0000	0.0000
1.0000	1.0000	0.0000

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

B0 (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000	-1.0000	1.0000
--------	---------	--------

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	0.0000	100.0000	200.0000
0.0000	100.0000	100.0000	0.0000
100.0000	100.0000	200.0000	200.0000
200.0000	0.0000	200.0000	401.0000

COEFFICIENT OF DETERMINATION = .9975

CORRELATION MATRIX OF X, Y...

1.0000	0.0000	.7071	.9988
0.0000	1.0000	.7071	0.0000
.7071	.7071	1.0000	.7062
.9988	0.0000	.7062	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:			
	-18.3276	-15.0855	-33.4131	-36.0779
OBSN NO	2:			
	13.3911	-20.8667	-7.4756	28.2815
OBSN NO	3:			
	-14.8901	3.3520	-11.5381	-28.3591

PROGRAM MREG  
MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFIL1

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	4	4	3
INDICES OF INDEP VARS IN REGRESSION...			
1	2	3	

DATA...

(FIRST 3 OBSNS)  
-18.3276 -15.0855 -33.4131 -36.0779  
13.3911 -20.8667 -7.4756 28.2815  
-14.8901 3.3520 -11.5381 -28.3591

MEANS FOR ALL VARIABLES...  
.0505 -1.4574 -1.4069 1.0312

STANDARD DEVIATIONS FOR ALL VARIABLES...  
9.2663 10.3026 15.0878 18.5535

CORRELATION MATRIX...  
1.0000  
.1866 1.0000  
.7416 .7975 1.0000  
.9981 .1892 .7422 1.0000

DETERMINANT = .1582E-05

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

23011935\*9  
25585351\*528446825\*0  
746875\*004165908\*4361008056\*2

PRODUCT OF MATRIX AND INVERSE...  
SHOULD BE IDENTITY MATRIX. IF NOT, ROUND OFF ERRORS  
HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0312			
.0313	1.0000		
.0313	.0313	.9375	

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.2707E+02			
.2707E+02	.2707E+02		
-.2707E+02	-.2707E+02	.2707E+02	

INDEX OF DEP VAR IN REGRESSION = 4

INDICES OF INDEP VARS IN REGRESSION...

1	2	3
REGRESSION COEFFICIENTS...		
2.1900	.1688	-.1921

T-VALUES...  
.2764 .0213 -.0242

STD DEVS OF REG COEFS...  
7.9237 7.9237 7.9237

INTERCEPT = .8963

RESIDUAL STANDARD ERROR = 1.5229  
SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9967  
SAMPLE COEFFICIENT OF DETERMINATION = .9935  
SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 33856.4766  
DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 3.0000  
MEAN SQUARE OF SSAR = 11285.4922  
SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = 222.6484  
DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 96.0000  
MEAN SQUARE OF SSDR = 2.3193  
F-VALUE = 4865.9995

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	4	4	3
INDICES OF INDEP VARS IN REGRESSION...			
1	2	3	

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-33.4131	-36.0779
13.3911	-20.8667	-7.4756	28.2815
-14.8901	3.3520	-11.5381	-28.3591

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.0505	-1.4574	-1.4069	1.0312
-------	---------	---------	--------

MEDIANS...

-.0464	-1.8042	-2.1631	.9846
--------	---------	---------	-------

INTERQUARTILE RANGES...

13.2813	14.8438	20.9375	26.5625
---------	---------	---------	---------

REGRESSION COEFFICIENTS, INTERCEPT...

1.9979	.0036	.0031	.9399
1.9931	.0008	.0068	.9413
1.9891	-.0029	.0106	.9415
1.9852	-.0067	.0145	.9416
1.9814	-.0105	.0183	.9416
1.9775	-.0144	.0222	.9416
1.9737	-.0182	.0260	.9416
1.9698	-.0221	.0299	.9416
1.9660	-.0259	.0337	.9416
1.9621	-.0298	.0376	.9416

RESIDUAL VARIANCE = 1.3293  
RESIDUAL STANDARD ERROR = 1.1529  
COEFFICIENT OF DETERMINATION = .9961

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (N) = 6

COEFFICIENT MATRIX (C)...

COEFFT MATRIX IS IDENTITY MATRIX

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

B0 (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000 -1.0000 1.0000 1.0000 -1.0000 1.0000

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	0.0000	0.0000	0.0000	0.0000	0.0000	100.0000
0.0000	100.0000	0.0000	0.0000	0.0000	0.0000	-100.0000
0.0000	0.0000	100.0000	0.0000	0.0000	0.0000	100.0000
0.0000	0.0000	0.0000	100.0000	0.0000	0.0000	100.0000
0.0000	0.0000	0.0000	0.0000	100.0000	0.0000	-100.0000
0.0000	0.0000	0.0000	0.0000	0.0000	100.0000	100.0000
100.0000	-100.0000	100.0000	100.0000	-100.0000	100.0000	601.0000

COEFFICIENT OF DETERMINATION = .9983

CORRELATION MATRIX OF X, Y...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	.4079
0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	-.4079
0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	.4079
0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	.4079
0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	-.4079
0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	.4079
.4079	-.4079	.4079	.4079	-.4079	.4079	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST

FIRST 3 OBSNS PRINTED)

OBSN NO	1:					
	-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667 -55.1443
OBSN NO	2:					
	4.9927	-14.8901	3.3520	-20.2808	4.2114	6.8286 6.3284
OBSN NO	3:					
	16.4380	-6.5698	8.5474	11.7895	3.1567	2.6489 44.8635



PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6			
INDICES OF INDEP VARS IN REGRESSION...	1	2	3	4	5	6

DATA...  
 (FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-55.1443
4.9927	-14.8901	3.3520	-20.2808	4.2114	6.8286	6.3284
16.4380	-6.5698	8.5474	11.7895	3.1567	2.6489	44.8635

MEANS FOR ALL VARIABLES...

-.0472	.4075	-.5128	-.9081	-.1784	.8763	.1547
--------	-------	--------	--------	--------	-------	-------

STANDARD DEVIATIONS FOR ALL VARIABLES...

10.2626	9.8716	10.8903	10.3570	9.4954	9.5353	27.1463
---------	--------	---------	---------	--------	--------	---------

CORRELATION MATRIX...

1.0000						
-.0039	1.0000					
.2032	-.0813	1.0000				
.0771	.1454	-.0221	1.0000			
-.0281	-.0781	.0677	.0364	1.0000		
.2311	-.0366	.1645	-.0104	-.1342	1.0000	
.5834	-.3250	.5361	.3263	-.3464	.5620	1.0000

DETERMINANT = .8320E+00

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.0972						
-.0047	1.0358					
-.1901	.0698	1.0747				
-.0910	-.1515	.0302	1.0317			
.0170	.0866	-.0928	-.0521	1.0360		
-.2211	.0376	-.1424	.0143	.1530	1.0966	

PRODUCT OF MATRIX AND INVERSE...  
 SHOULD BE IDENTITY MATRIX. IF NOT, ROUND OFF ERRORS  
 HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000						
.0000	1.0000					
.0000	-.000*	1.0000				
.0000	.000*	-.000*	1.0000			
.0000	-.000*	-.000*	-.000*	1.0000		
-.0000	-.000*	-.000*	.000*	0.0000	1.0000	

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1052E-03						
-.4657E-06	.1074E-03					
-.1718E-04	.6557E-05	.9153E-04				
-.8649E-05	-.1496E-04	.2700E-05	.9716E-04			
.1760E-05	.9337E-05	-.0068E-05	-.5350E-05	.1161E-03		

-.2283E-04 .4034E-05 -.1385E-04 .1458E-05 .1707E-04 .1218E-03

INDEX OF DEP VAR IN REGRESSION = 7  
INDICES OF INDEP VARS IN REGRESSION...  
1 2 3 4 5 6  
REGRESSION COEFFICIENTS...  
1.0048 -.9907 1.0082 .9830 -1.0231 .9974  
T-VALUES...  
98.7807 -96.4196 106.2686 100.5752 -95.7703 91.1315  
STD DEVS OF REG COEFS...  
.0102 .0103 .0095 .0098 .0107 .0109  
INTERCEPT = .9589  
RESIDUAL STANDARD ERROR = .9916  
SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9994  
SAMPLE COEFFICIENT OF DETERMINATION = .9987  
SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 72863.9766  
DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 6.0000  
MEAN SQUARE OF SSAR = 12143.9961  
SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSSR = 91.4453  
DEGREES OF FREEDOM ASSOCIATED WITH SSSR = 93.0000  
MEAN SQUARE OF SSSR = .9833  
F-VALUE = 12350.4600

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6			
INDICES OF INDEP VARS IN REGRESSION...						
1	2	3	4	5	6	

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-55.1443
4.9927	-14.8901	3.3520	-20.2808	4.2114	6.8286	6.3284
16.4380	-6.5698	8.5474	11.7895	3.1567	2.6489	44.8635

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

-.0472	.4075	-.5128	-.9081	-.1784	.8763	.1547
--------	-------	--------	--------	--------	-------	-------

MEDIANS...

.3833	.4614	-1.1401	-2.4292	-.7105	.8911	-.9060
-------	-------	---------	---------	--------	-------	--------

INTERQUARTILE RANGES...

14.3750	13.9843	15.3906	13.3203	11.8750	11.6406	31.4258
---------	---------	---------	---------	---------	---------	---------

REGRESSION COEFFICIENTS, INTERCEPT...

1.4869	-.7377	1.0480	.8243	-.9862	.8140	.9222
1.0289	-.9806	1.0338	.9833	-1.0185	.9758	.9892
.9981	-.9968	1.0091	.9909	-1.0122	.9923	.9752
1.0007	-.9944	1.0049	.9898	-1.0103	.9926	.9713
1.0016	-.9937	1.0046	.9894	-1.0101	.9924	.9708
1.0018	-.9937	1.0046	.9894	-1.0101	.9923	.9708
1.0018	-.9936	1.0046	.9894	-1.0101	.9923	.9708
1.0018	-.9936	1.0046	.9894	-1.0101	.9923	.9708
1.0018	-.9936	1.0046	.9894	-1.0101	.9923	.9708

RESIDUAL VARIANCE = 1.0136

RESIDUAL STANDARD ERROR = 1.0068

COEFFICIENT OF DETERMINATION = .9986

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFIL1

DIMENSION (M) = 6

COEFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
.1000	1.0000	.1000	.1000	.1000	.1000
.2000	.2000	1.0000	.2000	.2000	.2000
.3000	.3000	.3000	1.0000	.3000	.3000
.4000	.4000	.4000	.4000	1.0000	.4000
.5000	.5000	.5000	.5000	.5000	1.0000

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

INTERCEPT = 1.0000

PARAMETER VECTOR (B)...

1.0000	-1.0000	1.0000	1.0000	-1.0000	1.0000
--------	---------	--------	--------	---------	--------

STD DEV (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	10.0000	20.0000	30.0000	40.0000	50.0000	150.0000
10.0000	105.0000	38.0000	52.0000	66.0000	80.0000	9.0000
20.0000	38.0000	120.0000	74.0000	92.0000	110.0000	194.0000
30.0000	52.0000	74.0000	145.0000	118.0000	140.0000	219.0000
40.0000	66.0000	92.0000	118.0000	180.0000	170.0000	174.0000
50.0000	80.0000	110.0000	140.0000	170.0000	225.0000	275.0000
150.0000	9.0000	194.0000	219.0000	174.0000	275.0000	656.0001

COEFFICIENT OF DETERMINATION = .9985

CORRELATION MATRIX OF X, Y...

1.0000	.0976	.1826	.2491	.2981	.3333	.5857
.0976	1.0000	.3385	.4214	.4801	.5205	.0343
.1826	.3385	1.0000	.5610	.6260	.6694	.6914
.2491	.4214	.5610	1.0000	.7304	.7751	.7101
.2981	.4801	.6260	.7304	1.0000	.8447	.5064
.3333	.5205	.6694	.7751	.8447	1.0000	.7153
.5857	.0343	.6914	.7101	.5064	.7153	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:						
	-18.3276	-19.4602	-22.7412	-20.6082	-15.4986	-39.8499	-56.2681
OBSN NO	2:						
	4.9927	-14.9797	-.4756	-18.9324	-3.7876	-4.4734	.6304
OBSN NO	3:						
	16.4380	-2.3118	14.0400	19.0559	16.2983	19.3293	16.9038

PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6		
INDICES OF INDEP VARS IN REGRESSION...					
1	2	3	4	5	6

DATA...

(FIRST 3 OBSNS)

-18.3276	-19.4602	-22.7412	-20.6082	-15.4986	-39.8499	-66.2681
4.9927	-14.9797	-.4756	-18.9324	-3.7876	-4.4788	.6304
16.4380	-2.3118	14.0400	19.0559	16.2983	19.3298	56.9038

MEANS FOR ALL VARIABLES...

-.0472	.3305	-.4828	-.7445	-.2521	.2568	-.1204
--------	-------	--------	--------	--------	-------	--------

STANDARD DEVIATIONS FOR ALL VARIABLES...

10.2626	10.1399	12.4193	13.1373	13.6447	16.1560	30.3060
---------	---------	---------	---------	---------	---------	---------

CORRELATION MATRIX...

1.0000						
.1462	1.0000					
.3869	.2952	1.0000				
.3890	.4894	.5968	1.0000			
.4330	.4555	.7089	.7764	1.0000		
.5377	.5005	.7596	.8064	.8517	1.0000	
.7104	.1115	.7875	.7222	.6205	.8239	1.0000

DETERMINANT = .1736E-01

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.4607					
.2377	1.4447				
.1210	.3079	2.5382			
.1233	-.2822	.2253	3.2263		
.0397	-.1100	-.6679	-1.0620	4.1532	
-1.1296	-.7634	-1.7600	-1.7933	-2.1397	6.5947

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
 HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000					
0.0000	1.0000				
.0000	.000*	1.0000			
.0000	.000*	.000*	1.0000		
.0000	.000*	.0000	.0000	1.0000	
.0000	.000*	.000*	.000*	.000*	1.0000

INVERSE OF THE PART OF THE CORR'S PROD MATRIX USED IN REGRESSION...

.1401E-03					
.2307E-04	.1419E-03				
.9592E-05	.2470E-04	.1662E-03			
.9236E-05	-.2140E-04	.1395E-03	.1888E-03		
.2866E-05	-.8030E-05	-.3981E-04	-.5985E-04	.2253E-03	

-.6882E-04 -.4707E-04 -.8860E-04 -.8535E-04 -.9804E-04 .2552E-03

INDEX OF DEP VAR IN REGRESSION = 7  
INDICES OF INDEP VARS IN REGRESSION...  
1 2 3 4 5 6  
REGRESSION COEFFICIENTS...  
1.0112 -.9825 1.0182 .9849 -1.0278 1.0077  
T-VALUES...  
86.1889 -83.2002 79.6708 72.3065 -69.0716 63.6332  
STD DEVS OF REG COEFS...  
.0117 .0118 .0128 .0136 .0149 .0158  
INTERCEPT = .9589  
RESIDUAL STANDARD ERROR = .9913  
SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9995  
SAMPLE COEFFICIENT OF DETERMINATION = .9990  
SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 90835.6094  
DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 6.0000  
MEAN SQUARE OF SSAR = 15139.2686  
SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSTR = 91.3828  
DEGREES OF FREEDOM ASSOCIATED WITH SSTR = 93.0000  
MEAN SQUARE OF SSTR = .9826  
F-VALUE = 15407.1855

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6		
INDICES OF INDEP VARS IN REGRESSION...					
1	2	3	4	5	6

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 2 ITERATIONS

DATA...

(FIRST	3 OBSNS)					
-18.3276	-19.4602	-22.7412	-20.6082	-15.4986	-39.8499	-66.2681
4.9927	-14.9797	-.4756	-18.9324	-3.7876	-4.4788	.6304
16.4380	-2.3118	14.0400	19.0559	16.2983	19.3298	56.9038

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

-.0472	.3305	-.4828	-.7445	-.2521	.2568	-.1204
--------	-------	--------	--------	--------	-------	--------

MEDIANS...

.3833	-.2219	.4932	-2.0652	-.4595	-1.3928	-.9712
-------	--------	-------	---------	--------	---------	--------

INTERQUARTILE RANGES...

14.3750	14.4844	17.7812	19.3985	20.6484	21.8750	41.3985
---------	---------	---------	---------	---------	---------	---------

REGRESSION COEFFICIENTS, INTERCEPT...

1.4147	-.4632	1.4790	.7226	-.5361	.2250	1.1584
1.1234	-.8886	1.3225	1.0682	-.6632	.4241	1.3838
1.0834	-.9831	1.1404	1.0977	-.6860	.5760	1.3025
1.0602	-.9976	1.0535	1.0454	-.7153	.7003	1.1860
1.0376	-.9978	1.0154	.9961	-.7622	.8017	1.0921
1.0189	-.9972	.9979	.9649	-.8154	.8810	1.0255
1.0054	-.9973	.9896	.9483	-.8653	.9404	.9807
.9963	-.9979	.9858	.9411	-.9072	.9829	.9518
.9904	-.9987	.9845	.9390	-.9399	1.0122	.9338
.9867	-.9994	.9844	.9396	-.9641	1.0318	.9231

RESIDUAL VARIANCE = 1.3218

RESIDUAL STANDARD ERROR = 1.1497

COEFFICIENT OF DETERMINATION = .9986

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 6

EFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
.1000	1.0000	.1000	.1000	.1000	.1000
.3000	.3000	1.0000	.3000	.3000	.3000
.5000	.5000	.5000	1.0000	.5000	.5000
.7000	.7000	.7000	.7000	1.0000	.7000
.9000	.9000	.9000	.9000	.9000	1.0000

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

(INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000	-1.0000	1.0000	1.0000	-1.0000	1.0000
--------	---------	--------	--------	---------	--------

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	10.0000	30.0000	50.0000	70.0000	90.0000	190.0000
10.0000	105.0000	52.0000	80.0000	108.0000	136.0000	65.0000
30.0000	52.0000	145.0000	140.0000	184.0000	228.0000	307.0000
50.0000	80.0000	140.0000	225.0000	260.0000	320.0000	395.0000
70.0000	108.0000	184.0000	260.0000	345.0000	412.0000	473.0000
90.0000	136.0000	228.0000	320.0000	412.0000	505.0000	595.0000
190.0000	65.0000	307.0000	395.0000	473.0000	595.0000	950.0001

COEFFICIENT OF DETERMINATION = .9989

CORRELATION MATRIX OF X, Y...

1.0000	.0976	.2491	.3333	.3769	.4005	.6164
.0976	1.0000	.4214	.5205	.5674	.5906	.2058
.2491	.4214	1.0000	.7751	.8227	.8426	.8272
.3333	.5205	.7751	1.0000	.9332	.9493	.8544
.3769	.5674	.8227	.9332	1.0000	.9871	.8262
.4005	.5906	.8426	.9493	.9871	1.0000	.8590
.6164	.2058	.8272	.8544	.8262	.8590	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO 1:	-18.3276	-19.4602	-27.2527	-31.5296	-37.1658	-55.0364	-75.2203
OBSN NO 2:	4.9927	-14.9797	-2.3894	-18.0335	-9.7869	-13.5247	-3.4312
OBSN NO 3:	16.4380	-2.3118	16.7864	23.9001	26.1545	32.6745	67.9829



PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6			
INDICES OF INDEP VARS IN REGRESSION...	1	2	3	4	5	6

DATA...  
 (FIRST 3 OBSNS)

-18.3276	-19.4602	-27.2527	-31.5296	-37.1658	-55.0364	-75.2203
4.9927	-14.9797	-2.3894	-18.0335	-9.7869	-13.5247	-3.4312
16.4380	-2.3118	16.7864	23.9001	26.1545	32.6745	67.9829

MEANS FOR ALL VARIABLES...

-.0472	.3305	-.4678	-.6354	-.3074	-.2388	-.4367
--------	-------	--------	--------	--------	--------	--------

STANDARD DEVIATIONS FOR ALL VARIABLES...

10.2626	10.1399	13.7158	16.5319	19.9104	24.6741	36.3758
---------	---------	---------	---------	---------	---------	---------

CORRELATION MATRIX...

1.0000						
.1462	1.0000					
.4448	.3738	1.0000				
.4830	.5431	.8001	1.0000			
.5294	.5416	.8655	.9471	1.0000		
.5623	.5633	.8792	.9577	.9886	1.0000	
.7218	.2374	.8849	.8695	.8761	.9041	1.0000

DETERMINANT = .4848E-03

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.5937					
.4234	1.6643				
.4623	.7992	4.7397			
.6213	.0557	1.4482	11.2243		
-.2875	-.3710	-2.8464	-6.8279	23.9330	
.2875	.3710	2.8464	6.8279	-13.8066	13.8066

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
 HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

2.1708					
1.1729	2.0632				
1.8305	1.6593	6.4417			
1.9940	1.8075	5.9277	12.7853		
2.0584	1.8659	6.1191	12.1657	1.0000	
2.1395	1.9358	6.2965	12.4811	.4413	-9.8914

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1528E-03					
.4110E-04	.1635E-03				
.3318E-04	.5805E-04	.2545E-03			
.3699E-04	.3358E-05	.6451E-04	.4148E-03		
-.1421E-04	-.1856E-04	-.1053E-03	-.2095E-03	.6098E-03	

.1147E-04 .1498E-04 .8496E-04 .1691E-03 -.2839E-03 .2291E-03

INDEX OF DEP VAR IN REGRESSION = 7

INDICES OF INDEP VARS IN REGRESSION...

	1	2	3	4	5	6
--	---	---	---	---	---	---

REGRESSION COEFFICIENTS...

	1.1545	-.8600	1.2206	1.2504	-.4873	13.4723
--	--------	--------	--------	--------	--------	---------

T-VALUES...

	.8738	-.6293	.7159	.5745	-.1847	8.3293
--	-------	--------	-------	-------	--------	--------

STD DEVS OF REG COEFS...

	1.3212	1.3665	1.7049	2.1767	2.6391	1.6175
--	--------	--------	--------	--------	--------	--------

INTERCEPT = 4.3353

RESIDUAL STANDARD ERROR = 106.8683

SAMPLE MULTIPLE CORRELATION COEFFICIENT = 3.0180

SAMPLE COEFFICIENT OF DETERMINATION = 9.1081

SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 1193133.7500

DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 6.0000

MEAN SQUARE OF SSAR = 198855.6250

SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSCR = -1062137.0000

DEGREES OF FREEDOM ASSOCIATED WITH SSCR = 93.0000

MEAN SQUARE OF SSCR = -11420.8281

F-VALUE = -17.4117

PROGRAM ITEREG

ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6			
INDICES OF INDEP VARS IN REGRESSION...	1	2	3	4	5	6

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 20 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-18.3276	-19.4602	-27.2527	-31.5296	-37.1658	-55.0364	-75.2203
4.9927	-14.9797	-2.3894	-18.0335	-9.7869	-13.5247	-3.4312
16.4380	-2.3118	16.7864	23.9001	26.1545	32.6745	67.9829

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

-.0472	.3305	-.4678	-.6354	-.3074	-.2388	-.4367
--------	-------	--------	--------	--------	--------	--------

MEDIANS...

.3833	-.2219	.4036	-2.3108	.6584	-1.5403	-2.3765
-------	--------	-------	---------	-------	---------	---------

INTERQUARTILE RANGES...

14.3750	14.4844	17.6015	24.9023	27.1836	33.7930	51.1875
---------	---------	---------	---------	---------	---------	---------

REGRESSION COEFFICIENTS, INTERCEPT...

1.1139	-.9267	1.1271	1.1911	-.4600	.3392	1.1458
1.0836	-.9421	1.0840	1.1023	-.4941	.4559	1.0903
1.0596	-.9564	1.0590	1.0516	-.5782	.5773	1.0531
1.0402	-.9690	1.0398	1.0164	-.6630	.6850	1.0247
1.0242	-.9796	1.0242	.9890	-.7376	.7763	1.0016
1.0109	-.9885	1.0114	.9666	-.8006	.8529	.9826
.9998	-.9960	1.0007	.9480	-.8535	.9168	.9667
.9905	-1.0022	.9918	.9326	-.8976	.9700	.9536
.9828	-1.0073	.9844	.9198	-.9343	1.0144	.9426
.9764	-1.0116	.9783	.9091	-.9649	1.0514	.9335

PRINT AFTER EVERY 20 ITERATIONS

.9710	-1.0152	.9731	.9002	-.9904	1.0822	.9258
.9666	-1.0182	.9688	.8927	-1.0117	1.1079	.9195
.9629	-1.0207	.9653	.8866	-1.0294	1.1293	.9142

RESIDUAL VARIANCE = 1.2873

RESIDUAL STANDARD ERROR = 1.1346

COEFFICIENT OF DETERMINATION = .9990

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 6

COEFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	1.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	1.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
0.0000	0.0000	0.0000	1.0000	1.0000	0.0000

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

MEAN (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000	-1.0000	1.0000	1.0000	-1.0000	1.0000
--------	---------	--------	--------	---------	--------

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	0.0000	0.0000	0.0000	0.0000	0.0000	100.0000
0.0000	100.0000	0.0000	0.0000	0.0000	0.0000	-100.0000
0.0000	0.0000	100.0000	0.0000	0.0000	0.0000	100.0000
0.0000	0.0000	0.0000	100.0000	0.0000	100.0000	200.0000
0.0000	0.0000	0.0000	0.0000	100.0000	100.0000	0.0000
0.0000	0.0000	0.0000	100.0000	100.0000	200.0000	200.0000
100.0000	-100.0000	100.0000	200.0000	0.0000	200.0000	701.0000

COEFFICIENT OF DETERMINATION = .9986

CORRELATION MATRIX OF X, Y...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	.3777
0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	-.3777
0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	.3777
0.0000	0.0000	0.0000	1.0000	0.0000	.7071	.7554
0.0000	0.0000	0.0000	0.0000	1.0000	.7071	0.0000
0.0000	0.0000	0.0000	.7071	.7071	1.0000	.5341
.3777	-.3777	.3777	.7554	0.0000	.5341	1.0000

OBSERVATIONS...

DEPENDENT VARIABLE LAST

(FIRST 3 OBSNS PRINTED)

OBSN NO	1:						
	-18.3276	-15.0855	-13.7183	-4.2261	13.3911	9.1650	-25.1126
OBSN NO	2:						
	4.9927	-14.8901	3.3520	-20.2808	4.2114	-16.0693	-16.5696
OBSN NO	3:						
	16.4380	-6.5698	8.5474	11.7895	3.1567	14.9463	57.1609

PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6		
INDICES OF INDEP VARS IN REGRESSION...					
1	2	3	4	5	6

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-4.2261	13.3911	9.1650	-25.1126
4.9927	-14.8901	3.3520	-20.2808	4.2114	-16.0693	-16.5696
16.4380	-6.5698	8.5474	11.7895	3.1567	14.9463	57.1609

MEANS FOR ALL VARIABLES...

-.0472	.4075	-.5128	-.9081	-.1784	-1.0865	-1.8081
--------	-------	--------	--------	--------	---------	---------

STANDARD DEVIATIONS FOR ALL VARIABLES...

10.2626	9.8716	10.8903	10.3570	9.4954	14.3035	27.7991
---------	--------	---------	---------	--------	---------	---------

CORRELATION MATRIX...

1.0000						
-.0039	1.0000					
.2032	-.0813	1.0000				
.0771	.1454	-.0221	1.0000			
-.0281	-.0781	.0677	.0364	1.0000		
.0372	.0534	.0289	.7483	.6902	1.0000	
.5096	-.2773	.4819	.7073	.0629	.5539	1.0000

DETERMINANT = .8157E-07

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.0526					
.0029	1.0345				
-.2188	.0747	1.0562			
4.7693	-.4235	.665658642455*0			
4.5012	-.1675	.507953763765*0	49290965*0		
-6.7083	.3750	-.87508098794*0	07425025*5011184811*0		

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS  
 HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000					
0.0000	1.0000				
-.0000	-.000*	1.0000			
0.0000	.0000	-.000*	1.0000		
0.0000	-.0000	0.0000	.5000	1.0000	
.0000	-.000*	0.0000	.5000	.5000	0.0000

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1010E-03					
.2901E-06	.1072E-03				
-.1978E-04	.7016E-05	.8996E-04			
.4532E-03	-.4184E-04	.5961E-04	.5522E+03		
.4666E-03	-.1805E-04	.4961E-04	.5522E+03	.5522E+03	

-.4616E-03 .2683E-04 -.5674E-04 -.5522E+03 -.5522E+03 .5522E+03

INDEX OF DEP VAR IN REGRESSION = 7  
INDICES OF INDEP VARS IN REGRESSION...  
1 2 3 4 5 6  
REGRESSION COEFFICIENTS...  
1.0043 -.9906 1.0079 1.3420 -.7319 .9718  
T-VALUES...  
11.5150 -11.0202 12.2417 .0066 -.0036 .0048  
STD DEVS OF REG COEFS...  
.0872 .0899 .0823 203.9857 203.9858 203.9859  
INTERCEPT = 1.3038  
RESIDUAL STANDARD ERROR = 8.6805  
SAMPLE MULTIPLE CORRELATION COEFFICIENT = 1.0448  
SAMPLE COEFFICIENT OF DETERMINATION = 1.0916  
SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 83513.8203  
DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 6.0000  
MEAN SQUARE OF SSAR = 13918.9697  
SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = -7007.7031  
DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 93.0000  
MEAN SQUARE OF SSDR = -75.3516  
F-VALUE = -184.7202

PROGRAM ITEREG  
 ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	7	7	6			
INDICES OF INDEP VARS IN REGRESSION...						
1	2	3	4	5	6	

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-4.2261	13.3911	9.1650	-25.1126
4.9927	-14.8901	3.3520	-20.2808	4.2114	-16.0693	-16.5696
16.4380	-6.5698	8.5474	11.7895	3.1567	14.9463	57.1609

MEANS FOR VARIABLES IN REGRESSION,  
 DEPENDENT VARIABLE LAST...

-.0472	.4075	-.5128	-.9081	-.1784	-1.0865	-1.8081
--------	-------	--------	--------	--------	---------	---------

MEDIANS...

.3833	.4614	-1.1401	-2.4292	-.7105	-2.5537	-3.3587
-------	-------	---------	---------	--------	---------	---------

INTERQUARTILE RANGES...

14.3750	13.9843	15.3906	13.3203	11.8750	16.3281	37.7148
---------	---------	---------	---------	---------	---------	---------

REGRESSION COEFFICIENTS, INTERCEPT...

1.1909	-.5521	.9881	1.8828	.0843	.0026	.7074
.9934	-.9575	1.0006	1.9690	-.0044	.0051	.9350
.9997	-.9880	1.0031	1.9819	-.0136	.0034	.9571
.9999	-.9919	1.0036	1.9856	-.0127	.0010	.9599
.9999	-.9923	1.0037	1.9883	-.0104	-.0014	.9603
.9999	-.9924	1.0037	1.9907	-.0080	-.0039	.9603
.9999	-.9924	1.0037	1.9932	-.0056	-.0064	.9603
.9999	-.9924	1.0037	1.9957	-.0031	-.0088	.9603
.9999	-.9924	1.0037	1.9981	-.0006	-.0113	.9603
.9999	-.9924	1.0037	2.0006	.0018	-.0137	.9603

RESIDUAL VARIANCE = 1.0021

RESIDUAL STANDARD ERROR = 1.0011

COEFFICIENT OF DETERMINATION = .9987

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 10

COEFFICIENT MATRIX (C)...  
COEFFT MATRIX IS IDENTITY MATRIX

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

B0 (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...  
1.0000 1.0000 -1.0000 1.0000 1.0000 -1.0000 1.0000 1.0000  
-1.0000 1.0000

SIG (MODEL ERROR STD DEV) :: 1.0000

NO OF OBSNS = 100  
SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...  
100.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 100.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 100.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 100.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 100.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 -100.0000 100.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 100.0000 0.0000 100.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 0.0000 0.0000 100.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 100.0000 0.0000 0.0000 100.0000 0.0000 0.0000  
0.0000 0.0000 -100.0000 0.0000 0.0000 0.0000 100.0000 0.0000  
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 100.0000  
0.0000 0.0000 100.0000 0.0000 0.0000 0.0000 0.0000 100.0000  
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
100.0000 0.0000 -100.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 100.0000 100.0000 100.0000 100.0000 -100.0000 100.0000 100.0000  
-100.0000 100.0000 100.0000 100.0000 100.0000

COEFFICIENT OF DETERMINATION = .9990

CORRELATION MATRIX OF X, Y...  
1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 .3161 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 .3161 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 -.3161 1.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 0.0000 1.0000 0.0000 0.0000 0.0000 0.0000  
0.0000 0.0000 .3161 0.0000 1.0000 0.0000 0.0000 0.0000



0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
0.0000	0.0000	.3161					
0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000
0.0000	0.0000	-.3161					
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
0.0000	0.0000	.3161					
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
0.0000	0.0000	.3161					
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1.0000	0.0000	-.3161					
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	1.0000	.3161					
.3161	.3161	-.3161	.3161	.3161	-.3161	.3161	.3161
-.3161	.3161	1.0000					

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
 (FIRST 3 OBSNS PRINTED)

OBSN NO	1:						
-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-6.9995	4.9927
-14.8901	3.3520	25.5442					
OBSN NO	2:						
4.2114	6.8286	-2.4292	16.4380	-6.5698	8.5474	11.7895	3.1567
2.6489	10.2661	38.9543					
OBSN NO	3:						
-10.1245	-8.1323	-8.0151	10.2270	6.5942	-8.9136	3.7036	4.4458
13.3130	20.3052	30.1770					

PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
INDICES OF INDEP VARS IN REGRESSION...								
1	2	3	4	5	6	7	8	
9	10							

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-6.9995	4.9927
-14.8901	3.3520	25.5442					
4.2114	6.8286	-2.4292	16.4380	-6.5698	8.5474	11.7895	3.1567
2.6489	10.2661	38.9543					
-10.1245	-8.1323	-8.0151	10.2270	6.5942	-8.9136	3.7036	4.4458
13.3130	20.3052	30.1770					

MEANS FOR ALL VARIABLES...

.1809	-.0144	-.8847	-.0300	-.4503	-.2456	.0841	-.6612
-.2816	.4231	1.9894					

STANDARD DEVIATIONS FOR ALL VARIABLES...

11.2156	10.1579	9.8358	10.7963	9.8302	9.7790	10.1817	8.9774
8.7924	10.4552	32.0968					

CORRELATION MATRIX...

1.0000							
.0715	1.0000						
.0254	.0648	1.0000					
-.0757	.2606	.1868	1.0000				
-.0717	.1217	.2235	.1323	1.0000			
.0586	.0191	-.1219	-.0018	-.0893	1.0000		
.0552	.0508	.1240	.1180	-.0356	.1012	1.0000	
-.0626	-.0388	-.0055	.0173	.1513	-.0569	.0051	1.0000
-.0012	.1444	.0992	.0022	-.1068	-.0386	.1414	-.1326
1.0000							
-.0664	-.0577	.0919	.1316	.1494	-.1056	-.0247	.0684
-.0353	1.0000						
.2760	.3818	-.0691	.4557	.4252	-.2753	.2650	.3824
-.2871	.3955	1.0000					

DETERMINANT = .6387E+00

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.0337							
-.1058	1.1368						
-.0619	.0245	1.1325					
.1087	-.3037	-.1702	1.1582				
.0725	-.1531	-.2427	-.0484	1.1465			
-.0457	-.0205	.1248	-.0244	.0498	1.0495		
-.0552	.0048	-.1244	-.1179	.0534	-.1206	1.0673	
.0485	.0342	.0384	-.0012	-.1457	.0527	-.0131	1.0187
.0438	-.1755	-.1156	.0636	.1394	.0635	-.1428	.1210
1.0906							

.0352	.1086	-.0447	-.1430	-.1252	.0866	.0271	-.0393
.0076	1.0630						

PRODUCT OF MATRIX AND INVERSE...

SHOULD BE IDENTITY MATRIX. IF NOT, ROUNDOFF ERRORS HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000								
.0000	1.0000							
.0000	-.000*	1.0000						
-.0000	.000*	-.000*	1.0000					
-.0000	-.000*	-.000*	-.000*	1.0000				
.0000	-.000*	-.000*	.000*	-.000*	1.0000			
.0000	-.000*	-.000*	-.000*	.000*	-.000*	1.0000		
.0000	.000*	-.000*	.000*	.000*	-.000*	.000*	1.0000	
-.0000	.000*	.000*	-.000*	-.000*	-.000*	.000*	-.000*	
1.0000								
0.0000	-.0000	-.000*	0.0000	-.0000	0.0000	-.0000	0.0000	
-.0000	1.0000							

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.8301E-04								
-.9382E-05	.1113E-03							
-.5672E-05	.2478E-05	.1182E-03						
.9068E-05	-.2797E-04	-.1619E-04	.1004E-03					
.6643E-05	-.1549E-04	-.2536E-04	-.4608E-05	.1198E-03				
-.4213E-05	-.2089E-05	.1311E-04	-.2332E-05	.5233E-05	.1109E-03			
-.4879E-05	.4724E-06	-.1255E-04	-.1083E-04	.5385E-05	-.1223E-04	.1040E-03		
.4861E-05	.3783E-05	.4391E-05	-.1294E-06	-.1668E-04	.6063E-05	-.4758E-03		
.4489E-05	-.1984E-04	-.1350E-04	.6772E-05	.1629E-04	.7464E-05	-.1612E-03		
.1425E-03								
.3029E-05	.1033E-04	-.4388E-05	-.1279E-04	-.1231E-04	.8555E-05	.2573E-03		
.8335E-06	.9823E-04							

INDEX OF DEP VAR IN REGRESSION = 11

INDICES OF INDEP VARS IN REGRESSION...

1	2	3	4	5	6	7	8
9	10						

REGRESSION COEFFICIENTS...

.9941	.9888	-.9915	.9903	.9758	-.9970	.9946	1.0221
-1.0133	.9899						

T-VALUES...

99.2642	85.2707	-82.9539	89.9310	81.0920	-86.1448	88.7299	81.1131
-77.2245	90.8634						

STD DEVS OF REG COEFS...

.0100	.0116	.0120	.0110	.0120	.0116	.0112	.0126
.0131	.0109						

INTERCEPT = 1.0589

RESIDUAL STANDARD ERROR = 1.0992

SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9995

SAMPLE COEFFICIENT OF DETERMINATION = .9989

SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 101882.7891

DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 10.0000

MEAN SQUARE OF SSAR = 10188.2793

SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSCR = 107.5312

DEGREES OF FREEDOM ASSOCIATED WITH SSCR = 89.0000

MEAN SQUARE OF SSCR = 1.2082

F-VALUE = 8432.4961

PROGRAM ITEREG  
 ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
INDICES OF INDEP VARS IN REGRESSION...								
1	2	3	4	5	6	7	8	
9	10							

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 1 ITERATIONS

DATA...

FIRST	3 OBSNS)							
-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-6.9995	4.9927	
-14.8901	3.3520	25.5442						
4.2114	6.8286	-2.4292	16.4380	-6.5698	8.5474	11.7895	3.1567	
2.6489	10.2661	38.9543						
-10.1245	-8.1323	-8.0151	10.2270	6.5942	-8.9136	3.7036	4.4458	
13.3130	20.3052	30.1770						

MEANS FOR VARIABLES IN REGRESSION,  
 DEPENDENT VARIABLE LAST...

.1809	-.0144	-.8847	-.0300	-.4503	-.2456	.0841	-.6612	
-.2816	.4231	1.9894						

MEDIANS...

.1880	.8130	-.7495	.5786	-.8276	-.2026	-1.4136	-1.2183	
-.0464	.1489	.3645						

INTERQUARTILE RANGES...

15.4688	14.0235	12.7734	16.7969	13.5547	10.8985	15.6250	10.9766	
12.5390	14.8828	38.5937						

REGRESSION COEFFICIENTS, INTERCEPT...

.7239	1.1009	-.1421	1.1834	1.3138	-1.3654	.8399	1.2278	
-.9845	.6267	2.2392						
1.0218	.7820	-1.0716	1.0741	1.0202	-1.0831	1.0582	1.0402	
-.9638	.9634	1.0131						
1.0310	.9516	-1.0316	.9993	.9905	-.9890	1.0075	1.0262	
-1.0061	.9872	1.0298						
1.0052	.9900	-.9849	.9904	.9751	-.9918	1.0021	1.0175	
-1.0234	.9835	1.0598						
.9931	.9934	-.9748	.9903	.9713	-.9948	1.0039	1.0147	
-1.0269	.9825	1.0652						
.9971	.9934	-.9736	.9903	.9709	-.9953	1.0046	1.0141	
-1.0273	.9824	1.0656						
.9970	.9934	-.9735	.9902	.9709	-.9953	1.0047	1.0141	
-1.0273	.9824	1.0656						
.9969	.9934	-.9735	.9901	.9708	-.9953	1.0047	1.0141	
-1.0274	.9824	1.0656						
.9969	.9934	-.9735	.9901	.9708	-.9953	1.0047	1.0140	
-1.0274	.9824	1.0656						
.9969	.9935	-.9735	.9901	.9708	-.9953	1.0047	1.0140	

-1.0274      .9824      1.0656

RESIDUAL VARIANCE =                      1.2819  
RESIDUAL STANDARD ERROR =                      1.1322  
COEFFICIENT OF DETERMINATION =                      .9988

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFILL

DIMENSION (M) = 10

COEFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000							
0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000							
.1000	.1000	1.0000	.1000	.1000	.1000	.1000	.1000	.1000
.1000	.1000							
0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000							
.2000	.2000	.2000	.2000	1.0000	.2000	.2000	.2000	.2000
.2000	.2000							
0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
0.0000	0.0000							
.3000	.3000	.3000	.3000	.3000	.3000	1.0000	.3000	.3000
.3000	.3000							
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	1.0000
0.0000	0.0000							
.4000	.4000	.4000	.4000	.4000	.4000	.4000	.4000	.4000
1.0000	.4000							
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	1.0000							

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

INTERCEPT = 1.0000

PARAMETER VECTOR (B)...

1.0000	1.0000	-1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000
-1.0000	1.0000						

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	0.0000	10.0000	0.0000	20.0000	0.0000	30.0000	0.0000
40.0000	0.0000	100.0000					
0.0000	100.0000	10.0000	0.0000	20.0000	0.0000	30.0000	0.0000
40.0000	0.0000	100.0000					
10.0000	10.0000	109.0000	10.0000	46.0000	10.0000	64.0000	10.0000
82.0000	10.0000	-41.0000					
0.0000	0.0000	10.0000	100.0000	20.0000	0.0000	30.0000	0.0000
40.0000	0.0000	100.0000					
20.0000	20.0000	46.0000	20.0000	136.0000	20.0000	98.0000	20.0000
124.0000	20.0000	144.0000					
0.0000	0.0000	10.0000	0.0000	20.0000	100.0000	30.0000	0.0000
40.0000	0.0000	-100.0000					
30.0000	30.0000	64.0000	30.0000	98.0000	30.0000	181.0000	30.0000
166.0000	30.0000	169.0000					
0.0000	0.0000	10.0000	0.0000	20.0000	0.0000	30.0000	100.0000

40.0000	0.0000	100.0000						
40.0000	40.0000	82.0000	40.0000	124.0000	40.0000	166.0000	40.0000	
244.0000	40.0000	124.0000						
0.0000	0.0000	10.0000	0.0000	20.0000	0.0000	30.0000	0.0000	
40.0000	100.0000	100.0000						
100.0000	100.0000	-41.0000	100.0000	144.0000	-100.0000	169.0000	100.0000	
124.0000	100.0000	831.0001						

COEFFICIENT OF DETERMINATION = .9988

CORRELATION MATRIX OF X, Y...

1.0000	0.0000	.0958	0.0000	.1715	0.0000	.2230	0.0000	
.2561	0.0000	.3469						
0.0000	1.0000	.0958	0.0000	.1715	0.0000	.2230	0.0000	
.2561	0.0000	.3469						
.0958	.0958	1.0000	.0958	.3778	.0958	.4556	.0958	
.5028	.0958	-.1362						
0.0000	0.0000	.0958	1.0000	.1715	0.0000	.2230	0.0000	
.2561	0.0000	.3469						
.1715	.1715	.3778	.1715	1.0000	.1715	.6246	.1715	
.6807	.1715	.4283						
0.0000	0.0000	.0958	0.0000	.1715	1.0000	.2230	0.0000	
.2561	0.0000	-.3469						
.2230	.2230	.4556	.2230	.6246	.2230	1.0000	.2230	
.7899	.2230	.4358						
0.0000	0.0000	.0958	0.0000	.1715	0.0000	.2230	1.0000	
.2561	0.0000	.3469						
.2561	.2561	.5028	.2561	.6807	.2561	.7899	.2561	
1.0000	.2561	.2754						
0.0000	0.0000	.0958	0.0000	.1715	0.0000	.2230	0.0000	
.2561	1.0000	.3469						
.3469	.3469	-.1362	.3469	.4283	-.3469	.4358	.3469	
.2754	.3469	1.0000						

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:							
	-18.3276	-15.0855	-19.5842	-4.2261	-3.7627	-20.8667	-26.6131	4.9927
	-37.8853	3.3520	17.6379					
OBSN NO	2:							
	4.2114	6.8286	3.3025	16.4380	5.7217	8.5474	24.7190	3.1567
	23.5444	10.2661	37.5481					
OBSN NO	3:							
	-10.1245	-8.1323	-4.8733	10.2270	9.9560	-8.9136	9.6135	4.4458
	17.3491	20.3052	32.2707					

PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
INDICES OF INDEP VARS IN REGRESSION...								
1	2	3	4	5	6	7	8	
9	10							

DATA...  
 (FIRST 3 OBSNS)

-18.3276	-15.0855	-19.5842	-4.2261	-3.7627	-20.8667	-26.6131	4.9927
-37.8853	3.3520	17.6379					
4.2114	6.8286	3.3025	16.4380	5.7217	8.5474	24.7190	3.1567
23.5444	10.2661	37.5481					
-10.1245	-8.1323	-4.8733	10.2270	9.9560	-8.9136	9.6135	4.4458
17.3491	20.3052	32.2707					

MEANS FOR ALL VARIABLES...

.1809	-.0144	-.9842	-.0300	-.7362	-.2456	-.5050	-.6612
-.9207	.4231	1.8532					

STANDARD DEVIATIONS FOR ALL VARIABLE...

11.2156	10.1579	11.0153	10.7963	12.6680	9.7790	15.3408	8.9774
16.7076	10.4552	29.6484					

CORRELATION MATRIX...

1.0000								
.0715	1.0000							
.1161	.2030	1.0000						
-.0757	.2606	.3178	1.0000					
.1219	.3381	.5962	.3737	1.0000				
.0586	.0191	-.0289	-.0018	.0647	1.0000			
.2317	.3488	.6098	.4160	.6869	.1958	1.0000		
-.0626	-.0388	.0732	.0173	.2289	-.0569	.1695	1.0000	
.2519	.4437	.6629	.4428	.7419	.1700	.8684	.1628	1.0000
1.0000								
-.0664	-.0577	.1822	.1316	.2812	-.1056	.2220	.0684	
.2746	1.0000							
.2985	.4192	.1470	.4788	.5795	-.3112	.5123	.3875	
.4474	.4197	1.0000						

DETERMINANT = .2064E-01

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.1878								
.1190	1.4176							
.1437	.3235	2.0519						
.2815	-.0652	.0275	1.3480					
.0597	-.1781	-.4746	-.0985	2.5064				
.0861	.1732	.3451	.1239	.0269	1.1623			
-.1331	.1639	-.2461	-.2198	-.3721	-.2769	4.3004		
.1608	.2041	.2125	.1265	-.2961	.1486	-.1257	1.1297	
-.5956	-1.0093	-1.1262	-.4446	-1.0076	-.4453	-3.1919	-.2302	
6.3430								



.1970      .3345      .1700      .0356      -.2294      .2253      .0804      .0799  
 -.6155      1.2307

PRODUCT OF MATRIX AND INVERSE...  
 SHOULD BE IDENTITY MATRIX. IF NOT, ROUND OFF ERRORS  
 HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.0000								
-.0000	1.0000							
.0000	.000*	1.0000						
.0000	.000*	.000*	1.0000					
.0000	.000*	.000*	.000*	1.0000				
.0000	.000*	.000*	.000*	.000*	1.0000			
.0000	.000*	-.000*	.000*	.000*	.000*	1.0000		
-.0000	-.000*	.000*	-.000*	.000*	-.000*	.000*	1.0000	
.0000	.000*	.000*	.000*	.000*	.000*	.000*	.000*	1.0000
1.0000								
.0000	.000*	.000*	.000*	-.000*	.000*	.000*	.000*	
-.0000	1.0000							

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.9538E-04								
.1055E-04	.1388E-03							
.1175E-04	.2920E-04	.1708E-03						
.2348E-04	-.6005E-05	.2334E-05	.1168E-03					
.4242E-05	-.1398E-04	-.3436E-04	-.7272E-05	.1578E-03				
.7931E-05	.1761E-04	.3236E-04	.1185E-04	.2191E-05	.1228E-03			
-.7813E-05	.1062E-04	-.1471E-04	-.1341E-04	-.1934E-04	-.1865E-04	.1846E-		
.1613E-04	.2261E-04	.2170E-04	.1318E-04	-.2629E-04	.1710E-04	-.9220E-		
-.3210E-04	-.6007E-04	-.6181E-04	-.2490E-04	-.4809E-04	-.2753E-04	-.1258E-		
.2295E-03								
.1697E-04	.3182E-04	.1491E-04	.3183E-05	-.1749E-04	.2226E-04	.5065E-		
-.3559E-04	.1137E-03							

INDEX OF DEP VAR IN REGRESSION = 11

INDICES OF INDEP VARS IN REGRESSION...

1	2	3	4	5	6	7	8
9	10						

REGRESSION COEFFICIENTS...

1.0007	.9954	-.9832	.9969	.9780	-.9903	1.0018	1.0288
-1.0111	.9965						

T-VALUES...

93.2158	76.8713	-68.4353	83.9122	70.8374	-81.3103	67.0802	78.6530
-60.7164	85.0077						

STD DEVS OF REG COEFS...

.0107	.0129	.0144	.0119	.0138	.0122	.0149	.0131
.0167	.0117						

INTERCEPT = 1.0589

RESIDUAL STANDARD ERROR = 1.0992

SAMPLE MULTIPLE CORRELATION COEFFICIENT = .9994

SAMPLE COEFFICIENT OF DETERMINATION = .9988

SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = 86916.2109

DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 10.0000

MEAN SQUARE OF SSAR = 8691.6211

SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SDR = 107.5391

DEGREES OF FREEDOM ASSOCIATED WITH SDR = 89.0000

MEAN SQUARE OF SDR = 1.2083

T-VALUE = 7193.2402

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFIL1

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
1	2	3	4	5	6	7	8	
9	10							

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 4 ITERATIONS

DATA...

FIRST 3 OBSNS)

-18.3276	-15.0855	-19.5842	-4.2261	-3.7627	-20.8667	-26.6131	4.9927
-37.8853	3.3520	17.6379					
4.2114	6.8286	3.3025	16.4380	5.7217	8.5474	24.7190	3.1567
23.5444	10.2661	37.5481					
-10.1245	-8.1323	-4.8733	10.2270	9.9560	-8.9136	9.6135	4.4458
17.3491	20.3052	32.2707					

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.1809	-.0144	-.9842	-.0300	-.7362	-.2456	-.5050	-.6612
-.9207	.4231	1.8532					

MEDIANS...

.1880	.8130	-.8577	.5786	-.0596	-.2026	.3088	-1.2183
-.1978	.1489	-.7137					

INTERQUARTILE RANGES...

15.4688	14.0235	15.0469	16.7969	17.5469	10.8985	23.4218	10.9766
24.0156	14.8828	36.7695					

REGRESSION COEFFICIENTS, INTERCEPT...

.9768	.7948	-1.0648	1.0780	.9434	-1.0708	.6114	.9883
-.5644	.8971	1.1667					
.9779	.9134	-1.0462	1.0049	.9431	-1.0295	.8290	1.0049
-.7777	.9509	1.0958					
.9858	.9527	-1.0003	1.0032	.9334	-1.0069	.9064	1.0132
-.8737	.9705	1.0865					
.9919	.9696	-.9829	1.0030	.9399	-.9985	.9424	1.0169
-.9231	.9779	1.0816					
.9951	.9782	-.9755	1.0022	.9450	-.9946	.9614	1.0188
-.9489	.9815	1.0786					

PRINT AFTER EVERY 4 ITERATIONS

.9967	.9827	-.9717	1.0016	.9476	-.9926	.9713	1.0198
-.9624	.9835	1.0770					
.9975	.9850	-.9698	1.0014	.9489	-.9915	.9765	1.0203
-.9693	.9845	1.0762					

RESIDUAL VARIANCE = 1.3789  
RESIDUAL STANDARD ERROR = 1.1743

EFFICIENT OF DETERMINATION = .9984

PROGRAM SIMULA  
GENERATES MULTIVARIABLE UNIVARIATE SAMPLE

FILE NAME: DFIL1

DIMENSION (M) = 10

COEFFICIENT MATRIX (C)...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000						
.1000	1.0000	.1000	.1000	.1000	.1000	.1000	.1000
.1000	.1000						
.2000	.2000	1.0000	.2000	.2000	.2000	.2000	.2000
.2000	.2000						
.3000	.3000	.3000	1.0000	.3000	.3000	.3000	.3000
.3000	.3000						
.4000	.4000	.4000	.4000	1.0000	.4000	.4000	.4000
.4000	.4000						
.5000	.5000	.5000	.5000	.5000	1.0000	.5000	.5000
.5000	.5000						
.6000	.6000	.6000	.6000	.6000	.6000	1.0000	.6000
.6000	.6000						
.7000	.7000	.7000	.7000	.7000	.7000	.7000	1.0000
.7000	.7000						
.8000	.8000	.8000	.8000	.8000	.8000	.8000	.8000
1.0000	.8000						
.9000	.9000	.9000	.9000	.9000	.9000	.9000	.9000
.9000	1.0000						

SIGMA (SCALING PARAMETER FOR INDEP VARIABLES) = 10.0000

CO (INTERCEPT) = 1.0000

PARAMETER VECTOR (B)...

1.0000	1.0000	-1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000
-1.0000	1.0000						

SIG (MODEL ERROR STD DEV) = 1.0000

NO OF OBSNS = 100

SKIP 0 RANDOM NOS PRIOR TO SIMULATION

VARIANCE MATRIX OF X, Y...

100.0000	10.0000	20.0000	30.0000	40.0000	50.0000	60.0000	70.0000
80.0000	90.0000	250.0000					
10.0000	109.0000	46.0000	64.0000	82.0000	100.0000	118.0000	136.0000
154.0000	172.0000	391.0000					
20.0000	46.0000	136.0000	98.0000	124.0000	150.0000	176.0000	202.0000
228.0000	254.0000	406.0000					
30.0000	64.0000	98.0000	181.0000	166.0000	200.0000	234.0000	268.0000
302.0000	336.0000	679.0002					
40.0000	82.0000	124.0000	166.0000	244.0000	250.0000	292.0000	334.0000
376.0000	417.9999	826.0000					
50.0000	100.0000	150.0000	200.0000	250.0000	325.0000	350.0000	400.0000
450.0001	500.0000	924.9999					
60.0000	118.0000	176.0000	234.0000	292.0000	350.0000	424.0000	466.0000
524.0000	582.0000	1126.0001					
70.0000	136.0000	202.0000	268.0000	334.0000	400.0000	466.0000	540.9999

597.9999	664.0001	1279.0002					
80.0000	154.0000	228.0000	302.0000	376.0000	450.0001	524.0000	597.9999
676.0000	745.9999	1426.0000					
90.0000	172.0000	254.0000	336.0000	417.9999	500.0000	582.0000	664.0001
745.9999	829.0000	1591.0002					
250.0000	391.0000	406.0000	679.0002	826.0000	924.9999	1126.0001	1279.0002
426.0000	1591.0002	3386.0010					

COEFFICIENT OF DETERMINATION = .9997

CORRELATION MATRIX OF X, Y...

1.0000	.0958	.1715	.2230	.2561	.2774	.2914	.3010
.3077	.3126	.4296					
.0958	1.0000	.3778	.4556	.5028	.5313	.5489	.5601
.5673	.5722	.6436					
.1715	.3778	1.0000	.6246	.6807	.7135	.7329	.7447
.7520	.7565	.5983					
.2230	.4556	.6246	1.0000	.7899	.8246	.8447	.8564
.8634	.8674	.8673					
.2561	.5028	.6807	.7899	1.0000	.8878	.9078	.9193
.9258	.9294	.9087					
.2774	.5313	.7135	.8246	.8878	1.0000	.9429	.9539
.9601	.9633	.8818					
.2914	.5489	.7329	.8447	.9078	.9429	1.0000	.9730
.9788	.9817	.9397					
.3010	.5601	.7447	.8564	.9193	.9539	.9730	1.0000
.9888	.9915	.9450					
.3077	.5673	.7520	.8634	.9258	.9601	.9788	.9888
1.0000	.9965	.9425					
.3126	.5722	.7565	.8674	.9294	.9633	.9817	.9915
.9965	1.0000	.9496					
.4296	.6436	.5983	.8673	.9087	.8818	.9397	.9450
.9425	.9496	1.0000					

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
(FIRST 3 OBSNS PRINTED)

OBSN NO	1:							
	-18.3276	-20.8147	-25.4502	-24.6717	-20.9165	-46.6223	-46.2266	-49.1668
	-60.8804	-64.8050	-113.0040					
OBSN NO	2:							
	4.2114	11.6345	9.0342	27.9729	18.0132	31.7175	37.6484	39.3684
	44.4399	50.4255	105.6835					
OBSN NO	3:							
	-10.1245	-4.9788	-1.7315	14.1799	13.3178	7.2448	15.5234	17.7160
	21.3852	23.0935	41.3710					

PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
INDICES OF INDEP VARS IN REGRESSION...								
1	2	3	4	5	6	7	8	
9	10							

DATA...  
 (FIRST 3 OBSNS)

-18.3276	-20.8147	-25.4502	-24.6717	-20.9165	-46.6223	-46.2266	-49.1668
-60.8804	-64.8050	-113.0040					
4.2114	11.6345	9.0342	27.9729	18.0132	31.7175	37.6484	39.3684
44.4399	50.4255	105.6835					
-10.1245	-4.9788	-1.7315	14.1799	13.3178	7.2448	15.5234	17.7160
21.3852	23.0935	41.3710					

MEANS FOR ALL VARIABLES...

.1809	-.2009	-1.0837	-.5849	-1.0220	-1.0626	-1.0941	-1.5141
-1.5599	-1.6493	-1.1327					

STANDARD DEVIATIONS FOR ALL VARIABLES...

11.2156	11.2752	12.9436	16.0995	17.7264	19.7445	23.7825	26.1492
29.5282	32.9943	67.4830					

CORRELATION MATRIX...

1.0000								
.1514	1.0000							
.1783	.5113	1.0000						
.1608	.6446	.7567	1.0000					
.2139	.6434	.8100	.8599	1.0000				
.2814	.6573	.7846	.8749	.9099	1.0000			
.2753	.6649	.8248	.8957	.9292	.9590	1.0000		
.2757	.6777	.8313	.9052	.9497	.9627	.9813	1.0000	
.2854	.6933	.8404	.9095	.9467	.9672	.9866	.9921	
1.0000								
.2853	.6878	.8412	.9138	.9523	.9684	.9866	.9945	
.9977	1.0000							
.3872	.7512	.7412	.9097	.9324	.9197	.9560	.9659	
.9638	.9680	1.0000						

DETERMINANT = .1048E-08

INVERSE OF THE PART OF THE CORR MATRIX USED IN REGRESSION...

1.2492								
.1726	2.0793							
.3164	.5738	3.7911						
.7514	-.0601	.3480	6.4310					
.7383	.2471	-.1940	.8414	11.6186				
.1618	.4591	1.6436	.8575	2.0441	17.2304			
.6074	1.2267	.5277	1.1366	5.0786	-.8493	42.3096		
-2.0589	-3.7660	-4.9050	-5.9452	-13.3671	-15.6544	-40.4673	-21.7312	
-2.0589	-3.7660	-4.9050	-5.9452	-13.3671	-15.6544	-40.4673	28.4020	
90.7163								

2.0589 3.7660 4.9050 5.9452 13.3671 15.6544 40.4673 -38.8296  
 -90.7163 55.6381

PRODUCT OF MATRIX AND INVERSE...  
 SHOULD BE IDENTITY MATRIX. IF NOT, ROUND OFF ERRORS  
 HAVE RUINED SOLUTION... DISREGARD REST OF ANALYSIS.

1.2556  
 .5865 1.0617  
 .7119 .0645 1.9018  
 .7687 .0606 .9678 4.0301  
 .7869 .0347 .9728 3.1224 7.2449  
 .8072 .0493 1.0068 3.1927 6.3799 5.4214  
 .8210 .0476 1.0224 3.2494 6.4941 4.4960 8.3334  
 .8035 -.0005 .9697 3.2154 6.4452 4.3768 6.9886 -85.4023  
 .8328 .0584 1.0445 3.2812 6.5137 4.5306 7.3777 -85.3785  
 49.9736  
 .8442 .0575 1.0576 3.3284 6.6131 4.5963 7.4849 -85.6612  
 49.7392 -63.9288

INVERSE OF THE PART OF THE CROSS PROD MATRIX USED IN REGRESSION...

.1003E-03  
 .1379E-04 .1652E-03  
 .2202E-04 .3971E-04 .2286E-03  
 .4204E-04 -.3344E-05 .1687E-04 .2506E-03  
 .3751E-04 .1249E-04 -.8540E-05 .2978E-04 .3735E-03  
 .7379E-05 .2083E-04 .6496E-04 .2725E-04 .5899E-04 .4464E-03  
 .2300E-04 .4621E-04 .1732E-04 .2999E-04 .1217E-03 -.1827E-04 .7556E-  
 -.7091E-04 -.1290E-03 -.1464E-03 -.1426E-03 -.2913E-03 -.3063E-03 -.6573E-  
 -.6280E-04 -.1143E-03 -.1296E-03 -.1263E-03 -.2580E-03 -.2712E-03 -.5821E-  
 .1051E-02  
 .5620E-04 .1023E-03 .1160E-03 .1131E-03 .2309E-03 .2427E-03 .5209E-  
 -.9405E-03 .5162E-03

INDEX OF DEP VAR IN REGRESSION = 11

INDICES OF INDEP VARS IN REGRESSION...  
 1 2 3 4 5 6 7 8  
 9 10

REGRESSION COEFFICIENTS...  
 .9687 .8740 -1.0836 .9737 .9768 -1.1008 .6742 -169.3762  
 -39.3268 14.5951

T-VALUES...  
 .1583 .1113 -.1173 .1007 .0827 -.0853 .0401 -15.4745  
 -1.9858 1.0515

STD DEVS OF REG COEFS...  
 6.1186 7.8522 9.2360 9.6712 11.8062 12.9079 16.7925 10.9455  
 19.8043 13.8804

INTERCEPT = -294.8914  
 RESIDUAL STANDARD ERROR = 610.9026  
 SAMPLE MULTIPLE CORRELATION COEFFICIENT = 8.5249  
 SAMPLE COEFFICIENT OF DETERMINATION = -72.6732  
 SUM OF SQUARES ATTRIBUTABLE TO REGRESSION, SSAR = -32764136.0000  
 DEGREES OF FREEDOM ASSOCIATED WITH SSAR = 10.0000  
 MEAN SQUARE OF SSAR = -3276413.5000  
 SUM OF SQUARES OF DEVIATIONS FROM REGRESSION, SSDR = 33214978.0000  
 DEGREES OF FREEDOM ASSOCIATED WITH SSDR = 89.0000  
 MEAN SQUARE OF SSDR = 373202.0000  
 T-VALUE = -8.7792

PROGRAM ITEREG  
ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
INDICES OF INDEP VARS IN REGRESSION...								
1	2	3	4	5	6	7	8	
9	10							

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 40 ITERATIONS

DATA...  
FIRST 3 OBSNS)

-18.3276	-20.8147	-25.4502	-24.6717	-20.9165	-46.6223	-46.2266	-49.1668
-60.8804	-64.8050	-113.0040					
4.2114	11.6345	9.0342	27.9729	18.0132	31.7175	37.6484	39.3684
44.4399	50.4255	105.6835					
-10.1245	-4.9788	-1.7315	14.1799	13.3178	7.2448	15.5234	17.7160
21.3852	23.0935	41.3710					

MEANS FOR VARIABLES IN REGRESSION,  
DEPENDENT VARIABLE LAST...

.1809	-.2009	-1.0837	-.5849	-1.0220	-1.0626	-1.0941	-1.5141
-1.5599	-1.6493	-1.1327					

MEDIANS...

.1880	.0173	-1.1221	-1.7224	.4272	-1.7786	-1.4766	-3.6473
-2.5523	-2.8128	-.5040					

INTERQUARTILE RANGES...

15.4688	15.2031	18.2188	24.4649	27.7266	30.3711	36.3750	39.9883
49.3359	51.7383	93.7500					

REGRESSION COEFFICIENTS, INTERCEPT...

1.0906	1.0271	-.9463	.8936	1.8787	1.0710	.2496	.0809
.1764	-.4496	1.3606					
1.0074	1.0439	-.8263	1.0221	1.0995	.2052	1.1456	.1096
.3533	-.4498	1.1670					
1.0093	1.0351	-.8215	1.1452	.9199	-.5670	1.5165	.3191
.2901	-.3364	1.0498					
1.0185	1.0238	-.8448	1.1827	.9041	-.8994	1.5305	.5827
.1307	-.2183	1.0337					
1.0240	1.0140	-.8707	1.1778	.9241	-.9924	1.4104	.8178
-.0271	-.1211	1.0600					

PRINT AFTER EVERY 40 ITERATIONS

1.0265	1.0061	-.8934	1.1602	.9481	-.9931	1.2661	.9991
-.1501	-.0466	1.0947					
1.0275	.9998	-.9118	1.1422	.9695	-.9673	1.1377	1.1284
-.2356	.0087	1.1252					
1.0278	.9949	-.9263	1.1272	.9871	-.9388	1.0349	1.2158
-.2907	.0490	1.1484					
1.0278	.9912	-.9374	1.1160	1.0012	-.9145	.9568	1.2720



-.3238	.0781	1.1650					
1.0277	.9883	-.9458	1.1078	1.0121	-.8958	.8995	1.3060
-.3421	.0988	1.1762					

RESIDUAL VARIANCE = 4.1393  
RESIDUAL STANDARD ERROR = 2.0345  
COEFFICIENT OF DETERMINATION = .9991



0.0000	100.0000	200.0000						
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
100.0000	100.0000	0.0000						
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	100.0000	
100.0000	200.0000	200.0000						
100.0000	100.0000	-100.0000	100.0000	100.0000	-100.0000	100.0000	200.0000	
0.0000	200.0000	1101.0000						

COEFFICIENT OF DETERMINATION = .9991

CORRELATION MATRIX OF X, Y...

1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	.3014						
0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	.3014						
0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	-.3014						
0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	.3014						
0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
0.0000	0.0000	-.3014						
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000
0.0000	0.0000	.3014						
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
0.0000	.7071	.6027						
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1.0000	.7071	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	.7071
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	.7071
.7071	1.0000	.4262						
.3014	.3014	-.3014	.3014	.3014	-.3014	.3014	.3014	.6027
0.0000	.4262	1.0000						

OBSERVATIONS...

DEPENDENT VARIABLE LAST  
 (FIRST 3 OBSNS PRINTED)

OBSN NO	1:							
	-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-6.9995	4.9927
	-14.8901	-9.8975	12.2946					
OBSN NO	2:							
	4.2114	6.8286	-2.4292	16.4380	-6.5698	8.5474	11.7895	3.1567
	2.6489	5.8057	34.4939					
OBSN NO	3:							
	-10.1245	-8.1323	-8.0151	10.2270	6.5942	-8.9136	3.7036	4.4458
	13.3130	17.7588	27.6306					

PROGRAM MREG  
 MULTIPLE LINEAR REGRESSION ANALYSIS

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
INDICES OF INDEP VARS IN REGRESSION...								
1	2	3	4	5	6	7	8	
9	10							

DATA...

(FIRST 3 OBSNS)

-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-6.9995	4.9927
-14.8901	-9.8975	12.2946					
4.2114	6.8286	-2.4292	16.4380	-6.5698	8.5474	11.7895	3.1567
2.6489	5.8057	34.4939					
-10.1245	-8.1323	-8.0151	10.2270	6.5942	-8.9136	3.7036	4.4458
13.3130	17.7588	27.6306					

MEANS FOR ALL VARIABLES...

.1809	-.0144	-.8847	-.0300	-.4503	-.2456	.0841	-.6612
-.2816	-.9428	.6235					

STANDARD DEVIATIONS FOR ALL VARIABLES...

11.2156	10.1579	9.8358	10.7963	9.8302	9.7790	10.1817	8.9774
8.7924	11.7031	32.6040					

CORRELATION MATRIX...

1.0000								
.0715	1.0000							
.0254	.0648	1.0000						
-.0757	.2606	.1868	1.0000					
-.0717	.1217	.2235	.1323	1.0000				
.0586	.0191	-.1219	-.0018	-.0893	1.0000			
.0552	.0508	.1240	.1180	-.0356	.1012	1.0000		
-.0626	-.0388	-.0055	.0173	.1513	-.0569	.0051	1.0000	
-.0012	.1444	.0992	.0022	-.1068	-.0386	.1414	-.1326	1.0000
1.0000								
-.0489	.0787	.0704	.0150	.0358	-.0726	.1102	.6675	
.6495	1.0000							
.2755	.4226	-.0722	.4118	.3835	-.2632	.3084	.5941	
-.0381	.4271	1.0000						

DETERMINANT = -.2833E-06

ZERO DETERMINANT, REGRESSION CANNOT BE PERFORMED

PROGRAM ITEREG  
 ESTIMATES REGRESSION COEFFICIENTS BY AN ITERATIVE METHOD

DATA FILE NAME: DFILL

NO OF OBS, NO OF VARS, INDEX OF DEP VAR  
 IN REGRESSION, NO OF INDEP VARS IN REGRESSION...

100	11	11	10					
INDICES OF INDEP VARS IN REGRESSION...								
1	2	3	4	5	6	7	8	
9	10							

REGRESSION METHOD USED AT EACH ITERATION: WALD'S METHOD

PRINT AFTER EVERY 2 ITERATIONS

DATA...

FIRST	3 OBSNS)							
-18.3276	-15.0855	-13.7183	-4.2261	13.3911	-20.8667	-6.9995	4.9927	
-14.8901	-9.8975	12.2946						
4.2114	6.8286	-2.4292	16.4380	-6.5698	8.5474	11.7895	3.1567	
2.6489	5.8057	34.4939						
-10.1245	-8.1323	-8.0151	10.2270	6.5942	-8.9136	3.7036	4.4458	
13.3130	17.7588	27.6306						

MEANS FOR VARIABLES IN REGRESSION,  
 DEPENDENT VARIABLE LAST...

.1809	-.0144	-.8847	-.0300	-.4503	-.2456	.0841	-.6612	
-.2816	-.9428	.6235						

MEDIANS...

.1880	.8130	-.7495	.5786	-.8276	-.2026	-1.4136	-1.2183	
-.0464	-1.6943	1.4197						

INTERQUARTILE RANGES...

15.4688	14.0235	12.7734	16.7969	13.5547	10.8985	15.6250	10.9766	
12.5390	15.0781	36.7578						

REGRESSION COEFFICIENTS, INTERCEPT...

.9995	.9142	-1.0840	.9995	.9548	-1.0296	1.0589	2.0564	
.0479	-.0360	.9540						
.9978	.9977	-.9735	.9893	.9673	-.9902	1.0010	2.0465	
.0039	-.0265	1.0633						
.9975	.9951	-.9748	.9900	.9668	-.9901	1.0019	2.0334	
-.0085	-.0135	1.0620						
.9975	.9951	-.9749	.9900	.9668	-.9901	1.0019	2.0205	
-.0214	-.0006	1.0619						
.9975	.9951	-.9749	.9900	.9668	-.9901	1.0019	2.0076	
-.0343	.0124	1.0619						

RESIDUAL VARIANCE = 1.2721  
 RESIDUAL STANDARD ERROR = 1.1279  
 COEFFICIENT OF DETERMINATION = .9988

END

DTIC

6-86