

**A MODIFICATION OF THE GAUSS-JORDAN PROCEDURE AS AN ADEQUATE
ALTERNATIVE TO ITERATIVE PROCEDURES IN MULTIPLE REGRESSION**

by 632

MELVIN HILDEBRAND

**B. S., Concordia Teachers College, 1967
Seward, Nebraska**

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

**KANSAS STATE UNIVERSITY
Manhattan, Kansas**

1970

Approved by:

Herbert E. Kaiser
Major Professor

LD
2668
R4
1970
H55
C.2

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Statement of the Problem	1
Purpose of the Study	3
Questions for the Study	3
Definition of Terms	4
II. RELATED LITERATURE	6
Introduction	6
Background in Matrix Algebra as it Applies to	
Multiple Regression	6
A Brief Study of the Modified Gauss-Jordan	8
Why the Iterative Procedures have been Used in	
Multiple Regression	10
Summary	13
III. METHODS AND PROCEDURES	14
Description of Multiple Regression Programs	
Used in the Study	14
Characteristics of the Data Used in the Programs	14
Organization of the Study	16
Summary	17

CHAPTER	PAGE
IV. ANALYSIS OF THE DATA	18
Introduction	18
Relating the Results and Efficiency of the Iterative and Modified Gauss-Jordan Procedures	18
Summary	22
V. SUMMARY AND CONCLUSION	23
Introduction	23
Summary	26
BIBLIOGRAPHY	28

LIST OF TABLES

TABLE	PAGE
I. Comparison of Least Squares Solution by Iterative and Modified Gauss-Jordan Programs	19
II. Comparative Analysis of Compile and Execution Time (in seconds) Used by Iterative and Modified Gauss-Jordan Programs.	20

CHAPTER I

INTRODUCTION

STATEMENT OF THE PROBLEM

This study is an investigation of the Gauss-Jordan procedure modified for use with correlation matrices having linear dependencies. The simplification of the matrix inversion portion in studies of multiple regression would mean educational research could be more easily conducted. More specifically, in the public school setting the guidance counselor would feel more comfortable doing applied research in establishing relationships and testing theories in the field settings.

The traditional approaches for deriving the multiple regression coefficients using the Gauss-Jordan procedure do not usually function in the presence of linear dependencies. Veldman¹ related the Gauss-Jordan procedure programmed in Fortran for the computer.

¹Donald J. Veldman, Fortran Programing for the Behavioral Sciences (New York: Holt, Rinehart and Winston, 1967), pp. 156-59.

Typically, the multiple regression program determines the determinant of the correlation matrix in order to check for linear dependency. If the determinant is zero there is linear dependency, execution ceases, and an error statement is printed by the computer. If the determinant is not zero, execution is completed, and the solution is determined, using the Gauss-Jordan procedure.

Another approach for deriving the multiple regression coefficients is the iterative procedure which yields a solution meeting the least squares criterion despite the presence of linear dependencies. Veldman² also reported the computer program using iterative techniques to derive the regression coefficients from the correlation matrix.

The Gauss-Jordan procedure provides a direct solution characterized by a specific process executed once, but will not function in the presence of linear dependencies. In contrast, the iterative technique is a cycle of approximations carried out repeatedly until a prescribed accuracy is achieved, but this method is so complicated it is difficult to share with the typical user.

²Donald J. Veldman, Fortran Programming for the Behavioral Sciences (New York: Holt, Rinehart and Winston, 1967), pp. 294-307.

PURPOSE OF THE STUDY

The purpose of the study was to investigate the validity and utility in the modification of the Gauss-Jordan procedure. The matrix inversion technique was suggested by Roscoe and Kittleson.³

QUESTIONS FOR THE STUDY

After reviewing literature related to the problem and critically analyzing the modified Gauss-Jordan, questions were developed for this study as follows:

Question 1 -- Does the modified Gauss-Jordan yield results comparable to the iterative procedure cited above in the presence of various kinds of linear dependencies?

Question 2 -- How does the efficiency (in terms of computer time used) of the modified Gauss-Jordan compare to that of the iterative procedure which is widely used by educational researchers?

Question 3 -- Can the modified Gauss-Jordan be conveniently altered to calculate the rank of the matrix?

³John T. Roscoe and Howard M. Kittleson, "A Modified Gauss-Jordan Procedure as an Alternative to Iterative Procedures in Multiple Regression" (Paper prepared for presentation at the annual convention of the American Educational Research Association, February, 1971, New York City, New York).

DEFINITION OF TERMS

Multiple Regression provides a prediction equation for a single criterion measure and two or more predictor measures.

Linear Dependencies occur when one vector in a set of vectors can be expressed as a linear combination of the other vectors.

Correlation Matrix is a matrix that has correlation coefficients as elements.

Rank of a Matrix as used in this paper is the number of ones in the canonical form under both row and column operations.

Iterative Procedure is a method for finding the inverse of a matrix by employing a cycle of computations.

Least Square Criterion is a procedure that minimizes the sum of the squared values of the element in a vector in order to make the errors as small as possible.

Modified Gauss-Jordan is a matrix inversion that was produced from the traditional Gauss-Jordan matrix inversion technique. The inversion executes despite linear dependencies, and is written with eleven instead of about sixty Fortran statements.

Augmented Matrix is a matrix that is formed by placing a constant column vector to the right of a matrix.

Matrix Inversion is the method of multiplying a given matrix by a second matrix in order to produce the identity matrix.

Regression Weights are constants that are multiplied by vectors to produce a linear combination of a set of vectors.

Rounding Error is the error introduced when the last desired digit of a number is approximated based on the character to the right of the desired ending digit.

CHAPTER II

RELATED LITERATURE

INTRODUCTION

The review of literature related to the questions of this study includes (1) background in matrix algebra as it applies to the Gauss-Jordan Procedure in multiple regression, (2) a study of the modified Gauss-Jordan computer program written in Fortran for multiple regression, and (3) the argument that is raised as to why the iterative procedures have been used in multiple regression programs.

BACKGROUND IN MATRIX ALGEBRA AS IT APPLIES TO MULTIPLE REGRESSION

Fuller⁴ described the matrix algebra needed to work with the multiple regression problem. A system of linear equations in standard form having variables and coefficients on the

⁴Leonard E. Fuller, Basic Matrix Theory (Englewood Cliffs: Prentice-Hall, Inc., 1962), pp. 39-69.

left side of the equation and constants on the right side is given as follows:

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

The problem is to determine the values of the variables.

If a series of algebraic operations is performed on the equations such that the equalities are retained, and the solution is represented by b_1 , b_2 , and b_3 the result may be shown as follows:

$$\begin{bmatrix} 1x_1 + 0x_2 + 0x_3 \\ 0x_1 + 1x_2 + 0x_3 \\ 0x_1 + 0x_2 + 1x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

The iterative procedure differs from the Gauss-Jordan only in the method of solving for the solution to the augmented matrix.

It can be shown how the Gauss-Jordan works in the multiple regression program if a correlation matrix is substituted for the original system of linear equations as follows:

$$\begin{bmatrix} r_{11}x_1 + r_{12}x_2 + r_{13}x_3 \\ r_{21}x_1 + r_{22}x_2 + r_{23}x_3 \\ r_{31}x_1 + r_{32}x_2 + r_{33}x_3 \end{bmatrix} = \begin{bmatrix} r_{14} \\ r_{24} \\ r_{34} \end{bmatrix}$$

The solution may be derived as shown above. Letting r_{nm} represent the elements of the correlation matrix for the independent variables, r_{14} , r_{24} , r_{34} represent the elements of correlation matrix for the dependent variables with each of the independent variables, and b_1 , b_2 , and b_3 represent the elements of the desired regression coefficients. These coefficients could be substituted back into the original system of linear equations and written as:

$$\begin{bmatrix} r_{11}b_1 + r_{12}b_2 + r_{13}b_3 \\ r_{21}b_1 + r_{22}b_2 + r_{23}b_3 \\ r_{31}b_1 + r_{32}b_2 + r_{33}b_3 \end{bmatrix} = \begin{bmatrix} r_{14} \\ r_{24} \\ r_{34} \end{bmatrix}$$

A BRIEF STUDY OF THE MODIFIED GAUSS-JORDAN

The modified Gauss-Jordan requires eleven Fortran statements:

```

      DO 400 I=1, NR
      BB = B(I,I)
      IF(ABS(BB).LT.0.0001) BB = 1.0
      DO 310 J=1, NC
310   B(I,J) = B(I,J)/BB
      DO 400 L=1, NR
      IF(I.EQ.L) GO TO 400
      AA = -B(L,I)
      DO 340 J=1, NC
340   B(L,J) = B(L,J) + B(I,J) * AA
400  CONTINUE

```

Where B is the augmented matrix described earlier. NR is the number of rows, and NC the number of columns (NC = NR + 1).

When the standard system of linear equations is used as follows:

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Row one, two and three are divided by elements a_{11} , a_{22} , and a_{33} respectively in order to put one in the diagonal. A problem is encountered in the above program if the value of $B(I,I)$ is zero. Division by zero is not defined in the Fortran compiler, and compiling ceases. This division by zero may be avoided by the third Fortran statement:

```

      IF(ABS(BB).LT.0.0001) BB = 1.0

```

Row one is multiplied by the negative of elements a_{21} and a_{31} , and added to rows two and three respectively. This proceeds to subtract a_{21} from a_{21} and a_{31} from a_{31} , and leaves zero in these positions.

Likewise, row two is multiplied by the negative of elements a_{12} and a_{32} , and added to rows one and three respectively. This subtracts elements a_{12} from a_{12} and a_{32} from a_{32} , and puts zero in these positions.

Lastly, row three is multiplied by the negative of elements a_{13} and a_{23} , and added to rows one and two respectively. This successfully zeroes the elements a_{13} and a_{23} .

With these operations successfully carried out the result is:

$$\begin{bmatrix} 1x_1 + 0x_2 + 0x_3 \\ 0x_1 + 1x_2 + 0x_3 \\ 0x_1 + 0x_2 + 1x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

The solution being represented by b_1 , b_2 , and b_3 .

WHY THE ITERATIVE PROCEDURES HAVE BEEN USED IN MULTIPLE REGRESSION

The researchers who use and report multiple regression programs are brief in their explanation as to why the iterative

procedures have been used instead of other techniques of matrix inversion to calculate the weights for the predictor equation.

Kelly⁵ reported that if more accuracy in determining the weights is desired the number of iterations should be increased. He made it known that this involves more computer time.

Ward⁶ stated that the iterative procedure came into use through the need for a computational procedure that guaranteed a solution, and tended to select the most efficient subset of predictors in the system.

In a journal article Jennings⁷ indicated that the iterative procedures were used to circumvent the difficulties encountered when the predictor matrix was singular; and to speed processing time.

⁵Francis J. Kelly and others, Research Design in the Behavioral Sciences: Multiple Regression Approach (London: Feffer & Simons, Inc., 1969), p. 254.

⁶Joe H. Ward, "A Computational Procedure for Solving Regression Problems," Computer Applications in the Behavioral Sciences, ed. H. Borko (Englewood Cliffs: Prentice-Hall, Inc., 1962), pp. 204-37.

⁷Earl Jennings and Joe H. Ward, "CDC 1604 Program for Intercorrelation and Multiple Regression Analysis," Behavioral Science, 44:81, 1963.

Veldman⁸ reported that the iterative procedure avoided the necessity of obtaining the inverse of the correlation matrix. The iterative procedure calculates a regression equation by adding variables to the predictor set or by adjusting the weights of variables already in the set in such a way as to maximize the increase in the square of the multiple correlation at each step.

Fuller⁹ identified the iterative procedures as an inversion method that can be used successfully with matrices whose powers approach the zero matrix. Instead of being a direct method, it involves a cycle of computations which is repeated until a prescribed accuracy is obtained. In the iterative technique the inverse of a matrix is approximated, and then multiplied times the original matrix. This product is subtracted from the identity matrix, and gives the error matrix. To test whether the approximated inverse has the desired accuracy the sum of the absolute values of the elements in each row or each column of the error matrix is determined. If the sums are

⁸Donald J. Veldman, Fortran Programing for the Behavioral Sciences (New York: Holt, Rinehart and Winston, 1967), pp. 294-307.

⁹Leonard E. Fuller, Basic Matrix Theory, (Englewood Cliffs: Prentice-Hall, Inc., 1962), pp. 177-87.

less than one, the inverse matrix has the needed accuracy. If the sums are not all less than one, the inverse matrix is approximated repeatedly until the desired property is obtained.

SUMMARY

Literature dealing with matrix algebra, the modified Gauss-Jordan, and reasons for using the iterative procedures have been surveyed in this chapter.

The chapter reports that the modified Gauss-Jordan does accomplish matrix inversion, and determines a solution to an augmented matrix. The reason for using the iterative procedure is that it guarantees a solution to an augmented matrix, and the calculation can be continued until the prescribed accuracy is obtained.

CHAPTER III

METHODS AND PROCEDURES

DESCRIPTION OF MULTIPLE REGRESSION PROGRAMS USED IN THE STUDY

The multiple regression programs used in the study for comparing results and efficiency of the modified Gauss-Jordan and the iterative procedures were the same except for the matrix inversion portion. Investigation of the two matrix inversion methods was accomplished by running the modified Gauss-Jordan and the iterative procedures simultaneously.

The multiple regression programs used were programmed for single precision accuracy so that the rounding errors were more likely to be evident. It is strongly urged that double precision accuracy be used by those who run these programs to insure more accuracy.

CHARACTERISTICS OF THE DATA USED IN THE PROGRAMS

Four sets of data were used to check the results and efficiency of the modified Gauss-Jordan and the iterative procedures. In the following description of data N is the

number of data cards, and NV is the number of variables on each card plus the number of variables to be generated.

The first set of data was the miscellaneous dependencies which was a small data deck ($N = 9$, $NV = 10$) developed by Roscoe and Kittleson¹⁰ for demonstration of linear dependencies. It included a design matrix for simple analysis of variance, a variable which is the sum of two other variables, a variable which is a constant, and a variable which is a blank field in the data deck.

The second comparison was made with the two dimensional analysis of variance which was a set of data ($N = 24$, $NV = 1$) for two dimensional analysis of variance with unequal but proportional cell frequencies. The problem was constructed by Roscoe¹¹ for classroom use in demonstrating variable generation techniques and equivalence of the analysis of variance the analysis of regression.

¹⁰ John T. Roscoe and Howard M. Kittleson, "A Modified Gauss-Jordan Procedure as an Alternative to Iterative Procedures in Multiple Regression" (Paper prepared for presentation at the annual convention of the American Educational Research Association, February, 1971, New York City, New York.

¹¹ Data constructed by John T. Roscoe in classroom notes, Kansas State University, 1970.

The third set of data was taken from the analysis of covariance problem in the text by Roscoe.¹²

The last investigation was conducted with a continuous data set which contained a set of data cards (N = 50, NV = 26) derived from public school research, modified by Roscoe¹³ to introduce a maximum amount of rounding error in digital statistical analysis.

ORGANIZATION OF THE STUDY

After multiple regression programs containing the iterative and modified Gauss-Jordan inversion procedures were assembled, the following information was collected from each program and analyzed.

1. The multiple regression program using a series of subroutines were modeled after Veldman¹⁴ and adapted to the IBM 360. The program was written in the Fortran computer language.

¹² John T. Roscoe, Fundamental Research Statistics: For the Behavioral Sciences (New York: Holt, Rinehart and Winston, Inc., 1969), pp. 254-84.

¹³ Data modified by John T. Roscoe in notes, Kansas State University, 1970.

¹⁴ Donald J. Veldman, Fortran Programing for the Behavioral Sciences (New York: Holt, Rinehart and Winston, 1967), pp. 294-307.

2. The modified Gauss-Jordan procedure was taken from a paper prepared by Roscoe and Kittleson.¹⁵

3. Four sets of data were used in the study. Each set of data was run in both programs at the same time. The compile time for the two programs was compared. For each set of data the R-squares and execution time were compared.

4. Tables were drawn to compare the results and time of the iterative and modified Gauss-Jordan procedures in multiple regression programs.

5. An investigation was made to see whether the modified Gauss-Jordan procedure could be conveniently altered to calculate the rank of the matrix.

SUMMARY

This chapter was a review of the research methods employed in this study. Included was a brief description of the programs analyzed, a summary of the data used to compare the two inversion procedures, and a list of the steps used to carry out the study.

¹⁵ John T. Roscoe and Howard M. Kittleson, "A Modified Gauss-Jordan Procedure as an Alternative to Iterative Procedures in Multiple Regression" (Paper prepared for presentation at the annual convention of the American Education Research Association, February, 1971, New York City, New York).

CHAPTER IV

ANALYSIS OF THE DATA

INTRODUCTION

This chapter presents an analysis of the data concerning the comparison of the iterative and modified Gauss-Jordan procedures in multiple regression. The results and efficiency of each program were analyzed.

RELATING THE RESULTS AND EFFICIENCY OF THE ITERATIVE AND MODIFIED GAUSS-JORDAN PROCEDURES

R-Square Values. (Table I) The R-square values are shown in the first table. Decimal accuracy was carried to four significant figures to the right of the decimal point.

Compile and Execution Time. (Table II) The compile and execution times were recorded in seconds, and carried to two significant figures to the right of the decimal point. The compile times used in carrying out the execution of the four sets of data were averaged for both the iterative and modified Gauss-Jordan procedures. The execution times for the four

Table 1
Comparison of Least Squares Solution by Iterative and
Modified Gauss-Jordan Programs

Data	Model	Iterative	Modified Gauss-Jordan
Miscellaneous Linear Dependencies	1	0.6900	0.6900
	2	0.6900	0.6900
	3	0.6900	0.6900
	4	0.3333	0.3333
	5	0.3333	0.3333
Two Dimensional Analysis of Variance	1	0.7784	0.7784
	2	0.6486	0.6486
	3	0.2595	0.2595
	4	0.6486	0.6486
Analysis of Covariance	1	0.8497	0.8497
	2	0.4719	0.4719
Continuous Data	1	0.7183	0.7183
	2	0.7181	0.7183
	3	0.6903	0.6904
	4	0.6771	0.6772
	5	0.6765	0.6766
	6	0.6765	0.6766
	7	0.6765	0.6766
	8	0.6764	0.6766
	9	0.6757	0.6758
	10	0.6628	0.6628

Table 2

Comparative Analysis of Compile and Execution Time (in seconds)
Used by Iterative and Modified Gauss-Jordan Programs

Procedure	Iterative	Modified Gauss-Jordan
Average Compile Time (Watfor Compiler)	16.84	7.19
Miscellaneous Linear Dependencies Problem (Execution)	9.59	9.27
Two Dimensional Analysis of Variance Problem (Execution)	19.36	14.44
Analysis of Covariance Problem (Execution)	2.47	2.90
Continuous Data Problem (Execution)	68.37	103.48

sets of data were recorded separately for the two procedures. Each procedure was programmed to have a maximum time of five minutes.

Determining the Rank of the Matrix. The investigation was carried on using for the definition of rank the number of ones in the canonical form under elementary row operations. A problem arises when the rank is found this way in the

modified Gauss-Jordan procedure, because of the following statement in the program:

```
IF(ABS(BB).LT. 0.0001) BB = 1.0
```

When the diagonal element $B(I,I)$ was less than 0.0001 it was left in that position, because the row in question was divided by one and not the diagonal element. In case that occurred the contribution to the rank by such an element was defined as zero. The number 0.0001 was an arbitrary cut off point, but such a rounding procedure was needed when the program is run on the computer.

The canonical rank of the matrix with the previous stipulation reported by Roscoe and Kittleson¹⁸ is shown in the following statements:

```
RANK(K) = 0.0
DO 410 J=1, NR
410 RANK(K) = RANK(K) + B(J,J)
```

These statements sum the diagonal elements after the inversion of the correlation matrix is completed.

¹⁸ John T. Roscoe and Howard M. Kittleson, "A Modified Gauss-Jordan Procedure as an Alternative to Iterative Procedures in Multiple Regression" (Paper prepared for presentation at the annual convention of the American Educational Research Association, February, 1971, New York City, New York).

SUMMARY

This chapter presented the findings of this study. The results and efficiency of the iterative and modified Gauss-Jordan procedures using four sets of data were reported. A method of calculating the rank was investigated which involved a definition of a zero rank.

CHAPTER V

SUMMARY AND CONCLUSION

INTRODUCTION

It was the purpose of this study to compare the results and efficiency of the modified Gauss-Jordan to iterative procedures, and to investigate whether the modified Gauss-Jordan could be altered to find the rank of the matrix.

In this chapter, the three questions investigated were reviewed, the evidence regarding each was summarized and conclusions were made.

Question 1 -- The two procedures have no significant difference with respect to the R-square values calculated using single precision on the programs. The only difference in the R-squares was obtained using the continuous data, and the greatest amount of difference was 0.0002. There was no difference in the R-squares carried out to four significant digits to the right of the decimal point in the other three sets of data used for this study. In order to determine which program was more accurate, the continuous data problem was run

in double precision. The results were the same as in single precision for both programs. The stop criteria of the iterative and modified Gauss-Jordan regression programs were changed from 0.0000001 and 0.0001 to 0.000000001 and 0.00001 respectively. When the continuous data problem was run in double precision and with the stop criterion changed, the R-square values of the iterative program were identical to the previously recorded modified Gauss-Jordan values. This would indicate that the Gauss-Jordan was the more accurate of the two original programs in this study.

Question 2 -- The two procedures have significantly different compile and execution times. The compile time for the iterative procedure ranged from 13.16 to 19.36 seconds averaging 16.84 seconds, compared to a range from 5.50 to 8.46 seconds averaging 7.19 seconds for the modified Gauss-Jordan.

The execution time of the miscellaneous linear dependencies, and the two dimensional analysis of variance problems differed 0.32 and 4.94 seconds respectively in favor of the modified Gauss-Jordan. The difference in the execution of the analysis of covariance, and the continuous data problems was 0.43 and 35.11 seconds respectively with the iterative being the quicker procedure.

Question 3 -- The modified Gauss-Jordan was altered in order to determine the rank of the matrix, but a problem arises when the rank is determined using the procedure. The difficulty lies in the statement:

IF(ABS(BB).LT.0.0001) BB = 1.0

According to Fuller¹⁹ the correct rank for all correlation matrices cannot be found by the above method. The exception being when the diagonal element is small. With the arbitrary definition of zero, the rank for the rows with diagonal elements between the absolute value of 1.0 and 0.0001 is one, and the rank for the rows with diagonal elements less than the absolute value of 0.0001 is zero. For all practical purposes the row ranks of less than 0.0001 when summed will be zero, when rounded to the nearest whole number. As the number in the diagonal approaches zero the computer is left to make the decision as to whether the rank is zero or one. When using the previous procedure it is suggested that the rows containing small diagonal elements be printed out so the user can analyze

¹⁹ Statement by Leonard E. Fuller, personal interview, September 23, 1970.

the results and accept or reject the decision made by the computer based on the accuracy desired.

SUMMARY

This study compared the results and efficiency of the iterative and modified Gauss-Jordan procedures, and investigated the possibility of altering the modified Gauss-Jordan to find the rank.

Of these factors there was no significant difference in the results of the R-squares using single precision. The continuous data problem was run in double precision, but the results were the same as in single precision. The stop criteria of the iterative and modified Gauss-Jordan regression programs were changed from 0.0000001 and 0.0001 to 0.000000001 and 0.00001 respectively. When the continuous data problem was run in double precision and with the stop criterion changed, the R-square values of the iterative program were identical to the previously record modified Gauss-Jordan values. This would indicate that the modified Gauss-Jordan is more accurate in single precision using the original stop criterion. There was significant difference in the efficiency with the modified Gauss-Jordan taking approximately half the time to compile while executing quicker on two of the four sets of

data in this study. Determining if the modified Gauss-Jordan can function more efficiently needs further investigation. The altering of the modified Gauss-Jordan in order to find the rank was investigated. It involved an arbitrary definition of the zero rank in order for it to run on the computer. This method of finding the rank will not work for all matrices, therefore some users may prefer to print out the rows containing small diagonal elements in order to analyze the results.

BIBLIOGRAPHY

Articles and Periodicals

Jennings, Earl. "Fixed Effects Analysis of Variance by Regression Analysis," Multivariate Behavioral Research, 2:95-108, January, 1967.

Jennings, Earl and Joe H. Ward. "CDC 1604 Program for Intercorrelation and Multiple Regression Analysis," Behavioral Science, 44:81, January, 1963.

Books

Borko, H. (ed.). Computer Application in the Behavioral Sciences. Englewood Cliffs: Prentice-Hall, 1962. 633 pp.

Cooley, William W. and Paul R. Lohnes. Multivariate Procedures for Behavioral Science, New York: John Wiley & Sons, 1962. 211 pp.

Fuller, Leonard E. Basic Matrix Theory, Englewood Cliffs: Prentice-Hall, 1962. 245 pp.

Kelly, Francis J. and others. Research Design in the Behavioral Sciences: Multiple Regression Approach, London: Feffer & Simons, 1969. 353 pp.

Roscoe, John T. Fundamental Research Statistics: For the Behavioral Sciences, New York: Holt, Rinehart and Winston, 1969. 336 pp.

Veldman, Donald J. Fortran Programing for the Behavioral Sciences, New York: Holt, Rinehart and Winston, 1967. 406 pp.

Unpublished Materials

- Bottenberg, Robert A. Applied Multiple Linear Regression.
U. S. Department of Commerce Office of Technical Service
Publication No. AD413128. Lackland Air Force Base,
Texas: Air Force Systems Command, March. 1963. 138 pp.
- Fuller, Leonard E. Personal interview. September 23, 1970.
- Roscoe, John T. Data constructed by John T. Roscoe in class-
room notes, Kansas State University, 1970.
- Roscoe, John T. Data modified by John T. Roscoe in notes,
Kansas State University, 1970.
- Roscoe, John T. and Howard M. Kittleson. "A Modified Gauss-
Jordan Procedure as an Alternative to Iterative Pro-
cedures in Multiple Regression." Paper prepared for
presentation at the annual convention of the American
Educational Research Association, February, 1971, New
York City, New York.

A MODIFICATION OF THE GAUSS-JORDAN PROCEDURE AS AN ADEQUATE
ALTERNATIVE TO ITERATIVE PROCEDURES IN MULTIPLE REGRESSION

by

MELVIN HILDEBRAND

B. S., Concordia Teachers College, 1967
Seward, Nebraska

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1970

ABSTRACT

STATEMENT OF THE PROBLEM

This study is an investigation and analysis of the modified Gauss-Jordan procedure as an adequate alternative to iterative procedures in multiple regression.

PURPOSE OF THE STUDY

The purpose of the study was to investigate the validity and utility in the modification of the Gauss-Jordan procedure. The matrix inversion technique was suggested by Roscoe and Kittleson.¹

PROCEDURES

Both multiple regression programs used in the study were written in Fortran, and the investigation was conducted using the Watfor compiler for the IBM model 360. The programs differed in the matrix inversion portion, one was written using the iterative procedure, and the other using the Roscoe modification of the Gauss-Jordan.

¹John T. Roscoe and Howard M. Kittleson, "A Modified Gauss-Jordan Procedure as an Alternative to Iterative Procedures in Multiple Regression" (Paper prepared for presentation at the annual convention for the American Educational Research Association, February, 1971, New York City, New York).

Four sets of data were run in the two programs comparing the R-squares, and the compile and execution time. An investigation was carried on to determine whether the modified Gauss-Jordan procedure could be altered to find the rank of the matrix.

FINDINGS

Based upon the analysis of the data, the following results are reported:

1. The two procedures have no significant difference with respect to the R-square values calculated. The greatest amount of difference in the R-squares was 0.0002, which was obtained using the continuous data. When the programs were run in double precision and the stop criterion was changed to insure greater accuracy, the difference disappeared. The R-square values of the iterative program were changed and matched the R-square values of the modified Gauss-Jordan in single precision. This would indicate that the modified Gauss-Jordan was the more accurate of the two original programs in this study.

2. The two procedures had significantly different compile and execution times. The compile time for the iterative procedure averaged 16.84 seconds, compared with an average of

7.19 seconds for the modified Gauss-Jordan. The execution time of the miscellaneous linear dependencies, and the two dimensional analysis of variance problem differed 0.32 and 4.94 seconds respectively in favor of the modified Gauss-Jordan. The difference in the execution of the analysis of covariance, and the continuous data problem was 0.43 and 35.11 seconds respectively with the iterative being the quicker procedure.

3. The altering of the modified Gauss-Jordan in order to find the rank was investigated. It involved an arbitrary definition of the zero rank in order for it to run on the computer. A zero rank would result when the diagonal element was less than 0.0001. This method of finding the rank would not work for all matrices, therefore some users might prefer to print out the rows containing small diagonal elements in order to analyze the results.