COMPUTER PROGRAMS FOR MULTIVARIATE ANALYSIS IN GEOLOGY

Edited by

DANIEL F. MERRIAM

COMPUTER CONTRIBUTION 20
State Geological Survey
The University of Kansas, Lawrence
1968

Editor's Remarks

The following quote is reproduced with permission from the Stanford Research Institute Journal (no. 17 November, 1967, p. 13). We believe it is of considerable interest to all users of the COMPUTER CONTRIBUTION series.

> The U.S. market for computer systems is expected to grow from about $2 billion in 1965 to about $5.2 billion by 1975, representing nearly 0.5 percent of the total gross national product. With programming, operating, and support costs, it seems reasonable to predict that at least 1 percent of the nation's efforts will be devoted to computer activity by 1975.
>
> Third generation computers will continue to be marketed into the early 1970's; the major advance during this period will be in applications and systems use. The single most dramatic development will be the widespread adoption of communications in conjunction with computers. Dedicated systems (i.e., within a company or organization) for on-line and real-time operations will replace many of the present batch systems. Systems design and implementation will become quite advanced. Time-sharing, on-line systems also will be developed but will continue to be a relatively small factor until 1972.
>
> Fourth generation computers featuring large-scale integrated circuits (LSI) are expected to be available in 1972. Their design will facilitate further expansion of the time-sharing philosophy, utilizing large data bases for the concurrent operation of many dissimilar applications. This will expand the use of computers to a number of applications that previously could not support computers. By 1975, time-sharing should encompass such areas as management information systems, computer-aided instruction, hospital-patient monitoring, and others. Today such time-sharing is being used in the laboratory or experimentally.
>
> As a result of the effect of communications and time-sharing, the terminal unit will become the most important single segment of the computer peripheral market.
>
> Total systems are unlikely to become a reality by 1975. However, many discrete systems will be combined in time-sharing. This certainly will be an important step toward the total system concept. It is doubtful that a total system with all applications integrated will become a practical reality except in a few special cases--at least not in the foreseeable future.
>
> The use of computers in specific industries will remain essentially unchanged, with government and commerce being the big users. Any organization that uses extensive communications in the clerical conduct of its business will expand its data processing needs in the next few years. Banks, department stores, government agencies, widespread manufacturers, and distributors will expand and maintain their present leadership in computer use.
>
> The universities will continue to provide leadership in new applications and innovations of computer use, as well as becoming one of the big users of computers.

An up-to-date list of other COMPUTER CONTRIBUTIONS and related publications may be obtained by writing the Editor, COMPUTER CONTRIBUTION Series, Kansas Geological Survey, The University of Kansas, Lawrence, Kansas 66044.

COMPUTER PROGRAMS FOR MULTIVARIATE ANALYSIS IN GEOLOGY

Edited by

DANIEL F. MERRIAM

1968

CONTENTS

FORTRAN IV PROGRAM FOR MULTIVARIATE PALEONTOLOGIC ANALYSIS USING AN IBM SYSTEM/360 MODEL 40 COMPUTER

by

J. A. Wolleben, R. J. Pauken, and J. A. Dearien

University of Missouri

INTRODUCTION

Paleontologists, in attempting to make comparisons between sample sets based on large numbers of variables, are confronted with the problem of working with these variables simultaneously in order to obtain the most meaningful results. If the variables can be quantified then the computer program described in this paper can be used to aid in making paleontologic decisions.

A population described by p variables may be pictured as a cluster of sample points in p-dimensional space. The distribution of these points is specified by the mean vector and covariance matrix of the sample. Equality of sample covariance matrices is a specific characteristic that is of primary concern. R. A. Reyment has used homogeneity and heterogeneity of covariance matrices as an aid in the interpretation of evolutionary changes in ostracodes and physical sorting of fossil samples. The degree of distinctness of several samples can be measured by the "generalized distance" between the multivariate means of the samples. This if referred to as Mahanolobis' D^2 and has been used by Reyment (1960a, 1960b, 1962a, 1962b, 1963), Giles (1960), and Lerman (1965) as an aid in the interpretation of evolutionary changes of different fossil groups. The linear discriminant function also has been used by Reyment (1962b) to classify an unknown sample as belonging to one of several groups.

The computer program described here provides the following:

(a) sums of squares and cross-products matrices and mean vectors for each sample,

(b) the necessary values for a homogeneity test of covariance matrices,

(c) generalized distance (D^2) values and F values for determining the significance of the distances, and

(d) the constant terms of the linear discriminant function and the discriminant index for the classification of an unknown into one of several populations.

The program differs from previously described programs in several ways. A basic requirement for most multivariate methods including D^2 and discriminant function analysis is that the covariance matrices be homogeneous. This program computes the necessary values for a homogeneity test of covariance matrices. In addition the program makes generalized distance comparisons and provides discriminant functions for q-sample sets. Previously described programs such as Davis and Sampson (1966) compared only two sample sets at a time. In taxonomic and biostratigraphic analysis the former approach is much more applicable than the latter.

MATHEMATICAL DEVELOPMENT

A more detailed mathematical development of each technique can be found in Anderson (1958). In the following discussion, matrices and vectors are denoted by boldface characters.

The mean vector $\bar{\mathbf{X}}$ is a vector composed of the arithmetic means of two or more different variables in a sample. The covariance matrix $\mathbf{S}$ is derived from the matrix of the sums of squares and cross products $\mathbf{A}$. The sample covariance matrix is given by

$$\mathbf{S} = \frac{1}{N-1}\mathbf{A}. \qquad (1)$$

The test for homogeneity of covariance matrices is a generalization of the Bartlett test for homogeneity of variance in the univariate case and is developed by Anderson (1958). In the multivariate case the homogeneity of covariance is tested by

$$W_1 = V_1 \cdot \frac{n^{\frac{1}{2}pn}}{\prod_{g=1}^{q} n_g^{\frac{1}{2}pn_g}}, \qquad (2)$$

where V_1 is defined by

$$V_1 = \frac{\prod_{g=1}^{q} |\mathbf{A}_g|^{\frac{1}{2}n_g}}{|\mathbf{A}|^{\frac{1}{2}n}} \qquad (3)$$

and $n_g = N_g - 1$, $n = \Sigma n_g = N - q$, where q is the number of sample sets and p is the number of variables. The significance is found by calculating $-\log_e W_1$, and then using the asymptotic expansion of $-\rho \log_e W_1$, where ρ is defined as

$$\rho = 1 - \left(\Sigma \frac{1}{n_g} - \frac{1}{n} \right) \frac{2p^2 + 3p - 1}{6(p+1)(q-1)} , \quad (4)$$

and

$$W_2 = \frac{p(p+1) \left[(p-1)(p+2) \left(\Sigma \frac{1}{n_g^2} - \frac{1}{n^2} \right) - 6(q-1)(1-\rho^2) \right]}{48\rho^2} . \quad (5)$$

Thus

$$Pr\left\{-2\rho \log_e W_1 \leq z\right\} = Pr\left\{\chi_f^2 \leq z\right\} + w_2 \left[Pr\left\{\chi_{f+4}^2 \leq z\right\} - Pr\left\{\chi_f^2 \leq z\right\} \right] + O(n^{-3}) \quad (6)$$

and the number of degrees of freedom, f, is $\frac{(q-1)\, p\, (p+1)}{2}$ (Anderson, 1958). This test takes account of the relative degree of scatter but not orientation heterogeneity.

The generalized statistical distance between populations i and j as defined by Mahanolobis is:

$$D^2 = (\bar{\mathbf{X}}_i - \bar{\mathbf{X}}_j)' \mathbf{S}^{-1} (\bar{\mathbf{X}}_i - \bar{\mathbf{X}}_j) , \quad (7)$$

where $\bar{\mathbf{X}}_i$ and $\bar{\mathbf{X}}_j$ are the respective sample mean vectors and $\mathbf{S}^{-1} = \left[\frac{1}{q} (\mathbf{S}_1 + \mathbf{S}_2 + \ldots . \mathbf{S}_q) \right]^{-1}$, the pooled inverse covariance matrix of all samples. Significance of the difference between multivariate means is tested by the variance ratio

$$F = \frac{(N_1 N_2)(N_1 + N_2 - p - 1)}{p(N_1 + N_2)(N_1 + N_2 - 2)} D^2 \quad (8)$$

with p and $(N_1 + N_2 - p - 1)$ degrees of freedom.

The linear discriminant function coefficients for three or more samples are found by the equation

$$\lambda'_{ij} = \mathbf{S}^{-1} (\bar{\mathbf{X}}_i - \bar{\mathbf{X}}_j). \quad (9)$$

The discriminant index is found by the equation

$$R_{ij} = -\tfrac{1}{2} (\bar{\mathbf{X}}_i + \bar{\mathbf{X}}_j) \mathbf{S}^{-1} (\bar{\mathbf{X}}_i - \bar{\mathbf{X}}_j). \quad (10)$$

From equations 9 and 10 we obtain the discriminant function

$$\mu_{ij} = \lambda'_{ij} \bar{\mathbf{X}}' + R_{ij} , \quad (11)$$

also $\mu_{ji} = -\mu_{ij}$. If there are equal a priori probabilities, the best set of regions of classification are

$$R_1 : \mu_{12} \geq 0;\ \mu_{13} \geq 0;\ \mu_{1m} \geq 0$$
$$R_2 : \mu_{21} \geq 0;\ \mu_{23} \geq 0;\ \mu_{2m} \geq 0$$
$$\vdots \qquad \qquad \vdots$$
$$R_m : \mu_{m1} \geq 0;\ \mu_{m2} \geq 0;\ \mu_{mm} \geq 0 \ .$$

This method divides the variable space into m regions of classification; an unknown is placed into one of these regions by applying the discriminant function.

PROGRAM DESCRIPTION

This program computes mean vectors and sum of squares and cross-products matrices for 20 sample sets, using up to 15 variables, and having up to 99 samples in each set. The determinants of each matrix are computed using a pivotal condensation method described by McCormick and Salvadori (1964) and become terms in the W_1 equation for determining heterogeneity of covariance matrices. Sums of squares and cross products are accumulated and used to obtain the pooled variance - covariance matrix of all the samples. The pooled covariance matrix is inverted and the inverted matrix used in the solution of the D^2 equations and the linear discriminant function equations.

Output from the program consists of determinants of each sums of squares and cross-products matrix and the determinant of the pooled sums of squares and cross-products matrix; V_1, W_1, ρ, W_2, and $-\rho \log_e W_1$ values and degrees of freedom for testing homogeneity of covariance matrices; mean vectors of each sample; pooled variance - covariance matrix; inverted variance - covariance matrix; Mahalanobis' generalized distance (D^2) values, F-values, and degrees of freedom for all sample comparisons; constant terms of the linear discriminant functions; and the discriminant indices.

The program was written to run on an IBM System/360 Model 40 computer with 65K bits of core storage, automatic divide, and indirect addressing. Minor alterations make possible the use of this program on an IBM 1620 computer.

Data are read into the program by the following procedure. The total number of sample sets to be considered and the number of variables measured on the sample sets are identified on the first data card. All samples must have the same number of variables.

Card 1. This card contains two two-digit numbers in the first four columns. These must be right justified in columns 2 and 4. The first number (columns 1 and 2) is the number of sample sets to be read-in. The second number is the number of variables and also the horizontal dimension of the input matrix.

Cards 2 and 3 supply information to the program about each sample. Therefore, each set of sample data cords will be preceded by these two cards each with information about that set.

Card 2. This card contains a two-digit number in the first two columns and is right justified. This number identifies that set as being group 1, 2, 3, ... or 20.

Card 3. This card also contains a two-digit number in the first two columns that is right justified. This digit, the number of specimens measured for that sample, is the vertical dimension of the input matrix.

The data format, statement 6 in the program listing (6 FORMAT 8F 10.0), may be adjusted by the programmer. If the format is unaltered the data are punched continuously in 10-digit fields, one row of the input matrix at a time from left to right. If, for example, the matrices were formed by measuring six variables the format would read (6 FORMAT 6F 10.0), again each number being placed in a 10-digit field.

All possible comparisons are punched in a triangular output matrix. Rows of the matrix as well as the specific comparisons are identified as they are punched.

The program has been tested using data published by various authors. The covariance matrix homogeneity test values were obtained from data given by Anderson (1958, p. 256) and correspond with the values obtained by Anderson. D^2 values were obtained from data given by Lerman (1965, p. 427) and correspond with his D^2 values. The discriminant function portion of the program was tested using data published by Reyment (1962b, p. 198, 199) and Miller and Kahn (1962, p. 265, 266, 281).

The length of time required to complete the program depends upon the number of samples, sample sets and variables used. Approximately 2 1/4 minutes are required for compiling and execution of an analysis of 17 sample sets with 13 variables and 20 samples per set.

The program was written by the authors at the University of Missouri as a part of a quantitative paleontology project. Machine time was donated by the University of Missouri Computer Center.

REFERENCES

Anderson, T. W., 1958, An introduction to multivariate statistical analysis: John Wiley & Sons, New York, 374 p.

Davis, J. C., and Sampson, R. J., 1966, FORTRAN II program for multivariate discriminant analysis using an IBM 1620 computer: Kansas Geol. Survey Computer Contr. 4, 8 p.

Giles, E., 1960, Multivariate analysis of Pleistocene and Recent coyotes (*Canis latrans*) from California: Univ. California Publ. Geol. Sci., v. 36, no. 8, p. 369-390.

Lerman, A., 1965, Evolution of *Exogyra* in the Late Cretaceous of the southeastern United States: Jour. Paleontology, v. 39, no. 3, p. 414-435.

McCormick, J. M., and Salvadori, M. G., 1964, Numerical methods in FORTRAN: Prentice-Hall, Englewood Cliffs, New Jersey, 324 p.

Miller, R. L., and Kahn, J. S., 1962, Statistical analysis in the geological sciences: John Wiley & Sons, New York, 483 p.

Reyment, R. A., 1960a, Notes on the study of evolutionary changes in ostracods: 21st Intern. Geol. Congr. Rep., Copenhagen, sec. 6, p. 7-17.

Reyment, R. A., 1960b, Studies on Nigerian Upper Cretaceous and lower Tertiary Ostracoda, pt. I: Stockholm Contr. Geol., v. 7, 238 p.

Reyment, R. A., 1962a, Observations on homogeneity of covariance matrices in paleontologic biometry: Biometrics, v. 18, p. 1-11.

Reyment, R. A., 1962b, Biometric study of *Actinocamax verus* s.l. from the Upper Cretaceous of the Russian Platform: Stockholm Contr. Geol., v. 9, p. 147-206.

Reyment, R. A., 1963, Studies on Nigerian Upper Cretaceous and Lower Tertiary Ostracoda, pt. 2: Stockholm Contr. Geol., v. 10, 286 p.

```
/WATFOR         000,TIME=5
C *********************************************************
C      MULTIVARIATE PALEONTOLOGIC ANALYSIS OF UP TO 20 SAMPLE SETS
C      USING UP TO 15 VARIABLES AND 99 SAMPLES PER SET
C      VALUES FOR TEST OF VARIANCE-COVARIANCE HOMOGENIETY, GENERALIZED
C      DISTANCE, AN F-VALUE FOR A TEST OF EQUALITY OF MULTIVARIATE
C      MEANS, AND DISCRIMINANT EQUATIONS ARE COMPUTED
C      COMPUTER= IBM 360-40
C      LANGUAGE= FORTRAN IV
C *********************************************************
C      JAMES A. WOLLEBEN, ROBERT J. PAUKEN, AND JOHN A. DEARIEN
C      UNIVERSITY OF MISSOURI  GEOLOGY DEPARTMENT   5-17-67
C *********************************************************
C      THE FIRST DATA CARD CONTAINS TWO NUMBERS OF TWO DIGITS EACH
C      THESE NUMBERS GO IN THE FIRST FOUR COLUMNS, RIGHT JUSTIFIED
C      THE FIRST NUMBER IS THE NUMBER OF SAMPLES
C      THE SECOND NUMBER IS THE NUMBER OF VARIABLES
C *********************************************************
      DIMENSION XA(15,99),XB(15,15),XC(15,15),XVS(20,15),XV(15),R(20)
      DIMENSION VM(15),DMV(15),XVCVI(15),DISC(15),KM(20)
      DIMENSION SMV(15),A(15,15),DS(20),CC(20)
  616 READ (5,1) N1,M
    1 FORMAT (I2,I2)
      DO 22 I=1,M
      DO 22 K=1,M
   22 XC(I,K)=0.0
      DF=0.0
      M2=0
      M3=0
      M4=0
      WRITE (6,619)
  619 FORMAT (/,25X,19HRESULTS OF ANALYSIS,//)
      WRITE (6,55)
   55 FORMAT(2X,69HDETERMINANTS AND SUMS OF SQUARES AND CROSS PRODUCTS O
     1F INPUT MATRICES,/)
      NUB=0
 2222 CONTINUE
      JL=NUB+1
C *********************************************************
C      TWO TWO-DIGIT CONTROL CARDS GO IN FRONT OF EACH SAMPLE
C      THE FIRST CARD HAS A NUMBER SPECIFYING THE GROUP NAME
C      IN THE FIRST TWO COLUMNS, RIGHT JUSTIFIED
C      THE SECOND CARD HAS THE NUMBER OF SPECIMENS FOR THAT SAMPLE
C      IN THE FIRST TWO COLUMNS, RIGHT JUSTIFIED
C *********************************************************
      READ (5,2) N6
      READ (5,2) N5
    2 FORMAT (I2)
      KM(JL)=N6
      DO 3 J=1,N5
    3 READ (5,6) (XA(I,J),I=1,M)
C *********************************************************
C      THE FOLLOWING FORMAT SHOULD BE CHANGED TO FIT THE
C      NUMBER OF VARIABLES IN THE INPUT MATRIX
C *********************************************************
    6 FORMAT (4F8.0)
      R(JL)=N5
      DF=DF+R(JL)
      SN=N1
      AN=N5
```

```
      DO 7 I=1,M
      SUM=0.0
      DO 8 J=1,N5
    8 SUM=SUM+XA(I,J)
    7 XV(I)=SUM/AN
      DO 511 I=1,M
      DO 511 K=1,M
      SUM=0.0
      DO 9 J=1,N5
    9 SUM=SUM+(XA(I,J)-XV(I))*(XA(K,J)-XV(K))
      XB(I,K)=SUM
  511 CONTINUE
  515 WRITE (6,525) JL
  525 FORMAT(2X,47HSUMS OF SQUARES AND CROSS PRODUCTS MATRIX GROUP,I2,/)
      DO 517  I=1,M
  517 WRITE (6,13) (XB(I,K),K=1,M)
      GO TO 34
   36 CONTINUE
      DO 37  K=1,M
      DO 37 J=1,M
   37 A(K,J)=XC(K,J)
   34 M4=M4+1
      IF (N1-M4) 61,60,60
   60 DO 99 K=1,M
      DO 99 J=1,M
   99 A(K,J)=XB(K,J)
   61 N=M
      K=2
      L=1
    5 DO 10 I=K,N
      IF(A(L,L)) 385,386,385
  385 RATIO=A(I,L)/A(L,L)
      DO 10 J=K,N
   10 A(I,J)=A(I,J)-A(L,J)* RATIO
      IF(K-N) 15,20,20
   15 L=K
      K=K+1
      GO TO 5
   20 DETERM=1
      DO 25 L=1,N
   25 DETERM=DETERM*A(L,L)
      IF (N1-M4) 41,38,38
   38 WRITE (6,35) JL,DETERM
   35 FORMAT(2X,17HDETERMINANT GROUP,I2,2X,2HIS,E16.8,/)
      DS(JL)=DETERM
      DO 11 I=1,M
      DO 11 K=1,M
      XVS(JL,I)=XV(I)
   11 XC(I,K)=XC(I,K)+XB(I,K)
      NUB=NUB+1
      IF (NUB-N1)2222,48,48
   48 CONTINUE
      GO TO 389
  386 WRITE (6,387)
  387 FORMAT(2X,49HTHE DETERMINANT OF THIS SAMPLE IS ZERO, JOB ENDED)
      GO TO 616
  389 DF1=DF-SN
      IF (N1-M4) 41,36,41
   41 SDET=DETERM
      WRITE (6,42) SDET
```

```
 42 FORMAT(2X,28HDETERMINANT OF POOLED MATRIX,E16.8,//)
    WRITE (6,618)
618 FORMAT(2X,27HVALUES FOR HOMOGENIETY TEST,/)
    HT=0.0
    DO 43 I=1,N1
    CC(I)=((R(I)-1.0)*0.5)*ALOG(ABS(DS(I)))
 43 HT=HT+CC(I)
    HOT=HT-(0.5*DF1)*ALOG(ABS(SDET))
    RP=M
    HH=0.0
    DO 304 I=1,N1
304 HH=HH+(0.5*RP*(R(I)-1.0))*ALOG(R(I)-1.0)
    GG=(0.5*RP*DF1)*ALOG(DF1)-HH+HOT
    WRITE (6,305) GG
305 FORMAT(2X,14HLOG W-ONE IS  ,E16.8)
    SS=0.0
    DO 306 I=1,N1
306 SS=SS+(1.0/(R(I)-1.0))
    XW=SS-(1.0/DF1)
    RO=1.0-XW*((2.0*(RP**2.0))+((3.0*RP)-1.0))/(6.0*(RP+1.0)*(SN-1.0))
    WRITE (6,307) RO
307 FORMAT(2X,10HRHO IS    ,E16.8)
    RR=0.0
    DO 308 I=1,N1
308 RR=RR+(1.0/((R(I)-1.0)**2.0))
    RW=(RR-(1.0/(DF1*DF1)))*(RP-1.0)*(RP+2.0)
    WW=RW-(6.0*(SN-1.0)*(1.0-RO)**2.0)
    QW=(WW*RP*(RP+1.0))/(48.0*RO*RO)
    WRITE (6,309) QW
309 FORMAT(2X,10HW-TWO IS  ,E16.8)
    RLW1=RO*GG
    WRITE (6,310) RLW1
310 FORMAT(2X,15HRHO LOG W1 IS  ,E16.8)
    DEG=((SN-1.0)*RP*(RP+1.0))/2.0
    WRITE (6,311) DEG
311 FORMAT(2X,20HDEGREES OF FREEDOM  ,F5.0,//)
    WRITE (6,57)
 57 FORMAT(2X,30HMEAN VECTORS OF INPUT MATRICES,/)
    DO 18 L=1,N1
    WRITE (6,17) L
 17 FORMAT(2X,17HMEAN VECTOR GROUP,I2)
 18 WRITE (6,13) (XVS(L,I),I=1,M)
 13 FORMAT (13F9.3)
    DO 28 I=1,M
    DO 28 K=1,M
 28 XC(I,K)=(XC(I,K))/(DF-SN)
    WRITE (6,14)
 14 FORMAT (//,2X,24HPOOLED COVARIANCE MATRIX,/)
    DO 16 I=1,M
 16 WRITE (6,13) (XC(I,K),K=1,M)
    DO 26 I=1,M
    Y=XC(I,I)
    XC(I,I)=1.00
    DO 23 J=1,M
 23 XC(I,J)=XC(I,J)/Y
    DO 26 K=1,M
    IF (K-I)24,26,24
 24 Y=XC(K,I)
    XC(K,I)=0.0
    DO 27 J=1,M
```

```
 27 XC(K,J)=XC(K,J)-Y*XC(I,J)
 26 CONTINUE
    WRITE (6,83)
 83 FORMAT (//,2X,33HINVERTED POOLED COVARIANCE MATRIX,/)
    DO 31 I=1,M
 31 WRITE (6,13) (XC(I,J),J=1,M)
    N3=N1
 69 M3=M3+1
    WRITE (6,77) M3
 77 FORMAT (//,2X,3HROW,1X,I2,1X,16HOF OUTPUT MATRIX,/)
    LL=M3
    N2=N1-1
    N3=N3-1
    M2=M2+1
    KK=M2
    ID1=KM(KK)
    DO 68 K=1,N3
    LL=LL+1
    ID2=KM(LL)
    DO 44 I=1,M
 44 VM(I)=XVS(LL,I)
    DO 51 I=1,M
    J=M2
 51 DMV(I)=XVS(J,I)-VM(I)
    DO 73 I=1,M
    J=M2
 73 SMV(I)=(VM(I)+XVS(J,I))/2.00
    DO 54 J=1,M
    SUM =0.0
    DO 19 I=1,M
 19 SUM=SUM+DMV(I)*XC(I,J)
 54 XVCVI(J)=SUM
    SUM=0.0
    J=M2
    X1=R(J)
    X2=R(LL)
    DOD=(X1+X2-1.0-RP)
    WRITE (6,151) ID1,ID2
151 FORMAT (2X,6HGROUP ,I2,1X,20HCOMPARED WITH GROUP ,I2)
    DO 32 J=1,M
 32 SUM=SUM+XVCVI(J)*DMV(J)
    DSQ=SUM
    WRITE (6,617)
617 FORMAT (/)
    WRITE (6,56) DSQ
 56 FORMAT(2X,14HD-SQUARED IS  ,E16.8)
    D=SQRT(ABS(DSQ))
    DO 59 I=1,M
    SUM=0.0
    DO 58 J=1,M
 58 SUM=SUM+XC(I,J)*DMV(J)
 59 DISC(I)=SUM
    SUM=0.0
    DO 79 I=1,M
 79 SUM=SUM+SMV(I)*DISC(I)
    XINDX=SUM
    EPH=(((X1*X2)*(X1+X2-RP-1.0))/((RP*(X1+X2))*(X1+X2-2.0)))* DSQ
    WRITE (6,97) EPH
 97 FORMAT(2X,14HTHE F VALUE IS,E16.8)
    WRITE (6,98) RP,DOD
```

```
   98 FORMAT(2X,23HF IS DISTRIBUTED WITH   ,F5.0,2X,5HAND   ,F5.0,2X,18HDE
     1GREES OF FREEDOM)
      WRITE (6,33)
   33 FORMAT(2X,42HTHE DISCRIMINANT FUNCTION COEFFICIENTS ARE)
      WRITE (6,13) (DISC(I),I=1,M)
      WRITE (6,93) XINDX
   93 FORMAT(2X,21HDISCRIMINANT INDEX IS,E16.8)
      WRITE (6,617)
   68 CONTINUE
      IF (N2-M3)49,49,69
   49 CONTINUE
      WRITE (6,47)
   47 FORMAT(2X,17HREAD IN MORE DATA)
      GO TO 616
      END


/DATA
0404  INPUT DATA- OREODONTS- MILLER AND KAHN (1962)
01    SUBDESMATOCHOERUS SP.
11
  47.0     99.0     26.0     15.0
  42.0     93.0     26.0     16.0
  40.0     90.0     22.0     13.0
  46.0    100.0     22.0     11.0
  46.0     96.0     24.0     16.0
  42.0     88.0     26.0     15.0
  43.0     89.0     23.0     14.0
  44.0     78.0     23.0     13.0
  44.0     90.0     25.0     11.0
  47.0     99.0     27.0     15.0
  47.0     92.0     27.0     13.0

02    MEGOREODON GIGAS LOOMISI
11
  78.0    165.0     35.0     18.0
  77.0    165.0     37.0     19.0
  65.0    148.0     30.0     20.0
  74.0    163.0     31.0     15.0
  65.0    169.0     31.0     16.0
  70.0    176.0     34.0     23.0
  69.0    161.0     28.0     13.0
  67.0    178.0     31.0     14.0
  65.0    174.0     34.0     18.0
  64.0    168.0     28.0     13.0
  68.0    166.0     32.0     15.0

03    O. OSBORNI
15
  42.0     81.0     15.0      8.0
  48.0     83.0     18.0      8.6
  45.0     87.0     18.0      9.0
  48.0     83.0     17.0      8.0
  46.0     84.0     16.0      6.1
  51.0     87.0     21.0      7.9
  46.0     80.0     17.0      7.0
  50.0     90.0     18.0      8.1
  46.0     85.0     16.0      6.5
```

48.0	85.0	15.0	7.2
47.0	85.0	17.0	8.0
49.0	83.0	18.0	7.7
43.0	79.0	15.0	7.1
47.0	87.0	19.0	7.5
46.0	87.0	18.0	8.0

04 PSEUDODESMATOCHOERUS
08

60.0	114.0	27.0	20.0
60.0	118.0	31.0	19.0
60.0	111.0	31.0	21.0
58.0	102.0	30.0	20.0
55.0	116.0	28.0	20.0
59.0	117.0	29.0	17.0
59.0	114.0	24.0	17.0
60.0	121.0	25.0	19.0

/END

RESULTS OF ANALYSIS

DETERMINANTS AND SUMS OF SQUARES AND CROSS PRODUCTS OF INPUT MATRICES

SUMS OF SQUARES AND CROSS PRODUCTS MATRIX GROUP 1

58.545	82.272	18.454	-.272
82.272	407.636	30.727	18.363
18.454	30.727	36.545	15.272
-.272	18.363	15.272	31.636

DETERMINANT GROUP 1 IS .12073238E+08

SUMS OF SQUARES AND CROSS PRODUCTS MATRIX GROUP 2

248.181	-29.909	90.272	37.818
-29.909	676.545	64.636	-8.090
90.272	64.636	80.909	58.727
37.818	-8.090	58.727	100.181

DETERMINANT GROUP 2 IS .28057376E+09

SUMS OF SQUARES AND CROSS PRODUCTS MATRIX GROUP 3

80.400	57.200	37.600	4.039
57.200	125.600	39.800	9.820
37.600	39.800	38.400	7.860
4.039	9.820	7.860	8.157

DETERMINANT GROUP 3 IS .78431321E+06

SUMS OF SQUARES AND CROSS PRODUCTS MATRIX GROUP 4

20.875	12.125	.125	-1.875
12.125	230.874	-39.125	-22.125
.125	-39.125	48.875	12.875
-1.875	-22.125	12.875	14.875

DETERMINANT GROUP 4 IS .21030275E+07

DETERMINANT OF POOLED MATRIX .92238759E+10

VALUES FOR HOMOGENIETY TEST

LOG W-ONE IS   -.34342310E+02
RHO IS   .81371670E+00
W-TWO IS   .11569229E+00
RHO LOG W1 IS   -.27944911E+02
DEGREES OF FREEDOM   30.

MEAN VECTORS OF INPUT MATRICES

MEAN VECTOR GROUP 1
44.363 92.181 24.636 13.818
MEAN VECTOR GROUP 2
69.272 166.636 31.909 16.727
MEAN VECTOR GROUP 3
46.800 84.400 17.200 7.646
MEAN VECTOR GROUP 4
58.875 114.125 28.125 19.125

POOLED COVARIANCE MATRIX

9.951	2.968	3.572	.968
2.968	35.137	2.342	-.049
3.572	2.342	4.993	2.310
.968	-.049	2.310	3.776

INVERTED POOLED COVARIANCE MATRIX

.139	-.004	-.112	.033
-.004	.029	-.016	.011
-.112	-.016	.384	-.206
.033	.011	-.206	.382

ROW 1 OF OUTPUT MATRIX

GROUP 1 COMPARED WITH GROUP 2

D-SQUARED IS   .20265159E+03
THE F VALUE IS   .23684904E+03
F IS DISTRIBUTED WITH   4. AND   17. DEGREES OF FREEDOM
THE DISCRIMINANT FUNCTION COEFFICIENTS ARE
-2.428 -2.038 1.837 -1.298
DISCRIMINANT INDEX IS -.36960935E+03

GROUP 1 COMPARED WITH GROUP 3

D-SQUARED IS   .21973263E+02
THE F VALUE IS   .30503747E+02
F IS DISTRIBUTED WITH   4. AND   21. DEGREES OF FREEDOM
THE DISCRIMINANT FUNCTION COEFFICIENTS ARE
-1.004 .192 1.731 .834
DISCRIMINANT INDEX IS .16335360E+02

```
GROUP  1 COMPARED WITH GROUP  4


D-SQUARED IS      .42713104E+02
THE F VALUE IS    .40729523E+02
F IS DISTRIBUTED WITH     4.  AND    14.  DEGREES OF FREEDOM
THE DISCRIMINANT FUNCTION COEFFICIENTS ARE
-1.709   -.599   1.752  -2.046
DISCRIMINANT INDEX IS  -.13758357E+03



ROW  2 OF OUTPUT MATRIX

GROUP  2 COMPARED WITH GROUP  3


D-SQUARED IS      .23320833E+03
THE F VALUE IS    .32374472E+03
F IS DISTRIBUTED WITH     4.  AND    21.  DEGREES OF FREEDOM
THE DISCRIMINANT FUNCTION COEFFICIENTS ARE
 1.423   2.230   -.105   2.133
DISCRIMINANT INDEX IS   .38594469E+03


GROUP  2 COMPARED WITH GROUP  4


D-SQUARED IS      .84478001E+02
THE F VALUE IS    .80554873E+02
F IS DISTRIBUTED WITH     4.  AND    14.  DEGREES OF FREEDOM
THE DISCRIMINANT FUNCTION COEFFICIENTS ARE
  .719   1.438   -.085   -.748
DISCRIMINANT INDEX IS   .23202580E+03



ROW  3 OF OUTPUT MATRIX

GROUP  3 COMPARED WITH GROUP  4


D-SQUARED IS      .64892646E+02
THE F VALUE IS    .72550782E+02
F IS DISTRIBUTED WITH     4.  AND    18.  DEGREES OF FREEDOM
THE DISCRIMINANT FUNCTION COEFFICIENTS ARE
 -.704   -.791    .020  -2.881
DISCRIMINANT INDEX IS  -.15391892E+03


READ IN MORE DATA
```

KANSAS GEOLOGICAL SURVEY COMPUTER PROGRAM
THE UNIVERSITY OF KANSAS, LAWRENCE

PROGRAM ABSTRACT

Title (If subroutine state in title):

FORTRAN IV program for multivariate paleontologic analysis using an IBM System/360 Model 40 computer

Date: May 17, 1967

Author, organization: J.A. Wolleben, R.J. Pauken, and J.A. Dearien, Dept. of Geology, University of Missouri, Columbia, Missouri

Direct inquiries to: Authors, or

Name: D.F. Merriam   Address: Kansas Geological Survey
Lawrence, Kansas 66044

Purpose/description: Computes linear discriminant functions and generalized distances for up to 20 groups, each having up to 15 variables and 99 samples. A test for homogeneity is also made.

Mathematical method:

Restrictions, range: 20 groups, with up to 15 variables and 99 samples per group.

Computer manufacturer: IBM   Model: System/360 Model 40

Programming language: FORTRAN IV

Memory required: 65 K   Approximate running time: 2 1/4 min. for 17 groups with 13 variables.

Special peripheral equipment required: None

Remarks (special compilers or operating systems, required word lengths, number of successful runs, other machine versions, additional information useful for operation or modification of program)

The program has been adapted to run on the IBM 1620 with 40 K memory (FORTRAN II) and the GE 625.

R-MODE FACTOR ANALYSIS PROGRAM IN FORTRAN II FOR THE IBM 1620 COMPUTER

by

Robert J. Sampson
Idaho State University

INTRODUCTION

Factor analysis is a well-known multivariate procedure for examining relationships between variables. The purpose of the technique is to reduce a multiplicity of variables to a smaller set of underlying variables, or "factors." Originated in the early 1900's by Spearman (1904) for the analysis of psychological traits, factor analysis has been developed into a powerful psychological tool by Cyril Burt and associates (see Burt, 1941, 1949, 1950). Although originally created by nonstatisticians, a statistical framework for the procedure has been erected by Kendall (1950), Bartlett (1950), Lawley and Maxwell (1963), and Jöreskog (1963). In recent years, factor analysis has played a large role in the development of numerical taxonomy (Sokal and Sneath, 1963). The technique also has been applied in geology to problems of paleontologic taxonomy (Kaesler, 1966; Pitcher, 1966), petrology (Harbaugh and Demirmen, 1964; Imbrie and Purdy, 1962; Imbrie and Van Andel, 1964; Klovan, 1966; Toomey, 1966) and mineral exploration (Harris, 1966). Excellent texts have been written by Fruchter (1954), Harman (1960), and Kendall (1957).

Several computer programs have been written to perform factor analysis including those by Cooley and Lohnes (1962), Imbrie (1963), Manson and Imbrie (1964), Dixon (1964), and IBM (1967). The program listed here is similar to other programs but differs in certain aspects.

This program computes factors in the R-mode. That is, it searches for structure within the variables and the factors express sources of variation affecting observations. This is in contrast to Q-mode analysis, which separates observations into factors that may be regarded as subsets of the original sample. The program is designed for operation on an IBM 1620 computer and may be adapted easily to other small computer systems.

Any number of data points may be processed with this program because data are not stored internally. Other programs store raw data in a core matrix or on a disc, which limits the number of observations to a few hundred. In some applications, this restriction is not severe, but many problems in the earth sciences require analysis of hundreds or thousands of observations. For example, the 100-brine analyses used in the accompanying example represent only a part of 3,016 observations contained in the original data set. The unlimited input feature has made this program sufficiently desirable so that it has been translated into FORTRAN IV for the GE 625 and the IBM 1130.

Theory and computational details of factor analysis are not described here; those unfamiliar with the procedure are referred to Manson and Imbrie (1964) or IBM (1967). This program operates in R-mode and generates the factor matrix, rotated (varimax) factor matrix, and oblique projection matrix. Up to 30 variables per observation (3 data cards per observation) are read in, and up to 20 of these may be selected for factor analysis. The program will compute and list the following information:

(1) correlation matrix,
(2) for each variable, the range, standard deviation, mean, standard error of the mean, variance, and coefficient of variation,
(3) optionally, the multiple regression of all variables on the first variable,
(4) eigenvalues and eigenvectors (principal components),
(5) factor matrix for specified number of factors,
(6) rotated factor matrix,
(7) communalities,
(8) variance at each iteration step, and
(9) oblique projection matrix.

Machine time for this project was donated by Idaho State University Computer Center on equipment supplied in part by NSF Grant GP-2275. Additional time was donated by the Computation Center at Wichita State University.

OPERATING INSTRUCTIONS

The factor analysis program consists of eight object decks which are called and loaded by the program as they are needed. The message LOAD DATA will appear as each of these enters the machine, but should be ignored after Program 1. Additional information will be requested via the console typewriter as required. Decks should be compiled with subroutines prior to execution. All output is punched unless otherwise noted.

PROGRAM 1

Input

Control Card

Columns 1-2. Total numbers of variables, maximum of 30.

Columns 3-4. Number of variables that are to be used, maximum of 20.

Columns 5-6. Number of first variable to be used.

Columns 7-8. Number of second variable to be used.

⋮ ⋮

Columns 43-44. Number of twentieth variable to be used.

Title Card

Columns 1-72. Any desired alphanumeric information.

Columns 73-80. Reserved for page numbering.

Data Cards

This is the first of any number of data cards; there may be as many as 3 cards per data set. Data are punched 10 variables to a card in 7 column fields.

End Card

Columns 1-7. Contain the number 9.0E48

Calculations

The correlation matrix is calculated for all variables and stored in COMMON. Incidental statistics are computed and punched.

Output

Output from Program 1 appears on two pages.

Page 1- Range, mean, and standard error of mean for each variable.

Page 2- Standard deviation, variance, and coefficient of variation for each variable.

Note: From this point, the next object deck is loaded by the one presently in execution. Object decks should be stacked in the input hopper and will be entered automatically by the EXIT routine incorporated in the preceding deck. Variables are labeled with their original variable numbers, as defined on the control card, in output from Program 1. Variables, however, are numbered sequentially on output from all subsequent programs.

PROGRAM 2

Calculations

This program contains format and output routines for the correlation matrix.

Output

Program 2 produces pages 3 and 4 containing the correlation matrix.

PROGRAM 3

Calculations

Program 3 calculates regression of variables on the first variable, goodness of fit, and multiple correlation coefficient.

Output

Program 3 produces page 5 containing the goodness of fit and multiple correlation coefficient of a multiple regression on the first variable and the standard partial and partial regression equations.

PROGRAM 4

Calculations

Eigenvalues and eigenvectors are computed and stored in COMMON.

Output

There is no output from Program 4.

PROGRAM 5

Output

Program 5 produces page 6, which contains eigenvalues and percent contribution of eigenvalues calculated by Program 4. These are punched and typed out on the console typewriter. The program also produces page 7 containing the eigenvectors. The message ENTER NUMBER OF FACTORS TO ROTATE will be typed on the console typewriter. The number of factors desired, up to 14, should be entered as a four digit, right-justified number on the typewriter.

Calculations

After entering the number of rotated factors, the factor matrix is computed.

Output

Page 8, containing the factor matrix, is then produced.

PROGRAM 6

Calculations

The factor matrix is rotated by varimax rotation to produce rotated factor matrix, and stored in COMMON.

Output

If SENSE SWITCH 1 is on, the variance is typed out at each step of the rotation iteration.

PROGRAM 7

Calculations

This program punches the rotated matrix and oblique projection matrix.

Output

Page 9- Variance of each iteration.

Page 10- Rotated factor matrix.

Page 11- Communalities.

PROGRAM 8

Calculations

Oblique projection matrix is calculated. The program then checks to see if any variables should be substituted. If so, substitutions are made and the oblique matrix recalculated.

Output

Oblique projection matrix.

The factor analysis program may be modified to perform a variety of calculations that stop short of a complete factor analysis. This may be done by using only part of the factor analysis program set. Possible program combinations include:

Program	Purpose
1, 2	Produces statistics of variables and correlation matrix.
1, 2, 3	Performs multiple regression on the first variable. Program 2 is optional.
1, 2, 4, 5	Performs principal components analysis. Program 2 is optional.

Oblique projection may be omitted by omitting Program 8. Program 3 may be omitted if multiple regression is not desired. Program 2 may be omitted if the correlation matrix is not required as output.

TEST EXAMPLE

The factor analysis program was tested on sample data listed in the Appendix. These data consist of 122 analyses of subsurface brines from the Arbuckle Group (Cambrian-Ordovician) in Kansas. The complete data set contains approximately 7,000 brine analyses obtained from the U.S. Bureau of Mines, Halliburton Company, Dow Chemical Company, and from published sources. The subset used as test data was selected for a study of brine type distributions in the Arbuckle and has 11 variables measured on each sample. The variables, in order, are:

(1) Calcium, in ppm,
(2) Magnesium, in ppm,
(3) Sodium, in ppm,
(4) Bicarbonate, in ppm,
(5) Sulphate, in ppm,
(6) Chloride, in ppm,
(7) Total dissolved salts, in ppm,
(8) Specific gravity,
(9) Temperature, C°,
(10) Electrical resistivity, in ohms/m^2/m, and
(11) pH.

In the analysis, four principal components accounted for about 70 percent of the variation in the system and were retained for factoring and rotation. The resulting factors and factor loadings are given on the final page of output. These factors correspond closely to those extracted by Betty Miller (personal communication, 1967) from brine data published by Ostroff (1967, p. 411-413), suggesting that they reflect fundamental characteristics of subsurface brines in general. Although no conclusions may be drawn from this preliminary study, it would seem that factor analysis may be useful in the characterization of oil-field waters.

REFERENCES

Bartlett, M. S., 1950, Tests of significance in factor analysis: British Jour. Psychology Stat. Sec., v. 3, p. 77-85.

Burt, C., 1941, Factors of the mind; an introduction to factor analysis in psychology: The Macmillan Co., New York, 509 p.

Burt, C., 1949, Alternative methods of factor analysis: British Jour. Psychology Stat. Sec., v. 2, p. 98-121.

Burt, C., 1950, Group factor analysis: British Jour. Psychology Stat. Sec., v. 3, p. 40-75.

Cooley, W. W., and Lohnes, P. R., 1962, Multivariate procedures for the behavorial sciences: John Wiley & Sons, New York, 211 p.

Dixon, W. J., ed., 1964, BMD biomedical computer programs: Health Ser. Computing Facility, Univ. California at Los Angeles, 620 p.

Fruchter, B., 1954, Introduction to factor analysis: D. Van Nostrand Co., Princeton, New Jersey, 280 p.

Harbaugh, J. W., and Demirmen, F., 1964, Application of factor analysis to petrologic variations of Americus Limestone (Lower Permian), Kansas and Oklahoma: Kansas Geol. Survey Sp. Dist. Publ. 15, 40 p.

Harmon, H. H., 1960, Modern factor analysis: Univ. Chicago Press, Chicago, 471 p.

Harris, D. V. P., 1966, Factor analysis, a tool for quantitative studies in mineral exploration: Pennsylvania State Univ., Mineral Industries, v. 2, p. GG1-GG37.

Imbrie, J., 1963, Factor and vector analysis programs for analyzing geologic data: Office of Naval Research, Geography Branch, ONR Task No. 389-135, Tech. Rpt. No. 6, 83 p.

Imbrie, J., and Purdy, E. G., 1962, Classification of modern Bahamian carbonate sediments, in Classification of carbonate rocks: Am. Assoc. Petroleum Geologists Mem. 1, p. 253-272.

Imbrie, J., and Van Andel, T. H., 1964, Vector analysis of heavy-mineral data: Geol. Soc. America Bull., v. 75, no. 11, p. 1131-1155.

IBM, 1967, 1130 statistical system (1130-CA-06s) user's Manual: IBM Appl. Prog. H 20-0333-0, Tech. Publ. Dept., White Plains, New York, p. 31-71.

Jöreskog, K. G., 1963, Statistical estimation in factor analysis: Almqvist and Wiksells Boktryckeri AB, Uppsala, 145 p.

Kaesler, R. L., 1966, Quantitative re-evaluation of ecology and distribution of Recent Foraminifera and Ostracoda of Todos Santos Bay, Baja California, Mexico: Univ. Kansas Paleontological Contr. Paper 10, 50 p.

Kendall, M. G., 1950, Factor analysis as a statistical technique: Jour. Roy. Stat. Soc., v. B12, p. 60-73.

Kendall, M. G., 1957, A course in multivariate analysis: Charles Griffin and Company, London, 185 p.

Klovan, J. E., 1966, The use of factor analysis in determining depositional environments from grain-size distributions: Jour. Sed. Pet., v. 36, no. 1, p. 115-125.

Lawley, D. N., and Maxwell, A. E., 1963, Factor analysis as a statistical method: Butterworth and Co., London, 117 p.

Manson, V., and Imbrie, J., 1964, FORTRAN program for factor and vector analysis of geologic data using an IBM 7090 or 7094/1401 computer system: Kansas Geol. Survey Sp. Dist. Publ. 13, 46 p.

Ostroff, A.G., 1967, Comparison of some formation water classification systems: Am. Assoc. Petroleum Geologists Bull., v. 51, no. 3, p. 404-416.

Pitcher, M. G., 1966, A factor analytic scheme for grouping and separating types of fossils, in Colloquium on classification procedures: Kansas Geol. Survey Computer Contr. 7, p. 30-41.

Sokal, R. R., and Sneath, P. H. A., 1963, Principles of numerical taxonomy: W. H. Freeman and Co., San Francisco, 359 p.

Spearman, C., 1904, General intelligence, objectively determined and measured: Am. Jour. Psychology, v. 15, p. 201-293.

Toomey, D. F., 1966, Application of factor analysis to a facies study of the Leavenworth Limestone (Pennsylvanian-Virgilian) of Kansas environs: Kansas Geol. Survey Sp. Dist. Publ. 27, 28 p.

Appendix.

```
C      CORRELATION
C      MULTIPLE REGRESSION
C      FACTOR ANALYSIS
C      PROGRAM 1
C      CORRELATION MATRIX GENERATION
C      ROBERT SAMPSON -- PROGRAMMER
C       4 / 26 / 1967
       BEGIN TRACE
       COMMON R(210),T(19),N,IPAGE,CONST,AM(20),SD(20),NS
       DIMENSION CP(210),SX(20),XMIN(20),XMAX(20),X(20),V(30),K(20)
       DO 100 I=1,210
  100  CP(I)=0.0
       DO 101 I=1,20
       SX(I)=0.0
       XMIN(I)=.99E 49
  101  XMAX(I)=-.99E 49
       T(19)=.57414745
       NS=0
       READ 1000,NI,N,K(1),K(2),K(3),K(4),K(5),K(6),K(7),K(8),K(9),
      1K(10),K(11),K(12),K(13),K(14),K(15),K(16),K(17),K(18),K(19),K(20)
       READ 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
      1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18)
    1  READ 1002,V(1),V(2),V(3),V(4),V(5),V(6),V(7),V(8),V(9),V(10)
       IF (V(1)-9.0E 48) 2,3,3
    2  IF (NI-10) 4,4,5
    5  READ 1002,V(11),V(12),V(13),V(14),V(15),V(16),V(17),V(18),
      1V(19),V(20)
       IF (NI-20) 4,4,14
   14  READ 1002,V(21),V(22),V(23),V(24),V(25),V(26),V(27),V(28),
      1V(29),V(30)
    4  NS=NS+1
       DO 110 I=1,N
       J=K(I)
  110  X(I)=V(J)
       DO 102 I=1,N
       SX(I)=SX(I)+X(I)
       IS=I*(I-1)/2
       DO 103 J=1,I
       JS=IS+J
       CP(JS)=CP(JS)+X(I)*X(J)
  103  CONTINUE
       IF (XMIN(I)-X(I)) 6,6,7
    7  XMIN(I)=X(I)
    6  IF (X(I)-XMAX(I)) 102,102,8
    8  XMAX(I)=X(I)
  102  CONTINUE
       GO TO 1
    3  AN=NS
       IPAGE=1
       PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
      1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
       PUNCH 1005,NS
       PUNCH 1003
       DO 104 I=1,N
       AM(I)=SX(I)/AN
       II=I*(I+1)/2
```

```
      SD(I)=SQRT((CP(II)-(SX(I)**2)/AN)/(AN-1.0))
      SE=SD(I)/SQRT(AN)
  104 PUNCH 1004,K(I),XMIN(I),XMAX(I),AM(I),SE
      IPAGE=2
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1006
      DO 107 I=1,N
      VAR=SD(I)**2
      CV=100.0*SD(I)/AM(I)
  107 PUNCH 1004,K(I),SD(I),VAR,CV
      AN1=AN-1.0
      DO 105 I=1,N
      IS=I*(I-1)/2
      DO 106 J=1,I
      JS=IS+J
      R(JS)=(CP(JS)-SX(I)*SX(J)/AN)/(SD(I)*SD(J)*AN1)
  106 CONTINUE
  105 CONTINUE
      A=EXIT(1.0)
 1000 FORMAT (22I2)
 1001 FORMAT (19A4,I3,1H-)
 1002 FORMAT (10F7.2)
 1003 FORMAT (8HVARIABLE,6X,7HMINIMUM,8X,7HMAXIMUM,24X,8HSTANDARD,/,
     17H NUMBER,8X,5HVALUE,10X,5HVALUE,11X,4HMEAN,6X,13HERROR OF MEAN)
 1004 FORMAT (/,I6,4X,4F15.7)
 1005 FORMAT (/,29HNUMBER OF SAMPLES EXAMINED = ,I6,/)
 1006 FORMAT (/,8HVARIABLE,7X,8HSTANDARD,21X,11HCOEFFICIENT,/,
     17H NUMBER,7X,9HDEVIATION,8X,8HVARIANCE,4X,12HOF VARIATION)
      END TRACE
      END
```

```
C      CORRELATION
C      MULTIPLE REGRESSION
C      FACTOR ANALYSIS
C      PROGRAM 2
C      CORRELATION MATRIX OUTPUT
C      ROBERT SAMPSON -- PROGRAMMER
C       4 / 26 / 1967
C      12244 - 12298 - 12580
       BEGIN TRACE
       COMMON R(210),T(19),N,IPAGE,C1,AM(20),SD(20),NS
       DIMENSION A(20,20)
       DO 100 I=1,N
       IS=I*(I-1)/2
       DO 101 J=1,I
       JS=IS+J
       A(I,J)=R(JS)
   101 A(J,I)=A(I,J)
   100 CONTINUE
       ID=0
       IND=N
       IT=1
    33 IPAGE=IPAGE+1
```

```
      PUNCH 1000,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1002
      DO 102 I=1,IND
      J=I+ID
      GO TO (1,2,3,4,5,6,7,8,9,10,10,10,10,10,10,10,10,10,10,10),I
    1 PUNCH 1001,J,A(J,IT)
      GO TO 102
    2 PUNCH 1001,J,A(J,IT),A(J,IT+1)
      GO TO 102
    3 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2)
      GO TO 102
    4 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2),A(J,IT+3)
      GO TO 102
    5 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2),A(J,IT+3),A(J,IT+4)
      GO TO 102
    6 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2),A(J,IT+3),A(J,IT+4),
     1A(J,IT+5)
      GO TO 102
    7 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2),A(J,IT+3),A(J,IT+4),
     1A(J,IT+5),A(J,IT+6)
      GO TO 102
    8 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2),A(J,IT+3),A(J,IT+4),
     1A(J,IT+5),A(J,IT+6),A(J,IT+7)
      GO TO 102
    9 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2),A(J,IT+3),A(J,IT+4),
     1A(J,IT+5),A(J,IT+6),A(J,IT+7),A(J,IT+8)
      GO TO 102
   10 PUNCH 1001,J,A(J,IT),A(J,IT+1),A(J,IT+2),A(J,IT+3),A(J,IT+4),
     1A(J,IT+5),A(J,IT+6),A(J,IT+7),A(J,IT+8),A(J,IT+9)
  102 CONTINUE
      IF (IND-10) 31,31,32
   32 ID=10
      IT=11
      IND=N-10
      GO TO 33
   31 B=EXIT(1.0)
 1000 FORMAT (19A4,I3,1H-,/)
 1001 FORMAT (/,I4,10F7.4)
 1002 FORMAT (27HCORRELATION COEFFICIENT (R),/)
      END TRACE
      END
```

```
C     CORRELATION
C     MULTIPLE REGRESSION
C     FACTOR ANALYSIS
C     PROGRAM 3
C     MULTIPLE REGRESSION
C     ROBERT SAMPSON -- PROGRAMMER
C      4 / 26 / 1967
C     09318 - 09732 - 12650
      BEGIN TRACE
      COMMON R(210),T(19),N,IPAGE,C1,AM(20),SD(20),NS
      DIMENSION A(20,20)
```

```
      DO 106 I=1,N
      IS=I*(I-1)/2
      DO 107 J=1,I
      JS=IS+J
      A(I,J)=R(JS)
  107 A(J,I)=A(I,J)
  106 CONTINUE
      DO 100 I=2,N
      DO 101 J=2,N
      IF (I-J) 1,101,1
    1 F=-A(J,I)/A(I,I)
      A(J,1)=A(J,1)+F*A(I,1)
      DO 102 K=I,N
  102 A(J,K)=A(J,K)+F*A(I,K)
  101 CONTINUE
  100 CONTINUE
      DO 105 I=2,N
  105 A(I,1)=A(I,1)/A(I,I)
      RR=0.0
      CONST=AM(1)
      DO 103 I=2,N
      A(I,2)=A(I,1)*SD(1)/SD(I)
      J=(I*I-I)/2+1
      RR=RR+A(I,1)*R(J)
  103 CONST=CONST-A(I,2)*AM(I)
      G=SQRT(RR)
      IPAGE=IPAGE+1
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1002,RR,G,CONST
      DO 104 I=2,N
      J=I-1
  104 PUNCH 1003,J,A(I,1),A(I,2)
      RR=EXIT(1.0)
 1001 FORMAT (19A4,I3,1H-,/)
 1002 FORMAT (18HGOODNESS OF FIT = ,F10.6,/,23HMULTIPLE CORRELATION = ,
     1F10.6,//,23X,E18.8)
 1003 FORMAT (I5,2E18.8)
      END TRACE
      END
```

```
C     CORRELATION
C     MULTIPLE REGRESSION
C     FACTOR ANALYSIS
C     PROGRAM 4
C     FACTOR ANALYSIS -- EIGENVALUES AND EIGENVECTORS
C     04 / 27 / 1967
C     10788 - 11222 - 12620
C     ROBERT SAMPSON -- PROGRAMMER
      BEGIN TRACE
      COMMON A(210),T(19),N,IPAGE,CONST,R(400)
      AN=N
      ANORM=0.0
```

```
      IA=0
      DO 100 I=1,N
      DO 101 J=1,N
      IA=IA+1
      R(IA)=0.0
      IF (I-J) 101,1,101
    1 R(IA)=1.0
  101 CONTINUE
      DO 103 J=I,N
      IF (I-J) 2,103,2
    2 JA=I+(J*J-J)/2
      ANORM=ANORM+A(JA)**2
  103 CONTINUE
  100 CONTINUE
      IF (ANORM) 3,3,4
    4 ANORM=SQRT(2.0*ANORM)
      ANRMX=ANORM*1.0E-06/AN
      IND=0
      THR=ANORM
    5 THR=THR/AN
    6 L=1
    7 M=L+1
    8 MQ=(M*M-M)/2
      LQ=(L*L-L)/2
      LM=L+MQ
      IF (ABS(A(LM))-THR) 10,11,11
   11 IND=1
      LL=L+LQ
      MM=M+MQ
      X=(A(LL)-A(MM))/2.0
      Y=-A(LM)/SQRT(A(LM)**2+X*X)
      IF (X) 13,14,14
   13 Y=-Y
   14 SINX=Y/SQRT(2.0*(1.0+SQRT(1.0-Y*Y)))
      SINX2=SINX*SINX
      COSX=SQRT(1.0-SINX2)
      COSX2=COSX*COSX
      SINCS=SINX*COSX
      ILQ=N*(L-1)
      IMQ=N*(M-1)
      DO 102 I=1,N
      IQ=(I*I-I)/2
      IF (I-L) 16,17,16
   16 IF (I-M) 18,17,20
   18 IM=I+MQ
      GO TO 19
   20 IM=M+IQ
   19 IF (I-L) 21,22,22
   21 IL=I+LQ
      GO TO 23
   22 IL=L+IQ
   23 X=A(IL)*COSX-A(IM)*SINX
      A(IM)=A(IL)*SINX+A(IM)*COSX
      A(IL)=X
   17 ILR=ILQ+I
      IMR=IMQ+I
      X=R(ILR)*COSX-R(IMR)*SINX
      R(IMR)=R(ILR)*SINX+R(IMR)*COSX
      R(ILR)=X
  102 CONTINUE
```

```
      X=2.0*A(LM)*SINCS
      Y=A(LL)*COSX2+A(MM)*SINX2-X
      X=A(LL)*SINX2+A(MM)*COSX2+X
      A(LM)=(A(LL)-A(MM))*SINCS+A(LM)*(COSX2-SINX2)
      A(LL)=Y
      A(MM)=X
   10 IF (M-N) 25,26,25
   25 M=M+1
      GO TO 8
   26 IF (L-N+1) 27,28,27
   27 L=L+1
      GO TO 7
   28 IF (IND-1) 29,30,29
   30 IND=0
      GO TO 6
   29 IF (THR-ANRMX) 3,3,5
    3 X=EXIT(1.0)
      END TRACE
      END
```

```
C     CORRELATION
C     MULTIPLE REGRESSION
C     FACTOR ANALYSIS
C     PROGRAM 5
C     FACTOR ANALYSIS -- EIGENVALUES AND EIGENVECTORS CONTINUED AND
C     CUMULATIVE PERCENTAGE OF EIGENVALUES
C     04 / 27 / 1967
C     09190 - 09244 - 12990
C     ROBERT SAMPSON -- PROGRAMMER
      BEGIN TRACE
      COMMON A(210),T(19),N,IPAGE,CONST,R(400),K
      DIMENSION D(20)
      AN=N
      DO 100 I=1,N
      IQ=N*(I-1)
      LL=(I*I+I)/2
      JQ=N*(I-2)
      DO 101 J=I,N
      JQ=JQ+N
      MM=(J*J+J)/2
      IF (A(LL)-A(MM)) 1,101,101
    1 X=A(LL)
      A(LL)=A(MM)
      A(MM)=X
      DO 102 K=1,N
      ILR=IQ+K
      IMR=JQ+K
      X=R(ILR)
      R(ILR)=R(IMR)
  102 R(IMR)=X
  101 CONTINUE
  100 CONTINUE
    4 IPAGE=IPAGE+1
```

```
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1002
      CP=0.0
      DO 104 I=1,N
      J=(I*I+I)/2
      CP=CP+A(J)*100.0/AN
      PRINT 1003,I,A(J),CP
  104 PUNCH 1003,I,A(J),CP
      IG=1
   16 IQ=0
      IPAGE=IPAGE+1
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1006
      DO 105 I=1,N
      DO 106 J=1,20
  106 D(J)=0.
      DO 107 L=1,N
      IQ=IQ+1
  107 D(L)=R(IQ)
      GO TO (11,12),IG
   11 PUNCH 1005,I,D(1),D(2),D(3),D(4),D(5),D(6),D(7),D(8),D(9),D(10)
      GO TO 105
   12 PUNCH 1005,I,D(11),D(12),D(13),D(14),D(15),D(16),D(17),D(18),
     1D(19),D(20)
  105 CONTINUE
      IF (N-11) 13,14,14
   14 GO TO (15,13),IG
   15 IG=2
      GO TO 16
   13 L=0
      JJ=0
      PRINT 1010
      ACCEPT 1003,K
      DO 110 J=1,K
      JJ=JJ+J
      SQ=SQRT(A(JJ))
      DO 111 I=1,N
      L=L+1
  111 R(L)=SQ*R(L)
  110 CONTINUE
      IPAGE=IPAGE+1
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1007,K
      DO 112 I=1,N
      DO 113 J=1,20
  113 D(J)=0.0
      DO 114 J=1,K
      L=N*(J-1)+I
  114 D(J)=R(L)
      PUNCH 1008,I,D(1),D(2),D(3),D(4),D(5),D(6),D(7),D(8),D(9),D(10),
     1D(11),D(12),D(13),D(14)
  112 CONTINUE
      IF (K-1) 20,20,21
   20 PRINT 1009,K
      STOP 100
   21 C=EXIT(1.0)
 1001 FORMAT (19A4,I3,1H-)
```

```
1002 FORMAT (/,9X,10HEIGENVALUE,6X,7HPERCENT,/)
1003 FORMAT (I4,2F15.7)
1005 FORMAT (/,6HVECTOR,I4,9X,6F10.6,/,4F10.6,39X,1H2)
1006 FORMAT (/,12HEIGENVECTORS,/)
1007 FORMAT (/,15HFACTOR MATRIX (,I4,9H FACTORS),/)
1008 FORMAT (/,I4,3X,9F8.4,/,5F8.4,39X,1H2)
1009 FORMAT (4HONLY,I4,33H FACTORS, NO ROTATION CAN BE MADE)
1010 FORMAT (27HNUMBER OF FACTORS TO ROTATE)
     END TRACE
     END
```

```
C     CORRELATION
C     MULTIPLE REGRESSION
C     FACTOR ANALYSIS
C     PROGRAM 6
C     FACTOR ANALYSIS -- VARIMAX ROTATION
C     04 / 28 / 1967
C     ROBERT SAMPSON -- PROGRAMMER
      BEGIN TRACE
      COMMON TV(210),T(19),N,IPAGE,CONST,A(400),K,H(20),NC
      DO 106 I=1,N
      H(I)=0.0
      DO 107 J=1,K
      L=N*(J-1)+I
  107 H(I)=H(I)+A(L)**2
      HI=SQRT(H(I))
      DO 108 J=1,K
      L=N*(J-1)+I
  108 A(L)=A(L)/HI
  106 CONTINUE
      EPS=.00116
      TVLT=0.0
      LL=K-1
      NV=1
      NC=0
      FN=N
      FFN=FN*FN
      CONS=.7071066
    1 TV(NV)=0.0
      DO 100 J=1,K
      AA=0.0
      BB=0.0
      LB=N*(J-1)
      DO 101 I=1,N
      L=LB+I
      CC=A(L)*A(L)
      AA=AA+CC
  101 BB=BB+CC*CC
  100 TV(NV)=TV(NV)+(FN*BB-AA*AA)/FFN
      IF (SENSE SWITCH 1) 30,31
   30 PRINT 1000,NV,TV(NV)
 1000 FORMAT (I4,F15.8)
   31 IF (NV-51) 2,3,3
```

```
    2 IF (TV(NV)-TVLT-1.0E-07) 4,4,5
    4 NC=NC+1
      IF (NC-3) 5,5,3
    5 DO 102 J=1,LL
      L1=N*(J-1)
      II=J+1
      DO 103 KI=II,K
      L2=N*(KI-1)
      AA=0.0
      BB=0.0
      CC=0.0
      DD=0.0
      DO 104 I=1,N
      L3=L1+I
      L4=L2+I
      U=(A(L3)+A(L4))*(A(L3)-A(L4))
      TN=A(L3)*A(L4)*2.0
      CC=CC+(U+TN)*(U-TN)
      DD=DD+2.0*U*TN
      AA=AA+U
  104 BB=BB+TN
      TN=DD-2.0*AA*BB/FN
      B=CC-(AA*AA-BB*BB)/FN
      IF (TN-B) 7,8,9
    8 IF (TN+B-EPS) 103,10,10
   10 COS4T=CONS
      SIN4T=CONS
      GO TO 11
    7 TAN4T=ABS(TN/B)
      IF (TAN4T-EPS) 12,13,13
   13 COS4T=1.0/SQRT(1.0+TAN4T*TAN4T)
      SIN4T=TAN4T*COS4T
      GO TO 11
   12 IF (B) 14,103,103
   14 SINP=CONS
      COSP=CONS
      GO TO 15
    9 CTN4T=ABS(TN/B)
      IF (CTN4T-EPS) 16,17,17
   17 SIN4T=1.0/SQRT(1.0+CTN4T*CTN4T)
      COS4T=CTN4T*SIN4T
      GO TO 11
   16 COS4T=0.0
      SIN4T=1.0
   11 COS2T=SQRT((1.0+COS4T)/2.0)
      SIN2T=SIN4T/(2.0*COS2T)
      COST=SQRT((1.0+COS2T)/2.0)
      SINT=SIN2T/(2.0*COST)
      IF (B) 19,19,20
   20 COSP=COST
      SINP=SINT
      GO TO 23
   19 COSP=CONS*(COST+SINT)
      SINP=ABS(CONS*(COST-SINT))
   23 IF (TN) 24,24,15
   24 SINP=-SINP
   15 DO 109 I=1,N
      L3=L1+I
      L4=L2+I
      AA=A(L3)*COSP+A(L4)*SINP
      A(L4)=-A(L3)*SINP+A(L4)*COSP
```

```
  109 A(L3)=AA
  103 CONTINUE
  102 CONTINUE
      NV=NV+1
      TVLT=TV(NV-1)
      GO TO 1
    3 NC=NV-1
      C=EXIT(1.0)
      END TRACE
      END
```

```
C     CORRELATION
C     MULTIPLE REGRESSION
C     FACTOR ANALYSIS
C     PROGRAM 7
      COMMON TV(210),T(19),N,IPAGE,C1,A(400),K,H(20),NC
      DIMENSION D(20),F(20),E(20)
      DO 100 I=1,N
      DO 101 J=1,K
      L=N*(J-1)+I
  101 A(L)=A(L)*SQRT(H(I))
  100 CONTINUE
      DO 103 I=1,N
      F(I)=0.0
      DO 104 J=1,K
      L=N*(J-1)+I
  104 F(I)=F(I)+A(L)*A(L)
  103 D(I)=H(I)-F(I)
      IPAGE=IPAGE+1
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1002
      NV=NC+1
      DO 105 I=1,NV
      J=I-1
  105 PUNCH 1003,J,TV(I)
      IPAGE=IPAGE+1
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1005,K
      DO 106 I=1,N
      DO 108 J=1,20
  108 E(J)=0.0
      DO 107 J=1,K
      L=N*(J-1)+I
  107 E(J)=A(L)
      PUNCH 1004,I,E(1),E(2),E(3),E(4),E(5),E(6),E(7),E(8),E(9)
      IF (K-10) 106,11,11
   11 PUNCH 1009,E(10),E(11),E(12),E(13),E(14)
  106 CONTINUE
      IPAGE=IPAGE+1
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1006
```

```
      DO 109 I=1,N
  109 PUNCH 1007,I,H(I),F(I),D(I)
      DO 110 I=1,K
      BIG=0.0
      JS=0
      DO 111 J=1,N
      L=N*(I-1)+J
      IF (BIG-ABS(A(L))) 20,111,111
   20 BIG=ABS(A(L))
      JS=J
  111 CONTINUE
      H(I)=JS
  110 CONTINUE
      M=K+1
      DO 112 I=M,N
      H(I)=0.0
      DO 114 L=1,N
      C1=L
      DO 113 J=1,I
      IF (H(J)-C1) 113,114,113
  113 CONTINUE
      H(I)=C1
      GO TO 112
  114 CONTINUE
  112 CONTINUE
      C1=EXIT(1.0)
 1001 FORMAT (19A4,I3,1H-)
 1002 FORMAT (/,9HITERATION,3X,9HVARIANCES,/)
 1003 FORMAT (I7,F16.7)
 1004 FORMAT (/,I4,2X,9F8.4)
 1005 FORMAT (/,23HROTATED FACTOR MATRIX (,I4,9H FACTORS),/)
 1006 FORMAT (/,22HCHECK ON COMMUNALITIES,/)
 1007 FORMAT (I6,3F20.7)
 1009 FORMAT (5F8.5,39X,1H2)
      END TRACE
      END
```

```
C     CORRELATION
C     MULTIPLE REGRESSION
C     FACTOR ANALYSIS
C     PROGRAM 8
      BEGIN TRACE
      COMMON V(14,15),T(19),N,IPAGE,C1,B(400),K,H(20)
      IR=0
    1 DO 100 I=1,K
      LINE=H(I)
      DO 101 J=1,K
      L=N*(J-1)+LINE
  101 V(I,J)=B(L)
  100 CONTINUE
      NP=K+1
      DO 102 I=1,K
      DO 103 J=1,K
  103 V(J,NP)=0.0
      V(I,NP)=1.0
```

```
      DIV=V(I,I)
      DO 104 J=1,NP
  104 V(I,J)=V(I,J)/DIV
      DO 105 J=1,K
      IF (I-J) 2,105,2
    2 FAC=V(J,I)
      DO 106 L=1,NP
  106 V(J,L)=V(J,L)-V(I,L)*FAC
  105 CONTINUE
      DO 107 J=1,K
  107 V(J,I)=V(J,NP)
  102 CONTINUE
      IPAGE=IPAGE+1
      IR=IR+1
      PUNCH 1001,T(1),T(2),T(3),T(4),T(5),T(6),T(7),T(8),T(9),T(10),
     1T(11),T(12),T(13),T(14),T(15),T(16),T(17),T(18),T(19),IPAGE
      PUNCH 1002,IR
      IS=1
      JS=1
      TB=0.0
      DO 108 I=1,N
      J=H(I)
      DO 109 L=1,14
  109 V(L,15)=0.0
      DO 110 M=1,K
      DO 111 MM=1,K
      NN=N*(MM-1)+J
  111 V(M,15)=V(M,15)+B(NN)*V(MM,M)
      IF (ABS(V(M,15))-TB) 110,110,4
    4 TB=ABS(V(M,15))
      IS=I
      JS=M
  110 CONTINUE
      PUNCH 1003,J,V(1,15),V(2,15),V(3,15),V(4,15),V(5,15),V(6,15),
     1V(7,15),V(8,15),V(9,15)
      IF (K-10) 108,10,10
   10 PUNCH 1004,V(10,15),V(11,15),V(12,15),V(13,15),V(14,15)
  108 CONTINUE
      IF (TB-1.00005) 5,5,6
    5 C=EXIT(1.0)
    6 AM=H(JS)
      H(JS)=H(IS)
      H(IS)=AM
      GO TO 1
 1001 FORMAT (19A4,I3,1H-)
 1002 FORMAT (/,25HOBLIQUE PROJECTION MATRIX,I4,/)
 1003 FORMAT (/,I4,3X,9F8.5)
 1004 FORMAT (5F8.5,39X,1H2)
      END TRACE
      END
```

```
1111 1 2 3 4 5 6 7 8 91011
FACTOR ANALYSIS OF ARBUCKLE BRINE DATA.
0.60381E 020.10047E 030.57380E 040.15210E 040.77000E 030.58000E 030.10425E 04
0.68236E 050.20963E 050.10760E 010.10000E 030.24800E 000.63667E 01
0.81229E 020.10775E 030.30900E 040.67209E 030.15500E 040.30267E 030.24759E 04
0.41303E 050.51472E 050.10444E 010.10000E 030.11200E 000.66111E 01
0.81245E 020.10675E 030.29789E 040.78714E 030.16500E 040.33667E 030.35208E 04
0.39018E 050.95806E 050.10484E 010.10000E 030.83500E-010.66167E 01
0.79251E 020.10672E 030.27200E 040.80500E 030.14300E 040.29350E 030.19033E 04
0.28833E 050.50827E 050.10343E 010.10000E 030.12100E 000.67500E 01
0.78255E 020.10670E 030.21876E 040.87057E 030.14400E 040.36200E 030.28425E 04
0.35400E 050.66314E 050.10410E 010.10000E 030.11700E 000.70375E 01
0.80490E 020.90734E 020.33900E 040.10558E 040.10100E 040.48350E 030.21600E 04
0.32326E 050.58141E 050.10415E 010.10000E 030.16200E 000.63000E 01
0.88138E 020.11186E 030.30100E 040.10223E 040.17500E 040.37800E 030.21350E 04
0.38925E 050.61162E 050.10387E 010.10000E 030.10700E 000.73333E 01
0.84314E 020.10180E 030.31188E 040.92775E 030.17300E 040.36950E 030.11175E 04
0.43413E 050.64642E 050.10445E 010.10000E 030.10050E 000.76000E 01
0.85392E 020.96812E 020.31390E 040.63750E 030.25400E 040.44500E 030.74650E 03
0.38800E 050.70645E 050.10480E 010.10000E 030.89000E-010.64000E 01
0.14417E 020.23531E 020.20750E 040.57517E 030.12360E 040.32550E 030.87540E 03
0.24600E 050.42237E 050.10298E 010.10000E 030.13750E 000.75000E 01
0.91129E 020.14364E 030.23710E 040.10595E 040.16300E 040.47950E 030.20550E 04
0.34900E 050.65670E 050.10445E 010.10000E 030.10700E 000.67000E 01
0.91191E 020.13964E 030.27325E 030.61000E 030.83600E 030.14320E 040.29340E 04
0.18400E 050.38225E 050.10245E 010.10000E 030.25200E 000.70000E 01
0.92378E 020.12766E 030.30550E 040.79167E 030.16800E 040.41700E 030.20067E 04
0.30433E 050.58963E 050.10393E 010.10000E 030.11400E 000.65000E 01
0.99021E 020.14777E 030.22803E 040.69150E 030.13300E 040.22067E 030.26613E 04
0.24325E 050.45821E 050.10307E 010.10000E 030.12567E 000.77667E 01
0.98215E 020.13675E 030.31119E 040.86641E 030.15700E 040.41971E 030.24006E 04
0.30824E 050.51958E 050.98294E 000.10000E 030.11033E 000.65167E 01
0.10298E 030.14884E 030.34863E 040.10723E 040.19243E 040.48850E 030.22650E 04
0.36163E 050.62264E 050.10431E 010.10000E 030.10100E 000.75000E 01
0.10600E 030.14690E 030.24550E 040.58000E 030.15300E 040.27000E 030.21175E 04
0.28550E 050.50150E 050.10335E 010.10000E 030.11750E 000.69000E 01
0.10307E 030.14384E 030.31200E 040.93100E 030.16700E 040.51500E 030.22150E 04
0.30550E 050.57903E 050.10390E 010.10000E 030.98000E-010.77000E 01
0.10340E 030.12584E 030.25488E 040.72675E 030.15300E 040.32100E 030.22069E 04
0.28538E 050.48045E 050.10312E 010.10000E 030.11825E 000.66000E 01
0.10443E 030.12386E 030.24130E 040.70980E 030.13700E 040.65000E 030.16848E 04
0.26347E 050.46540E 050.10310E 010.10000E 030.12100E 000.67750E 01
0.10445E 030.12286E 030.26000E 040.65100E 030.13550E 040.35700E 030.12933E 04
0.28400E 050.44844E 050.10310E 010.10000E 030.12900E 000.69000E 01
0.11326E 030.13203E 030.24167E 040.56767E 030.14250E 040.18400E 030.19400E 04
0.28367E 050.45798E 050.10323E 010.10000E 030.12850E 000.69000E 01
0.11524E 030.13307E 030.22967E 040.57767E 030.15500E 040.36700E 030.22667E 04
0.25533E 050.46978E 050.10327E 010.10000E 030.13000E 000.70000E 01
0.44350E 020.10832E 030.12774E 040.38867E 030.12000E 040.35717E 030.14572E 04
0.16026E 050.33226E 050.10202E 010.10000E 030.22633E 000.69650E 01
0.49314E 020.11036E 030.37778E 040.82033E 030.11900E 040.33650E 030.10796E 04
0.29073E 050.52964E 050.10369E 010.10000E 030.89000E-010.70625E 01
0.52321E 020.10839E 030.12993E 040.28925E 030.79550E 030.55000E 030.19200E 04
0.19825E 050.23419E 050.10235E 010.10000E 030.21200E 000.69500E 01
0.57207E 020.11744E 030.10506E 040.37450E 030.68300E 030.49813E 030.15989E 04
0.18819E 050.39682E 050.10211E 010.10000E 030.18786E 000.71671E 01
0.59229E 020.11446E 030.24734E 040.73473E 030.61900E 030.45471E 030.22045E 04
0.40069E 050.22416E 050.10415E 010.10000E 030.19533E 000.69950E 01
0.55248E 020.11442E 030.10293E 040.42340E 030.68500E 030.41600E 030.18264E 04
0.12239E 050.24269E 050.10162E 010.10000E 030.23600E 000.73840E 01
0.55259E 020.11342E 030.11476E 040.51380E 030.83900E 030.33667E 030.22300E 04
```

```
0.16860E 050.28767E 050.10204E 010.10000E 030.19300E 000.77500E 01
0.58299E 020.10845E 030.19041E 040.55310E 030.89600E 030.36925E 030.18905E 04
0.19763E 050.25206E 050.10235E 010.10000E 030.22757E 000.71417E 01
0.12528E 030.15968E 030.20608E 040.60713E 030.13259E 040.32786E 030.15054E 04
0.25414E 050.41664E 050.10290E 010.10000E 030.15829E 000.70000E 01
0.12236E 030.15661E 030.22420E 040.69120E 030.11387E 040.57733E 030.21960E 04
0.23032E 050.35888E 050.10264E 010.10000E 030.14967E 000.71000E 01
0.12143E 030.15359E 030.73700E 030.21050E 030.48400E 030.30950E 030.41500E 03
0.76650E 040.14494E 050.10085E 010.10000E 030.33900E 000.72000E 01
0.15422E 020.88039E 020.34010E 040.99489E 030.11200E 040.29860E 030.23711E 04
0.36373E 050.75832E 050.10437E 010.10000E 030.13733E 000.69125E 01
0.12145E 030.18241E 030.14744E 040.53280E 030.21000E 040.25633E 030.26420E 04
0.25248E 050.51659E 050.10316E 010.10000E 030.68000E-010.73333E 01
0.89437E 020.93878E 020.23097E 040.55833E 030.67300E 030.38033E 030.15483E 04
0.34169E 050.53932E 050.10363E 010.10000E 030.24300E 000.68000E 01
0.94397E 020.95964E 020.30510E 040.88314E 030.18500E 040.35050E 030.91657E 03
0.38617E 050.62079E 050.10436E 010.10000E 030.90000E-010.76417E 01
0.90033E 020.86791E 020.21433E 040.71100E 030.14850E 040.24100E 030.20000E 03
0.28900E 050.48753E 050.10347E 010.10000E 030.11000E 000.80000E 01
0.11544E 030.18327E 030.40415E 040.66675E 030.22500E 040.35167E 030.18758E 04
0.36225E 050.66453E 050.10437E 010.90000E 020.10100E 000.69333E 01
0.69203E 020.55788E 020.22923E 040.60275E 030.19500E 040.77400E 030.20550E 04
0.27750E 050.38818E 050.10317E 010.10000E 030.17400E 000.69000E 01
0.51269E 020.83263E 020.31774E 040.80029E 030.23300E 040.29343E 030.11257E 04
0.39286E 050.72951E 050.10464E 010.10000E 030.89167E-010.75000E 01
0.43369E 020.10631E 030.12310E 040.29150E 030.68700E 030.50625E 030.15775E 04
0.13137E 050.23003E 050.10157E 010.10000E 030.23050E 000.71800E 01
0.46402E 020.10133E 030.17356E 040.38260E 030.11193E 040.42600E 030.16876E 04
0.17812E 050.36271E 050.10200E 010.10000E 030.19400E 000.72000E 01
0.50374E 020.10337E 030.12483E 040.24025E 030.58700E 030.50000E 030.14640E 04
0.12400E 050.24363E 050.10160E 010.10000E 030.22967E 000.71667E 01
0.53406E 020.99397E 020.19623E 040.52991E 030.67100E 030.56950E 030.73000E 03
0.18952E 050.23363E 050.10206E 010.10000E 030.24100E 000.70000E 01
0.52417E 020.98388E 020.12867E 040.28233E 030.61700E 030.42700E 030.15367E 04
0.14567E 050.21638E 050.10170E 010.10000E 030.25500E 000.74000E 01
0.49422E 020.98360E 020.15003E 040.38800E 030.67900E 030.68800E 030.21987E 04
0.29212E 050.63117E 050.10330E 010.10000E 030.24000E 000.66500E 01
0.52427E 020.97388E 020.13630E 040.49833E 030.65400E 030.56400E 030.16743E 04
0.22267E 050.41495E 050.10273E 010.10000E 030.25000E 000.75000E 01
0.51438E 020.96378E 020.17623E 040.49133E 030.66900E 030.45600E 030.68720E 03
0.20062E 050.39643E 050.10240E 010.10000E 030.23100E 000.69400E 01
0.52456E 020.94388E 020.13354E 040.41588E 030.63500E 030.22775E 030.27194E 04
0.23399E 050.40863E 050.10306E 010.87000E 020.26650E 000.71143E 01
0.51457E 020.94378E 020.87700E 030.32100E 030.69000E 030.45800E 030.10450E 04
0.12100E 050.22368E 050.10150E 010.10000E 030.23400E 000.75000E 01
0.52465E 020.93388E 020.64500E 030.23600E 030.62800E 030.43300E 030.10035E 04
0.84500E 040.15291E 050.10095E 010.10000E 030.26300E 000.70000E 01
0.57313E 020.10744E 030.13653E 040.49300E 030.91600E 030.37500E 030.17700E 04
0.15700E 050.31713E 050.10200E 010.10000E 030.17500E 000.72500E 01
0.59306E 020.10746E 030.16593E 040.37214E 030.82000E 030.39100E 030.24325E 04
0.14992E 050.27042E 050.10164E 010.10000E 030.18250E 000.71286E 01
0.59339E 020.10446E 030.17033E 040.41733E 030.53900E 030.26050E 030.22200E 04
0.15197E 050.21497E 050.10193E 010.10000E 030.26000E 000.66500E 01
0.55381E 020.10142E 030.39460E 040.66975E 030.70700E 030.14700E 030.18300E 04
0.39080E 050.24031E 050.10477E 010.10000E 030.10700E 000.71000E 01
0.59372E 020.10146E 030.12950E 040.36500E 030.64700E 030.17400E 030.20400E 04
0.12250E 050.22902E 050.10155E 010.10000E 030.23400E 000.75000E 01
0.54404E 020.99407E 020.69250E 030.31150E 030.61500E 030.32900E 030.50800E 03
0.11150E 050.17471E 050.10115E 010.10000E 030.25450E 000.75000E 01
0.57398E 020.99438E 020.21639E 040.10153E 040.80100E 030.42200E 030.31650E 03
0.25318E 050.30237E 050.10218E 010.10000E 030.20900E 000.61833E 01
0.58396E 020.99449E 020.32215E 030.28350E 030.66800E 030.60500E 030.50000E 02
```

0.12525E	050.20125E	050.10120E	010.10000E	030.25200E	000.69000E	01	
0.59405E	020.98460E	020.10070E	040.16733E	030.66700E	030.56200E	030.17450E	03
0.15633E	050.23173E	050.10125E	010.10000E	030.24500E	000.70000E	01	
0.54424E	020.97407E	020.86027E	030.35500E	030.67100E	030.31067E	030.18133E	04
0.12133E	050.22500E	050.10157E	010.10000E	030.24100E	000.80000E	01	
0.58429E	020.96449E	020.27265E	040.10355E	040.19900E	040.39400E	030.12350E	04
0.29750E	050.52984E	050.10375E	010.10000E	030.11200E	000.72500E	01	
0.54464E	020.93407E	020.36229E	040.11342E	040.18600E	040.33700E	030.13434E	04
0.40019E	050.45553E	050.10510E	010.10000E	030.16825E	000.67833E	01	
0.56463E	020.93428E	020.81700E	030.21267E	030.69650E	030.65300E	030.91133E	03
0.18100E	050.32128E	050.10220E	010.10000E	030.24700E	000.72000E	01	
0.99161E	020.16812E	030.14963E	040.52089E	030.21600E	040.39880E	030.20961E	04
0.19559E	050.37085E	050.10237E	010.87167E	020.24767E	000.71062E	01	
0.97298E	020.16108E	030.34252E	040.79256E	030.14300E	040.31471E	030.20022E	04
0.32083E	050.60698E	050.10393E	010.10000E	030.12200E	000.67875E	01	
0.99370E	020.15712E	030.21480E	040.48250E	030.11300E	040.30150E	030.20050E	04
0.22200E	050.42704E	050.10290E	010.10000E	030.16000E	000.66000E	01	
0.10223E	030.16418E	030.17200E	040.52267E	030.20400E	040.22600E	030.24633E	04
0.25733E	050.63664E	050.10283E	010.10000E	030.92000E-010.73000E	01		
0.11127E	030.16136E	030.29200E	040.78979E	030.18162E	040.24993E	030.20029E	04
0.33800E	050.58675E	050.10413E	010.10000E	030.10315E	000.69000E	01	
0.10931E	030.15932E	030.29050E	040.72025E	030.50200E	040.22633E	030.20618E	04
0.43050E	050.79855E	050.10490E	010.10000E	030.80500E-010.70000E	01		
0.11930E	030.15954E	030.26740E	040.79690E	030.18960E	040.30350E	030.19555E	04
0.30994E	050.55383E	050.10372E	010.10000E	030.10850E	000.73750E	01	
0.11837E	030.15652E	030.24160E	040.79000E	030.11413E	040.30967E	030.13868E	04
0.24680E	050.43436E	050.10300E	010.10000E	030.16000E	000.80000E	01	
0.11848E	030.15152E	030.27013E	040.72767E	030.16300E	040.27400E	030.25000E	04
0.30524E	050.51592E	050.10367E	010.10000E	030.11550E	000.55000E	01	
0.91297E	020.10191E	030.14541E	040.48283E	030.67700E	030.42300E	030.18155E	04
0.19300E	050.23943E	050.10238E	010.10000E	030.18025E	000.72900E	01	
0.87563E	020.13201E	030.29075E	040.66888E	030.76000E	030.55750E	030.20983E	04
0.34313E	050.40925E	050.10366E	010.85333E	020.16067E	000.68000E	01	
0.69437E	020.12535E	030.20175E	040.58663E	030.83000E	030.35200E	030.14800E	04
0.23025E	050.44365E	050.10260E	010.10000E	030.17833E	000.67571E	01	
0.70470E	020.12236E	030.16155E	040.36800E	030.92200E	030.28100E	030.17850E	04
0.17425E	050.29627E	050.10185E	010.10000E	030.18100E	000.68000E	01	
0.66472E	020.12231E	030.26900E	040.87800E	030.19300E	040.55400E	030.13080E	04
0.27950E	050.48370E	050.10320E	010.10000E	030.92000E-010.75000E	01		
0.71481E	020.12137E	030.24150E	040.58175E	030.16100E	040.52350E	030.16913E	04
0.23150E	050.41237E	050.10282E	010.10000E	030.11700E	000.69500E	01	
0.68494E	020.12033E	030.65500E	040.34375E	040.92400E	030.36200E	030.16000E	04
0.26800E	050.50939E	050.10345E	010.10000E	030.16100E	000.60000E	01	
0.74443E	020.12440E	030.19907E	040.55186E	030.10300E	040.66033E	030.18176E	04
0.20530E	050.37592E	050.10259E	010.10000E	030.15600E	000.74000E	01	
0.74456E	020.12340E	030.30800E	040.18945E	040.11000E	040.54200E	030.25000E	04
0.52350E	050.96456E	050.10630E	010.10000E	030.14500E	000.60000E	01	
0.76454E	020.12343E	030.23900E	040.71950E	030.11300E	040.57300E	030.27600E	04
0.25850E	050.47059E	050.10315E	010.10000E	030.13200E	000.75000E	01	
0.73481E	020.12139E	030.22913E	040.64029E	030.10700E	040.44475E	030.29460E	04
0.34130E	050.43943E	050.10394E	010.10000E	030.14500E	000.71667E	01	
0.72493E	020.12038E	030.21766E	040.57620E	030.97300E	030.47480E	030.22931E	04
0.21620E	050.37076E	050.10253E	010.10000E	030.15500E	000.71750E	01	
0.80316E	020.13348E	030.52650E	040.10600E	040.63900E	030.43300E	030.15000E	04
0.45245E	050.22823E	050.10570E	010.10000E	030.23400E	000.68000E	01	
0.82324E	020.13251E	030.18871E	040.48200E	030.76700E	030.65550E	030.24150E	04
0.14268E	050.24197E	050.10157E	010.10000E	030.24000E	000.64667E	01	
0.83336E	020.13152E	030.14000E	040.46300E	030.82100E	030.33600E	030.29850E	04
0.34200E	050.27379E	050.10385E	010.10000E	030.19900E	000.66000E	01	
0.80438E	020.12448E	030.30325E	040.69850E	030.22650E	040.54800E	030.25500E	04
0.33133E	050.65701E	050.10363E	010.10000E	030.10400E	000.68875E	01	
0.78479E	020.12146E	030.28145E	040.77362E	030.27430E	050.53867E	030.29336E	04

```
0.37096E 050.64178E 050.10431E 010.10000E 030.10218E 000.70170E 01
0.79478E 020.12147E 030.22748E 040.58417E 030.11550E 040.52567E 030.24571E 04
0.27108E 050.48959E 050.10313E 010.10000E 030.13112E 000.70500E 01
0.85143E 020.14455E 030.19944E 040.62092E 030.15900E 040.68833E 030.31876E 04
0.31053E 050.52371E 050.10385E 010.10000E 030.11900E 000.67500E 01
0.84391E 020.12754E 030.24225E 040.75025E 030.10287E 040.48617E 030.13804E 04
0.26701E 050.38003E 050.10310E 010.10000E 030.15675E 000.67833E 01
0.84419E 020.12554E 030.32920E 040.13046E 040.15800E 040.12600E 030.15056E 04
0.33560E 050.61317E 050.10414E 010.10000E 030.11300E 000.72500E 01
0.64258E 020.79407E 020.24664E 040.77057E 030.21600E 040.34279E 030.22729E 04
0.24064E 050.32477E 050.10280E 010.10000E 030.61333E-010.76763E 01
0.65266E 020.78419E 020.14933E 040.50133E 030.72400E 030.31400E 030.38100E 04
0.26441E 050.58873E 050.10330E 010.10000E 030.14600E 000.65500E 01
0.65291E 020.76419E 020.14800E 040.36600E 030.69000E 030.44500E 030.18060E 04
0.13200E 050.24817E 050.10165E 010.10000E 030.26100E 000.75000E 01
0.71162E 020.84497E 020.18811E 040.87531E 030.10700E 040.40260E 030.14046E 04
0.34623E 050.63217E 050.10398E 010.82800E 020.17960E 000.70250E 01
0.66199E 020.83431E 020.32420E 040.10907E 040.88100E 030.46500E 030.13767E 04
0.36108E 050.74880E 050.10413E 010.10000E 030.19300E 000.72500E 01
0.67220E 020.81444E 020.15148E 040.42939E 030.81800E 030.55088E 030.21116E 04
0.20179E 050.31762E 050.10377E 010.10000E 030.18700E 000.74961E 01
0.67233E 020.80444E 020.12250E 040.39225E 030.91800E 030.48300E 030.17525E 04
0.14900E 050.27775E 050.10183E 010.10000E 030.18100E 000.75000E 01
0.69318E 020.73471E 020.15273E 040.39367E 030.78700E 030.60600E 030.20470E 04
0.17683E 050.30755E 050.10215E 010.10000E 030.19800E 000.73750E 01
0.69371E 020.69471E 020.17335E 040.46250E 030.97300E 030.15700E 030.27450E 04
0.20700E 050.31795E 050.10235E 010.10000E 030.17700E 000.70500E 01
0.74076E 020.89538E 020.41995E 040.82025E 030.15100E 040.30500E 030.85750E 03
0.42475E 050.49034E 050.10550E 010.10000E 030.97000E-010.68250E 01
0.75099E 020.87553E 020.21575E 040.82100E 030.12767E 040.38933E 030.16155E 04
0.23100E 050.38604E 050.10272E 010.95000E 020.16125E 000.70000E 01
0.76093E 020.87567E 020.19500E 040.52475E 030.10400E 040.42500E 030.22650E 04
0.22925E 050.34189E 050.10267E 010.82667E 020.19500E 000.76000E 01
0.74133E 020.85539E 020.20325E 040.71750E 030.15695E 040.41085E 030.20633E 04
0.25864E 050.45871E 050.10255E 010.95182E 020.16509E 000.73600E 01
0.76123E 020.85568E 020.16950E 040.55300E 030.11000E 040.35000E 030.21250E 04
0.19150E 050.35823E 050.10245E 010.85000E 020.19250E 000.75000E 01
0.73166E 020.83525E 020.20233E 040.57200E 030.12120E 040.43783E 030.18000E 04
0.24167E 050.41629E 050.10283E 010.10000E 030.14460E 000.75000E 01
0.77147E 020.83583E 020.28998E 040.72233E 030.10900E 040.24025E 030.26633E 04
0.25888E 050.55420E 050.10313E 010.10000E 030.14400E 000.59375E 01
0.75171E 020.82553E 020.17800E 040.41900E 030.11650E 040.20300E 030.25050E 04
0.20575E 050.36704E 050.10222E 010.10000E 030.17000E 000.68500E 01
0.73180E 020.82525E 020.22335E 040.49500E 030.99400E 030.51400E 030.19000E 04
0.19225E 050.33489E 050.10490E 010.10000E 030.17000E 000.75000E 01
0.74190E 020.81539E 020.21650E 040.60000E 030.14200E 040.26150E 030.26500E 04
0.21200E 050.40414E 050.10280E 010.10000E 030.14450E 000.70000E 01
0.73223E 020.79525E 020.14800E 040.39650E 030.83600E 030.44800E 030.22700E 04
0.15800E 050.30268E 050.10205E 010.10000E 030.19200E 000.75000E 01
0.75316E 020.72554E 020.14433E 040.45700E 030.10500E 040.56500E 030.19600E 04
0.17700E 050.31582E 050.10213E 010.10000E 030.16000E 000.75000E 01
0.73434E 020.64527E 020.19000E 040.44900E 030.11400E 040.43400E 030.27000E 04
0.24475E 050.35919E 050.10367E 010.10000E 030.13325E 000.72000E 01
0.72435E 020.64513E 020.24600E 040.57500E 030.12200E 040.42900E 030.22650E 04
0.26750E 050.48107E 050.10330E 010.10000E 030.12500E 000.70000E 01
0.83039E 020.88674E 020.33160E 040.14090E 040.86300E 030.26600E 030.24429E 04
0.34263E 050.62316E 050.10411E 010.86667E 020.11433E 000.71208E 01
0.79184E 020.80613E 020.18750E 040.68375E 030.12300E 040.49133E 030.34275E 04
0.28718E 050.56991E 050.10391E 010.69333E 020.17400E 000.70455E 01
0.86171E 020.79724E 020.27500E 040.62333E 030.12850E 040.38850E 030.23267E 04
0.32767E 050.40969E 050.10367E 010.10000E 030.14850E 000.72000E 01
                          9.0E25
```

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 1

NUMBER OF SAMPLES EXAMINED = 122

VARIABLE NUMBER	MINIMUM VALUE	MAXIMUM VALUE	MEAN	STANDARD ERROR OF MEAN
1	273.2500000	6550.0000000	2255.6043701	90.4923534
2	167.3299999	3437.5000000	666.9003754	33.7884774
3	484.0000000	27430.0000000	1433.3737335	221.3325558
4	126.0000000	1432.0000000	415.6301994	14.3571125
5	50.0000000	3810.0000000	1897.8546906	63.5190010
6	7665.0000000	68236.0000000	26361.5737305	881.7889252
7	14494.0000000	96456.0000000	43556.0898437	1500.3942719
8	0.9829400	1.0760000	1.0310697	0.0011189
9	69.3330002	100.0000000	98.7389221	0.4067703
10	0.0613330	0.3390000	0.1646294	0.0051083
11	5.5000000	8.0000000	7.0629427	0.0386358

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 2

VARIABLE NUMBER	STANDARD DEVIATION	VARIANCE	COEFFICIENT OF VARIATION
1	999.5207214	999041.6718750	0.443128E 02
2	373.2059364	139282.6699219	0.559613E 02
3	2444.6979980	5976548.2500000	0.170556E 03
4	158.5794926	25147.4553223	0.381540E 02
5	701.5903015	492228.9492187	0.369675E 02
6	9739.6771240	4861310.0000000	0.369465E 02
7	16572.3964844	4644324.0000000	0.380484E 02
8	0.0123588	0.0001527	0.119864E 01
9	4.4929244	20.1863694	0.455031E 01
10	0.0564229	0.0031835	0.342727E 02
11	0.4267461	0.1821122	0.604204E 01

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 3

CORRELATION COEFFICIENT (R)

1	1.0000	0.7881	0.1346	-0.2140	0.0192	0.7861	0.4866	0.7055	0.0249	-0.5286
2	0.7881	1.0000	0.0768	-0.0747	0.0312	0.5916	0.4828	0.5312	-0.0282	-0.3680
3	0.1346	0.0768	1.0000	0.0216	0.1579	0.2044	0.2549	0.1781	0.0262	-0.2647
4	-0.2140	-0.0747	0.0216	1.0000	0.0122	-0.1051	-0.1287	-0.0856	0.0243	0.2732
5	0.0192	0.0312	0.1579	0.0122	1.0000	0.1365	0.2969	0.1533	-0.1856	-0.2845
6	0.7861	0.5916	0.2044	-0.1051	0.1365	1.0000	0.6524	0.8716	-0.0421	-0.5465
7	0.4866	0.4828	0.2549	-0.1287	0.2969	0.6524	1.0000	0.5897	-0.0866	-0.6898
8	0.7055	0.5312	0.1781	-0.0856	0.1533	0.8716	0.5897	1.0000	-0.0621	-0.4893
9	0.0249	-0.0282	0.0262	0.0243	-0.1856	-0.0421	-0.0866	-0.0621	1.0000	-0.0815
10	-0.5286	-0.3680	-0.2647	0.2732	-0.2845	-0.5465	-0.6898	-0.4893	-0.0815	1.0000
11	-0.3542	-0.3384	-0.0137	-0.0312	-0.1537	-0.3318	-0.2380	-0.2296	-0.0544	0.0422

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 4

CORRELATION COEFFICIENT (R)

1	-0.3542
2	-0.3384
3	-0.0137
4	-0.0312
5	-0.1537
6	-0.3318
7	-0.2380
8	-0.2296
9	-0.0544
10	0.0422
11	1.0000

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 5

GOODNESS OF FIT = 0.836420

MULTIPLE CORRELATION = 0.914560

		-0.47153551E 04
1	0.48969497E 00	0.13115018E 01
2	0.93253375E-02	0.38126869E-02
3	-0.95391306E-01	-0.60124791E 00
4	-0.73227275E-01	-0.10432325E 00
5	0.41544549E 00	0.42634511E-01
6	-0.25986951E 00	-0.15673349E-01
7	0.99104295E-01	0.80150583E 04
8	0.36535178E-02	0.81278171E 00
9	-0.24023032E 00	-0.42556332E 04
10	-0.93469129E-01	-0.21892251E 03

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 6

	EIGENVALUE	PERCENT
1	4.3333752	39.3943195
2	1.3347967	51.5288348
3	1.1569419	62.0464878
4	1.0617283	71.6985626
5	0.8906040	79.7949629
6	0.7346410	86.4735174
7	0.5183883	91.1861382
8	0.4785199	95.5363188
9	0.2772021	98.0563374
10	0.1253838	99.1961899
11	0.0884105	99.9999208

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 7

EIGENVECTORS

VECTOR 1	0.416347	0.359419	0.135812	-0.099053	0.121448	0.434795	0.380415	0.405947	-0.017733	-0.346526
VECTOR 2	-0.276296	-0.295874	0.397028	-0.043072	0.597853	-0.083660	0.245765	-0.050012	-0.317075	-0.282026
VECTOR 3	0.047022	-0.104661	0.077496	-0.619472	-0.327850	-0.050470	0.017060	-0.045686	0.361443	-0.341223
VECTOR 4	-0.063559	-0.078736	0.528828	0.456103	-0.023643	-0.004136	0.023673	-0.032829	0.693628	-0.083665
VECTOR 5	0.053793	0.071215	0.282243	0.406063	-0.440511	0.152104	-0.031989	0.229164	-0.368531	0.109816
VECTOR 6	0.093640	0.081140	0.672125	-0.398922	-0.194215	-0.069604	-0.181628	-0.147734	-0.266477	0.261272
VECTOR 7	-0.160991	-0.674688	-0.024680	-0.145177	-0.078942	0.388515	-0.080104	0.513187	0.047626	0.200113
VECTOR 8	-0.274692	-0.178490	-0.070674	0.095641	-0.514855	-0.068556	0.609541	-0.218256	-0.181625	-0.244813
VECTOR 9	-0.322729	0.301981	0.026962	-0.197121	0.094485	-0.013455	0.513831	0.141158	0.210299	0.628123
VECTOR 10	-0.405224	0.268826	0.022663	-0.025929	-0.085899	-0.497983	-0.198281	0.611682	-0.032117	-0.268065
VECTOR 11	0.602219	-0.322406	-0.000835	0.052715	0.019868	-0.607945	0.280579	0.223879	0.007101	0.176967

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 8

EIGENVECTORS

VECTOR 1	-0.188458		
VECTOR 2	0.263114		
VECTOR 3	0.487441		
VECTOR 4	-0.108342	VECTOR 8	-0.310919
VECTOR 5	0.573486	VECTOR 9	0.182340
VECTOR 6	-0.370277	VECTOR 10	-0.147248
VECTOR 7	-0.167110	VECTOR 11	0.020003

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 9

FACTOR MATRIX (4 FACTORS)

1	0.8667	-0.3192	0.0506	-0.0655
2	0.7482	-0.3418	-0.1126	-0.0811
3	0.2827	0.4587	0.0834	0.5449
4	-0.2062	-0.0498	-0.6663	0.4700
5	0.2528	0.6907	-0.3526	-0.0244
6	0.9051	-0.0967	-0.0543	-0.0043
7	0.7919	0.2839	0.0184	0.0244
8	0.8451	-0.0578	-0.0491	-0.0338
9	-0.0369	-0.3663	0.3888	0.7147
10	-0.7214	-0.3258	-0.3670	-0.0862
11	-0.3923	0.3040	0.5243	-0.1116

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 10

ITERATION	VARIANCES
0	0.2836372
1	0.3825061
2	0.3960123
3	0.4096015
4	0.4203899
5	0.4210678
6	0.4210906
7	0.4210910
8	0.4210910
9	0.4210910
10	0.4210910
11	0.4210910

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 11

ROTATED FACTOR MATRIX (4 FACTORS)

1	0.9175	0.0233	0.0994	0.0872
2	0.8303	-0.0611	-0.0491	0.0191
3	0.0498	0.7553	-0.0722	0.1264
4	-0.1639	0.0852	-0.8214	0.0315
5	0.0388	0.5409	-0.1412	-0.5933
6	0.8823	0.2236	0.0249	-0.0501
7	0.6404	0.4919	0.1326	-0.1974
8	0.8157	0.2159	0.0445	-0.0839
9	-0.0167	0.1703	-0.0705	0.8736
10	-0.5240	-0.5751	-0.4042	0.0027
11	-0.5025	0.0903	0.5205	0.0459

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 12

CHECK ON COMMUNALITES

1	0.8599128	0.8599125	0.0000003
2	0.6958980	0.6958977	0.0000002
3	0.5942043	0.5942037	0.0000006
4	0.7098357	0.7098356	0.0000002
5	0.6659581	0.6659576	0.0000005
6	0.8315182	0.8315178	0.0000004
7	0.7086591	0.7086586	0.0000004
8	0.7210081	0.7210078	0.0000003
9	0.7975201	0.7975194	0.0000007
10	0.7686601	0.7686601	0.0000000
11	0.5336627	0.5336625	0.0000002

FACTOR ANALYSIS OF ARBUCKLE BRINE DATA. PAGE 13

OBLIQUE PROJECTION MATRIX 1

1	1.0000	-0.0000	-0.0000	0.0000
3	-0.0000	1.0000	-0.0000	-0.0000
4	-0.0000	0.0000	1.0000	0.0000
9	-0.0000	-0.0000	-0.0000	1.0000
2	0.9440	-0.1174	0.1896	-0.0622
5	0.0092	0.8806	0.1655	-0.8134
6	0.9550	0.3030	0.0756	-0.1992
7	0.6303	0.7330	-0.1160	-0.3907
8	0.8762	0.3055	0.0447	-0.2293
10	-0.4335	-0.8362	0.5004	0.1497
11	-0.6892	0.1978	-0.7447	0.1196

KANSAS GEOLOGICAL SURVEY COMPUTER PROGRAM
THE UNIVERSITY OF KANSAS, LAWRENCE

PROGRAM ABSTRACT

Title (If subroutine state in title):

R-mode factor analysis program in FORTRAN II for the IBM 1620 computer

Date:

Author, organization: Robert J. Sampson, Idaho State University, Pocatello, Idaho

Direct inquiries to: Author, or

Name: John C. Davis

Address: Kansas Geological Survey
Lawrence, Kansas 66044

Purpose/description: Completes R-mode factor analysis for 20 variables and performs verimax and oblique projections.

Mathematical method: Principal components solution.

Restrictions, range: Accepts samples having up to 30 variables. Up to 20 may be factored, and up to 14 may be rotated. Any number of samples may be entered.

Computer manufacturer: IBM

Model: 1620 Model 1

Programming language: PDQ FORTRAN

Memory required: 20 K

Approximate running time: 3 1/2 hours for 20 variables

Special peripheral equipment required: none

Remarks (special compilers or operating systems, required word lengths, number of successful runs, other machine versions, additional information useful for operation or modification of program)

Requires PDQ FORTRAN Processor C2 without reread (IBM User's Group Program 2.0.031). FORTRAN IV versions available for IBM 1130 and GE 625 computers.

Q-MODE FACTOR ANALYSIS PROGRAM IN FORTRAN IV FOR SMALL COMPUTERS

by

J. E. Klovan
University of Calgary

INTRODUCTION

Q-mode factor analysis is a multivariate quantitative technique that may be used to study and portray the interrelationships between geologic items such as rock samples, fossil collections, etc. The technique has been applied recently to geologic problems by Harbaugh and Demirmen (1964), Imbrie and Van Andel (1964), Manson and Imbrie (1964), and Klovan (1966), to name only a few. These authors have presented detailed descriptions of the method and its application.

Although programs for performing Q-mode factor analysis have been published previously (Imbrie, 1963; Manson and Imbrie, 1964), these programs are designed for large computers and require considerable effort to convert to small machines. The program described here is being used currently on an IBM System/360 Model 30 but is adapted easily to other computers. It has been written in a low-level subset of FORTRAN IV which allows its use on practically all IBM System/360 machines, and with modification of input-output statements, on the IBM 1620 as well.

PROGRAM DESCRIPTION

Size Restrictions and Data Format

A maximum of 60 items and 60 properties can be analyzed at one time (expansion of these size restrictions for larger machines is discussed later). The data are read in with items as rows of the data matrix and variables or properties as columns. Statement 52 of the main program controls the input format of data. The first 12 columns of a data card are reserved for identification purposes. Format statement 52 must always begin with 3A4.

A maximum of ten factors can be extracted by this program.

Machine Requirements

The version of the program listed here requires 64K memory with two scratch tapes or equivalent working area on disk. It has been run under the standard IBM DOS system and should be compatible with BOS and OS systems.

Operating Procedure

Two control cards precede the data deck:

Card 1. Columns 1-60. Any alphanumeric information.
Card 2. Columns 1- 2. NV = number of items (rows) in data matrix (maximum of 60).
Columns 3- 6. NS = number of variables (columns) in data matrix (maximum of 60).
Columns 7-11. QUIT = stopping criterion (usually set at 99.0).

The data deck follows the control cards. Because geological data seldom are obtained in a specific format, it may be most convenient to leave the main program in source form, change statement 52 to comply with the format of the data, and recompile with each run. If data need to be transformed by, for example, conversion to logarithms, additional FORTRAN statements may be added to the main program at the appropriate place.

Computational Methods

The main program reads in control cards and data. Means of the row vectors of data matrix X are computed as are square roots of the sums of squares, which are the lengths of the row vectors. A matrix of cosθ is next computed according to the equation

$$XX' = S,$$

or in scalar form

$$S_{kl} = \frac{\sum_{i=1}^{n} X_{ki} X_{li}}{\sqrt{\sum_{i=1}^{n} X_{ki}^2 \sum_{i=1}^{n} X_{li}^2}} ,$$

where n equals the number of variables, and where S_{kl} is the cosine of the angle between row vectors k and 1. The results of these computations are listed with appropriate labels. Row vectors then are normalized by dividing each element of each row by the row vector length and are stored on tape for computation of factor scores.

Subroutine PRINCP

The eigenvalues and eigenvectors of the cosθ matrix are extracted sequentially using a modified Hotelling iterative procedure. After each eigenvalue is extracted, the cumulative sum of squares accounted for is compared with the stopping criterion QUIT. Extraction of eigenvalues is terminated when QUIT is achieved, when 10 eigenvalues have been extracted, or when any eigenvalue is less than 0.01. The principal component factor matrix is listed and punched.

Subroutine VARMAX

The principal component factor matrix is rotated according to the standard varimax procedure for all principal factors extracted. Then, the last factor is deleted and remaining ones rotated. This process is continued until only two factors remain.

In the case of two and three factors, the varimax matrix is converted to factor components by squaring each element. This matrix is row normalized, resulting in normalized factor components which can be plotted directly on 3-component diagrams.

Subroutine FACSCO

This subroutine computes factor scores for each of the varimax matrices, according to the equation

$$F = Z A \Lambda^{-1} T ,$$

where F is the varimax factor score matrix, Z is the row normalized data matrix, A is the principal component factor matrix, Λ is the diagonal matrix of eigenvalues, and T is the varimax transformation matrix.

The column vectors of F show the relative amount of each variable in each factor. These numbers are pseudostandardized and should be used in a qualitative way only. A value of zero indicates that the factor contains roughly an average amount of that variable, a +1.0 means that the factor contains roughly one standard deviation above the average of that variable.

CONVERSION TO OTHER MACHINES

If a large computer is available, DIMENSION statements in the main program and each subroutine may be changed. All array and matrix sizes are determined by the maximum number of items and the maximum number of factors.

A version of this program, adapted for use on an IBM System/360 Model 40 with a 256K memory, can treat 200 samples at once. It should be emphasized that matrices larger than 90 x 90 cannot be handled under DOS and OS must be used. A version modified for the IBM 1620 can treat 40 samples at a time. This modification entails changing of all input-output statements to 1620-compatible format as well as changing the DIMENSION statements.

REFERENCES

Harbaugh, J. W., and Demirmen, F., 1964, Application of factor analysis to petrologic variations of Americus Limestone (Lower Permian), Kansas and Oklahoma: Kansas Geol. Survey Sp. Dist. Publ. 15, 40 p.

Imbrie, J., 1963, Factor and vector analysis programs for analyzing geologic data: Office of Naval Research, Geography Branch, ONR Task No. 389-135, Tech. Rpt. No. 6, 83 p.

Imbrie, J., and Van Andel, T. H., 1964, Vector analysis of heavy mineral data: Geol. Soc. America Bull., v. 75, no. 11, p. 1131-1155.

Klovan, J. E., 1966, The use of factor analysis in determining depositional environments from grain-size distributions: Jour. Sed. Pet., v. 36, no. 1, p. 115-125.

Manson, V., and Imbrie, J., 1964, FORTRAN program for factor and vector analysis of geologic data using an IBM 7090 or 7094/1401 computer system: Kansas Geol. Survey Sp. Dist. Publ. 13, 46 p.

```
C    Q - MODE FACTOR ANALYSIS
C  COS THETA PROGRAM
C
C    THIS PROGRAM WRITTEN FOR THE IBM 360/30
C                       BY
C              J.E. KLOVAN
C              DEPARTMENT OF GEOLOGY
C              UNIVERSITY OF CALGARY
C
C---------------------------------------------------------------------------
C
C    CONTROL CARD SET UP
C       CARD 1.
C           COLS. 1-60  JOB TITLE
C       CARD 2.
C           COLS. 1-2 (I2) NV= NUMBER OF ITEMS (ROWS) MAX OF 60.
C           COLS  3-6 (I4) NS= NUMBER OF VARIABLES (COLS) MAX OF 60
C           COLS  7-11 (F5.2) QUIT= STOPPING CRITERION
C
C---------------------------------------------------------------------------
C    NOTE.
C      DATA MATRIX HAS ITEMS AS ROWS
C
C      FIRST TWELVE COLUMNS OF DATA CARD RESERVED FOR ITEM IDENTIFICATION
C
C      STATEMENT 52 CONTROLS INPUT FORMAT
C
C
C      DIMENSIONING, READING CONTROL CARD AND INITIALIZATION
C
        DIMENSION X( 60),S( 60),XX( 60, 60),TITLE(15)
        DIMENSION NAME( 60,3)
        COMMON X,S,NV,NS,TITLE ,NAME
        REWIND 4
    1 READ(1,59) ( TITLE(I), I = 1,15)
   59 FORMAT(15A4)
      WRITE(3,60) (TITLE(I),I = 1,15)
   60 FORMAT(1H1,15A4)
      READ(1,51) NV,NS, QUIT
   51 FORMAT(I2,I4,F5.2)
      VN = NV
      SN = NS
      DO 2 I = 1, NV
      S(I) = 0.
      DO 2 J = I, NV
    2 XX(I,J) = 0.
C
C     READ RAW DATA
C
        DO 70 I = 1,NV
   70 READ(1, 52)(NAME(I,K), K = 1,3),(XX(I,J),J = 1,NS)
   52 FORMAT(3A4,10F3.1)
C
C  TRANSFORMS CAN GO HERE
C
C    PUT TRANSPOSED DATA MATRIX ON TAPE 4.
C
      DO 71 J = 1,NS
   71 WRITE(4) (XX(I,J), I = 1,NV)
      REWIND 4
C
      DO 75 I = 1,NV
      DO 75 J = I,NV
   75 XX(I,J) = 0.0
C
C     READING IN DATA AND COMPUTING SUMS AND RAW SUMS OF SQUARES AND
C     PRODUCTS
C
      DO 3 K = 1, NS
      READ(4) ( X(I), I = 1,NV)
C
C  TRANSFORMATIONS CAN GO HERE
C
      DO 3 I = 1, NV
      S(I) = S(I) + X(I)
      DO 3 J = I , NV
      XX(I,J) = XX(I,J) + X(I) * X(J)
    3 XX(J,I) = XX(I,J)
```

```
C
C
      WRITE(3,51)
      WRITE(3,53) NV,NS
   53 FORMAT(1H0,21HNUMBER OF SAMPLES    =  I3,15X,
     120HNUMBER OF VARS.    N=I5,///,37H          MEANS AND SUM OF SQUARES
     2     //,8X,9H SAMPLE  14X4HMEAN 11X8HVECTOR  /,8X,7H NAME   29X9H  L
     3ENGTH  /)
      DO 4 I = 1,NV
      X(I) = SQRT  (XX(I,I))
      S(I) = S(I) / SN
    4 WRITE(3,55)I,(NAME(I,J), J = 1,3),S(I),X(I)
   55 FORMAT(1H ,I5,2X,3A4,2F17.4)
C
C       COMPUTING COS THETA MATRIX
C
   11 DO 12 I = 1, NV
      DO 12 J = I, NV
      XX(I,J) = XX(I,J) / (X(I) * X(J))
   12 XX(J,I) = XX(I,J)
C
C  PRINTING  COS THETA MATRIX
C
   84 WRITE(3,58)
   58 FORMAT(///,'         COS THETA MATRIX  '  )
    6 DO 10 K= 1, NV, 15
      L = K + 14
    7 IF(NV - L ) 8,9,9
    8 L = L - 1
      GO TO 7
    9 WRITE(3,56)(J,J = K,L)
   56 FORMAT(1H0,7X,12HSAMPLES      ,3X,15(I5,2X),/ )
      DO 10I = 1, NV
   10 WRITE(3,57) I,(NAME(I,JK),JK=1,3),(XX(I,J),J=K,L)
   57 FORMAT(1H ,I5,2X,3A4,3X,15F7.3 )
C
      DO 72  I =1,NV
   72 X(I) = SQRT((X(I)* X(I))/ SN)
      REWIND 4
C
C     NORMALIZE DATA AND PUT ON TAPE 5
C
      REWIND 5
      DO 73  I= 1,NS
      READ(4) (S(J), J=1,NV)
      DO 74  J= 1,NV
   74 S(J) = S(J)/ X(J)
   73 WRITE(5)(S(J),J=1,NV)
      REWIND 4
      REWIND 5
      IF(QUIT)40,40,41
   41  CALL PRINCP(XX,QUIT)
   40 CALL EXIT
   50 FORMAT(1H1)
   54 FORMAT(I5,F17.4,F14.4)
       END
```

```
      SUBROUTINE PRINCP(RMAT,      QUIT)
C
C     PRINCIPAL AXES FACTOR ANALYSIS USING HOTELLING ITERATIVE PROCEDURE    4001
C
C     MAX = NO.OF SAMPLES , NS = NO. OF VARIABLES , NF = NO. OF FACTORS
C     RMAT CONTAINS COS THETA MATRIX
C     F  CONTAINS PRINCIPAL FACTOR MATRIX
C
C     DTEST IN STATEMENT 60+1 CONTROLS ACCURACY OF SOLUTION
C
       DIMENSION RMAT(60,60),V(60),VTEMP(60), F(60,10), VAR(10),CUM(10),
     1 TITLE(15), COM(60)
      DIMENSION NAME( 60,3)
      COMMON V,VTEMP,MAX,NS,TITLE ,NAME
1     FORMAT(I4,2F4.4)
    2 FORMAT(8X,8F8.3)
    4 FORMAT(1H ,6HFACTOR,I3, 9H EIGENV =,F12.8, 4H FOR,F7.2, 9H PER CEN
     1T)
      WRITE(3,60) (TITLE(I), I = 1,15)
   60 FORMAT(1H1,15A4, ///)
C
C     INITIALIZING
C
       DTEST = .005
      TRACE = 0.0                                                           4008
      DO 101  M = 1, MAX                                                    4009
101   TRACE = TRACE + RMAT(M,M)                                             4012
      NF = 0                                                                4014
      TVAR = 0.0                                                            4015
102   DO 103  M = 1, MAX                                                    4016
103   V(M) = 1.0                                                            4017
C
C     START MAIN CYCLE
C
104   BIGV = 0.0                                                            4018
      DO 107  M = 1,MAX                                                     4019
      VTEMP(M) = 0.0                                                        4020
      DO 105  N = 1,MAX                                                     4021
105   VTEMP(M) = VTEMP(M) + RMAT(M,N) * V(N)                                4022
      IF ( ABS (BIGV) - ABS (VTEMP(M))) 106,107,107
106   BIGV = VTEMP(M)                                                       4024
107   CONTINUE                                                              4025
      BIGD = 0.0                                                            4026
      DO 109  M = 1, MAX                                                    4027
      VTEMP(M) = VTEMP(M)/ BIGV                                             4028
      IF ( BIGD - ABS  ( V(M) - VTEMP(M))) 108,109,109
108   BIGD = ABS ( V(M) - VTEMP(M))
109   V(M) = VTEMP(M)                                                       4031
C
C     TEST FOR CONVERGENCE
C      IF BIGD LESS THAN DTEST DO IT AGAIN
C
      IF ( BIGD - DTEST) 110,110,104                                        4032
110   VSQ = 0.0                                                             4033
      DO 111  M = 1, MAX                                                    4034
111   VSQ = VSQ + VTEMP(M) ** 2                                             4035
      SUMSQ = 0.0                                                           4036
      NF = NF + 1                                                           4037
      BIGD = SQRT ( ABS (BIGV)/ VSQ)
      DO 112  M = 1, MAX                                                    4039
      V(M) = VTEMP(M) * BIGD                                                4040
  300 F(M,NF) = V(M)
C
C     SUMSQ IS THE EIGENVALUE FOR THIS ITERATION
C     BIGV IS THE CUMULATIVE % SUMS OF SQUARES EXPLAINED
C
112   SUMSQ = SUMSQ + V(M) ** 2                                             4042
      TVAR = TVAR + SUMSQ                                                   4043
      BIGV = TVAR * 100.0 / TRACE                                           4044
      WRITE(3,4) NF,SUMSQ,BIGV
      IF(SUMSQ - .01) 510,510,511
  510 NF = NF - 1
      GO TO 201
  511 IF(BIGV - QUIT) 200,201,201
200   DO 113  M = 1, MAX                                                    4047
      DO 113 N = 1, MAX                                                     4048
113   RMAT(M,N) = RMAT(M,N) - V(M) * V(N)                                   4049
      IF(NF - 10) 114,201,201
```

```
C
C    GO BACK AND GET NEXT EIGENVALUE
C
  114 GO TO 102
  201 WRITE(3,301)
  301 FORMAT(1H0,35H PRINCIPAL COMPONENT FACTOR MATRIX   //)
       WRITE(3,700) ( J, J = 1,NF)
  700 FORMAT(1H0,22X,'COMM.',4X,10(I5,4X),/ )
      DO 600 L = 1,MAX
  600   COM(L) = 0.0
      DO 601 L = 1,MAX
      DO 601 K = 1,NF
  601 COM(L) = COM(L) + F(L,K)*F(L,K)
          DO 305 I = 1,MAX
      WRITE(2,802) I,(NAME(I,JK),JK=1,3),(F(I,J),J=1,NF)
  802 FORMAT(I5,2X,3A4,10F6.3)
C
C     PUT  PRINC. COMP. MATRIX  ON  4
C
      WRITE(4)(F(I,J), J=1,NF)
  305 WRITE(3,311) I,(NAME(I,JK),JK=1,3),COM(I),(F(I,J),J=1,NF)
  311 FORMAT(1H ,I5,2X,3A4,11F9.4)
      DO 499 I = 1,NF
      CUM(I) = 0.
  499 VAR(I) = 0.
       FN = MAX
      DO 500 I = 1,NF
      DO 500 J = 1,MAX
  500 VAR(I) = VAR(I) + F(J,I) * F(J,I)
      DO 505 I = 1,NF
  505 VAR(I) =(VAR(I)/FN) * 100.
      CUM(1) = VAR(1)
      DO 501 I = 2,NF
  501 CUM(I) = CUM(I -1) + VAR(I)
      WRITE(3,502) (VAR(I), I = 1,NF)
  502 FORMAT(1H0,19X,8HVARIANCE,3X,10(F7.3 ,2X))
      WRITE(3,503)(CUM(I), I = 1,NF)
  503 FORMAT(1H0,19X,8HCUM. VAR,3X,10(F7.3 ,2X))
       NZ = 1
      DO 850  I=1,NF
  850 VAR(I)=(VAR(I) * FN)/100.
C
C    PUT EIGENVALUES ON TAPE 4
      WRITE(4)(VAR(I),I=1,MAX)
C
C    ROTATE NF FACTORS FIRST TIME,THEN DROP OFF THE LAST ONE UNTIL HAVE
C    ONLY TWO LEFT.
C
      ITER =0
  800 CALL VARMAX(F,NF,COM,VAR,CUM,ITER)
      ITER = ITER+1
       NF = NF - 1
        IF( NF - 2) 801,800,800
  801 RETURN
      END                                                               4051
```

```
      SUBROUTINE VARMAX(FF,MAXF,COM,VAR,CUM,ITER)
C     VARIMAX MATRIX ROTATION
C
C     MAXT = NO. OF SAMPLES, NS = NO. OF VARIABLES,MAXF = NO.OF FACTORS
C   T WILL CONTAIN THE VARIMAX TRANSFORMATION MATRIX
C
      DIMENSION F(60,10),COM(60),VAR(10),CUM(10), DUM(120),TITLE(15)
     1,FF(60,10),NAME(60,3)
      DIMENSION  T(10,10)
      COMMON DUM,MAXT,NS,TITLE   ,NAME
    1 FORMAT(2I4)
    2 FORMAT(4X 10F7.4 )
      DO 800  N = 1,MAXT
      DO 800 M = 1,MAXF
  800  F(N,M) = FF(N,M)
      DO 801 I=1,MAXF
      DO 801 J=1,MAXF
      IF(I-J) 803,802,803
  802 T(I,J)= 1.0
      GO  TO 801
  803 T(I,J) = 0.0
  801 CONTINUE
      EPS = 0.06993
  150 DO 103 N = 1, MAXT
      COM(N) = 0.0
      DO 102 M = 1, MAXF
  102 COM(N) = COM(N) + F(N,M) * F(N,M)
      COM(N) = SQRT (COM(N)   )
      DO 103 M = 1, MAXF
  103 F(N,M) = F(N,M)/COM(N)
      L = MAXF - 1
104   NORCT = 0
      DO 123 M = 1, L
      K = M + 1
      DO 123 MONE = K, MAXF
      A = 0.0
      B = 0.0
      C = 0.0
      D = 0.0
      DO 105 N = 1, MAXT
      U = F(N,M)**2  -F(N,MONE) **2
      V = F(N,M) * F(N,MONE)* 2.
      A = A + U
      B = B + V
      C = C + U ** 2 - V ** 2
105   D = D + U * V * 2.0
      R = MAXT
      QNUM = D - 2.0 * A * B / R
      QDEN = C - (A ** 2 - B ** 2) / R
      IF(ABS(QNUM) + ABS(QDEN)) 120,120,106
106   IF(ABS(QNUM) - ABS(QDEN)) 107,114,111
107   R = ABS(QNUM/QDEN)
      IF(R - EPS) 109,108,108
108   CS4TH = COS(ATAN(R))
      SN4TH = SIN(ATAN(R))
      GO TO 115
109   IF(QDEN) 110,120,120
110   SNPHI = .70710678
      CSPHI = SNPHI
      GO TO 121
111   R = ABS(QDEN/QNUM)
      IF(R - EPS) 113,112,112
112   SN4TH = 1.0/ SQRT(1.0 + R ** 2)
      CS4TH = SN4TH * R
      GO TO 115
113   CS4TH = 0.0
      SN4TH = 1.0
      GO TO 115
114   CS4TH = .70710678
      SN4TH = CS4TH
115   R = SQRT((1.0 + CS4TH) * 0.5)
      CSTH = SQRT((1.0 + R) * 0.5)
      SNTH = SN4TH/(4.0 * CSTH * R)
      IF(QDEN) 116,117,117
116   CSPHI =.70710678*(CSTH+SNTH)
      SNPHI =.70710678*(CSTH-SNTH)
      GO TO 118
117   CSPHI = CSTH
```

```
      SNPHI = SNTH                                                     1065
118   IF(QNUM) 119,121,121                                             1066
119   SNPHI = - SNPHI                                                  1067
      GO TO 121                                                        1068
120   NOROT = NOROT + 1                                                1069
      GO TO 123                                                        1070
121   DO 123 N = 1, MAXT                                               1071
      R = F(N,M) * CSPHI + F(N,MONE) * SNPHI
      F(N,MONE) = F(N,MONE) * CSPHI - F(N,M) * SNPHI
       F(N,M)= R
      IF(N-MAXF) 804,804,123
  804 TP = T(N,M)
      T(N,M)= TP * CSPHI + T(N,MONE) * SNPHI
      T(N,MONE) = -TP * SNPHI + T(N,MONE) * CSPHI
123   CONTINUE                                                         1075
      IF(NOROT -(MAXF * L)/2) 104,124,104                              1076
124   DO 125 N = 1, MAXT                                               1077
      DO 127 M = 1, MAXF
  127 F(N,M) = F(N,M) * COM(N)
  125 COM(N) = COM(N) ** 2
      WRITE(3,60) (TITLE(I), I = 1,15)
   60 FORMAT(1H1,15A4)
      WRITE(3,30)
   30 FORMAT(22HOVARIMAX FACTOR MATRIX  //)
      WRITE(3,40) (J,J = 1,MAXF)
   40 FORMAT(1H0,22X,'COMM.',4X,10(I5,4X),/ )
      DO 126 N = 1, MAXT                                               1081
  126 WRITE(3,20) N,(NAME(N,JK),JK=1,3),COM(N),(F(N,M),M=1,MAXF)
   20 FORMAT(1H ,I5,2X,3A4,11F9.4)
      DO 200 I = 1,MAXF
      VAR(I) =0.
  200 COM(I) = 0.
       FN = MAXT
      DO 201 I = 1,MAXF
      DO 201 J = 1,MAXT
  201 VAR(I) = VAR(I) + F(J,I) * F(J,I)
       DO 500 I = 1,MAXF
  500 VAR(I) = (VAR(I)/ FN) * 100.
       COM(1) = VAR(1)
       DO 202 I = 2, MAXF
  202 COM(I) = COM(I-1) + VAR(I)
      WRITE(3,502) (VAR(I), I = 1,MAXF)
  502 FORMAT(1H0,19X,8HVARIANCE,3X,10(F7.3,2X))
      WRITE(3,503)(COM(I), I = 1,MAXF)
  503 FORMAT(1H0,19X,8HCUM. VAR,3X,10(F7.3,2X))
C
C   IF ROTATING 3 OR 2 FACTORS, COMPUTE THE NORMALIZED FACTOR COMPONENTS
C
      IF(MAXF - 3) 600,600,601
  600 WRITE(3,60)(TITLE(I), I = 1,15)
        WRITE(3,602)
  602 FORMAT(1H0,'  NORMALIZED  VARIMAX FACTOR COMPONENTS     '  )
       WRITE(3,40)(J,J = 1,MAXF)
       DO 603 N = 1, MAXT
       DO 603 M = 1,MAXF
      S = F(N,M)
       F(N,M) = (F(N,M) * F(N,M))/COM(N)
      IF(S) 605,603,603
  605 F(N,M) = -F(N,M)
  603 CONTINUE
       DO 604 N = 1,MAXT
  604 WRITE(3,20) N,(NAME(N,JK),JK=1,3),COM(N),(F(N,M),M=1,MAXF)
  601 CALL FACSCO (MAXF,T,F,COM,ITER )
  100 RETURN
      END                                                              1084
```

```
      SUBROUTINE FACSCO (M,T,A,X ,ITER)
C
C    COMPUTES FACTOR SCORES
C
C     M= NO. OF FACTORS , NV = NO. OF SAMPLES , NS = NO. OF VARIABLES
C
      DIMENSION  A( 60,10),X( 60),FACTSC(10),VAR(10),T(10,10)
      DIMENSION  TITLE(15),DUM(120),NAME( 60,3)
      COMMON DUM,NV,NS,TITLE,NAME
      REWIND 4
      REWIND 5
      WRITE(3,60) TITLE
   60 FORMAT(1H1,15A4)
      WRITE(3,51)
   51 FORMAT(1H0,' VARIMAX FACTOR SCORE MATRIX ',//)
      WRITE(3,52)(J,J=1,M)
   52 FORMAT(1H0,25X,'FACTOR',10(I5,4X))
      WRITE(3,53)
   53 FORMAT(1H0, 18X, 'VARIABLE' ,//)
      IF(ITER)  6, 6,8
    6 DO 5 I=1,NV
    5 READ(4) (A(I,J),J=1,M)
      READ(4)(VAR(I),I =1,M)
    8 DO 40 LL=1,NS
      READ(5) (X(I),I=1,NV)
      DO  30  J=1,M
   25 FACTSC(J) = 0.
      DO  30  I=1,NV
   30 FACTSC(J) =(X(I)* A(I,J))+ FACTSC(J)
      DO  80  II = 1,M
   80 FACTSC(II) =FACTSC(II)/ VAR(II)
      DO 70  I=1,M
      X(I) = 0.0
      DO 70 J=1,M
   70 X(I) = X(I) + T(J,I) * FACTSC(J)
   40 WRITE (3,62) LL, (X(J), J=1,M)
   62 FORMAT(1H ,20X,I5,3X,10F9.4)
      REWIND 4
      REWIND 5
      RETURN
      END
```

SAMPLE INPUT

CONTROL CARD 1:

```
Q MODE TEST PROBLEM  AFTER IMBRIE AND VAN ANDEL
```

CONTROL CARD 2:

```
100010100.
```

DATA PUNCHED ACCORDING TO FORMAT : (3A4, 10F3.1)

```
LOC. 1       050250150050050200100050050050
LOC. 2       100300170170080080050040010000
LOC. 3       030060100130250150130080050020
LOC. 4       075275160110065140075045030025
LOC. 5       046212140066090190106056050044
LOC. 6       038136120098170170118068050032
LOC. 7       083266159142091111068046022012
LOC. 8       061227146102099154091053038029
LOC. 9       076242152138108118076050026014
LOC.10       039103112126213148119073046021
```

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

NUMBER OF SAMPLES = 10 NUMBER OF VARS. N= 10

MEANS AND SUM OF SQUARES

	SAMPLE NAME	MEAN	VECTOR LENGTH
1	LOC. 1	10.0000	38.7298
2	LOC. 2	10.0000	41.8091
3	LOC. 3	10.0000	37.7624
4	LOC. 4	10.0000	38.9808
5	LOC. 5	10.0000	36.6093
6	LOC. 6	10.0000	35.1875
7	LOC. 7	10.0000	38.8664
8	LOC. 8	10.0000	36.5871
9	LOC. 9	10.0000	37.5798
10	LOC.10	10.0000	36.0374

COS THETA MATRIX

	SAMPLES	1	2	3	4	5	6	7	8	9	10
1	LOC. 1	1.000	0.874	0.691	0.965	0.989	0.885	0.924	0.971	0.928	0.788
2	LOC. 2	0.874	1.000	0.648	0.970	0.873	0.802	0.990	0.939	0.978	0.753
3	LOC. 3	0.691	0.648	1.000	0.691	0.791	0.948	0.723	0.794	0.776	0.988
4	LOC. 4	0.965	0.970	0.691	1.000	0.960	0.870	0.990	0.986	0.986	0.795
5	LOC. 5	0.989	0.873	0.791	0.960	1.000	0.945	0.931	0.986	0.946	0.870
6	LOC. 6	0.885	0.802	0.948	0.870	0.945	1.000	0.872	0.939	0.908	0.983
7	LOC. 7	0.924	0.990	0.723	0.990	0.931	0.872	1.000	0.978	0.997	0.820
8	LOC. 8	0.971	0.939	0.794	0.986	0.986	0.939	0.978	1.000	0.987	0.879
9	LOC. 9	0.928	0.978	0.776	0.986	0.946	0.908	0.997	0.987	1.000	0.864
10	LOC.10	0.788	0.753	0.988	0.795	0.870	0.983	0.820	0.879	0.864	1.000

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

FACTOR 1 EIGENV = 9.05213642 FOR 90.52 PER CENT
FACTOR 2 EIGENV = 0.74166113 FOR 97.94 PER CENT
FACTOR 3 EIGENV = 0.20639962 FOR 100.00 PER CENT

PRINCIPAL COMPONENT FACTOR MATRIX

		COMM.	1	2	3
1	LOC. 1	1.0001	0.9506	-0.1684	0.2610
2	LOC. 2	1.0000	0.9310	-0.2859	-0.2270
3	LOC. 3	0.9999	0.8419	0.5343	-0.0752
4	LOC. 4	1.0000	0.9715	-0.2370	0.0079
5	LOC. 5	1.0001	0.9782	-0.0323	0.2053
6	LOC. 6	1.0000	0.9606	0.2699	0.0665
7	LOC. 7	1.0000	0.9723	-0.1970	-0.1263
8	LOC. 8	1.0001	0.9961	-0.0769	0.0448
9	LOC. 9	1.0000	0.9866	-0.1182	-0.1129
10	LOC.10	1.0000	0.9159	0.3966	-0.0613
		VARIANCE	90.521	7.417	2.064
		CUM. VAR	90.521	97.938	100.002

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

VARIMAX FACTOR MATRIX

		COMM.	1	2	3
1	LOC. 1	1.0001	0.6607	0.3965	0.6375
2	LOC. 2	1.0000	0.9132	0.3518	0.2055
3	LOC. 3	0.9999	0.3090	0.9333	0.1827
4	LOC. 4	1.0000	0.8180	0.3857	0.4269
5	LOC. 5	1.0001	0.6229	0.5281	0.5772
6	LOC. 6	1.0000	0.4898	0.7755	0.3983
7	LOC. 7	1.0000	0.8494	0.4346	0.2996
8	LOC. 8	1.0000	0.7265	0.5231	0.4456
9	LOC. 9	1.0000	0.8079	0.5042	0.3053
10	LOC.10	1.0000	0.4360	0.8658	0.2455
		VARIANCE	47.548	36.476	15.977
		CUM. VAR	47.548	84.024	100.002

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

NORMALIZED VARIMAX FACTOR COMPONENTS

		COMM.	1	2	3
1	LOC. 1	1.0001	0.4365	0.1572	0.4063
2	LOC. 2	1.0000	0.8340	0.1238	0.0422
3	LOC. 3	0.9999	0.0955	0.8711	0.0334
4	LOC. 4	1.0000	0.6690	0.1487	0.1822
5	LOC. 5	1.0001	0.3880	0.2789	0.3331
6	LOC. 6	1.0000	0.2399	0.6015	0.1586
7	LOC. 7	1.0000	0.7214	0.1889	0.0897
8	LOC. 8	1.0000	0.5278	0.2737	0.1985
9	LOC. 9	1.0000	0.6526	0.2542	0.0932
10	LOC.10	1.0000	0.1901	0.7497	0.0603

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

VARIMAX FACTOR SCORE MATRIX

FACTOR	1	2	3
VARIABLE			
1	0.8821	0.0346	-0.2956
2	2.4504	-0.4581	0.9455
3	1.1366	0.4271	0.4774
4	1.3324	0.9891	-1.3549
5	-0.1017	2.4238	-0.7589
6	-0.2062	0.9918	2.1590
7	-0.1788	1.0682	0.8027
8	0.0292	0.6698	0.1942
9	-0.2080	0.3984	0.6085
10	-0.2183	0.0938	0.8083

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

VARIMAX FACTOR MATRIX

		COMM.	1	2
1	LOC. 1	0.9320	0.8554	0.4474
2	LOC. 2	0.9485	0.9117	0.3424
3	LOC. 3	0.9943	0.3401	0.9373
4	LOC. 4	1.0000	0.9139	0.4059
5	LOC. 5	0.9579	0.7941	0.5720
6	LOC. 6	0.9956	0.5956	0.8005
7	LOC. 7	0.9841	0.8901	0.4381
8	LOC. 8	0.9980	0.8355	0.5477
9	LOC. 9	0.9873	0.8533	0.5092
10	LOC.10	0.9962	0.4828	0.8736
		VARIANCE	59.507	38.431
		CUM. VAR	59.507	97.938

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

NORMALIZED VARIMAX FACTOR COMPONENTS

		COMM.	1	2
1	LOC. 1	0.9320	0.7852	0.2148
2	LOC. 2	0.9485	0.8764	0.1236
3	LOC. 3	0.9943	0.1163	0.8837
4	LOC. 4	1.0000	0.8352	0.1648
5	LOC. 5	0.9579	0.6584	0.3416
6	LOC. 6	0.9956	0.3563	0.6437
7	LOC. 7	0.9841	0.8050	0.1950
8	LOC. 8	0.9980	0.6995	0.3005
9	LOC. 9	0.9873	0.7374	0.2626
10	LOC.10	0.9962	0.2340	0.7660

Q MODE TEST PROBLEM AFTER IMBRIE AND VAN ANDEL

VARIMAX FACTOR SCORE MATRIX

FACTOR VARIABLE	1	2
1	-1.0063	1.9130
2	-3.2205	6.3397
3	-2.8025	4.9963
4	-2.6010	4.3957
5	-4.0371	6.1161
6	-3.5680	5.9300
7	-2.5828	4.1243
8	-1.5430	2.4671
9	-1.0399	1.6572
10	-0.5917	0.9920

KANSAS GEOLOGICAL SURVEY COMPUTER PROGRAM
THE UNIVERSITY OF KANSAS, LAWRENCE

PROGRAM ABSTRACT

Title (If subroutine state in title):

Q-mode factor analysis program in FORTRAN IV for small computers

Date:

Author, organization: J.E. Klovan, Department of Geology, Univ. of Calgary, Calgary, Alberta

Direct inquiries to: Author, or

Name: D.F. Merriam

Address: Kansas Geological Survey
Lawrence, Kansas 66044

Purpose/description: Computes Q-mode factor analysis and performs varimax rotation.

Mathematical method: Principal components of cos θ similarity matrix.

Restrictions, range: Accepts up to 60 samples having up to 60 variables.

Computer manufacturer: IBM

Model: System/360 Model 30

Programming language: FORTRAN IV

Memory required: 64 K

Approximate running time:

Special peripheral equipment required:

Remarks (special compilers or operating systems, required word lengths, number of successful runs, other machine versions, additional information useful for operation or modification of program)

Operates under DOS, BOS, or OS systems. Versions exist for the IBM System/360 Model 40 which will handle up to 200 samples using OS, and for the IBM 1620 (40 K) which will accept up to 40 samples.

FORTRAN II PROGRAM FOR THE CALCULATION OF WILKS' Λ USING AN IBM 1620 COMPUTER

by

S.V.L.N. Rao

Indian Institute of Technology

Analysis of variance is one of the most widely used statistical techniques in geology. The purpose of most applications has been to estimate the significance of differences between means of populations. This is done by measuring the variability of the attribute under examination; if the variance between groups is greater than the variance within groups, this is considered evidence that the groups are indeed different. There are many schemes for partitioning variance into its sources and for controlling extraneous influences on the attribute being studied. Various models and geologic applications are given in Chapter 9 of Krumbein and Graybill (1965).

These techniques are based on a univariate model. That is, only one attribute or variate may be compared at a time. Most geologic problems are multivariate, however, and the appropriate model should consider sets of observation in which all variances (and covariances) are considered simultaneously. A single multivariate test is more sensitive than a series of tests based on individual variates because correlations between variables are taken into account. A sequence of univariate tests may lead to ambiguous results, as in a case discussed by Miller and Kahn (1962, p. 250). Univariate tests made of several variates measured on two groups of fossil oreodonts found significant differences with some variates but not with others. This conflict did not arise in a multivariate approach.

Wilks' Λ criterion, also known as analysis of dispersion, is a widely applicable multivariate technique. The procedure was developed by Wilks (1932) and is extensively discussed by Rao (1952). Miller and Kahn (1962, p. 254-257) present a table of computational equations which have been rewritten into a more standard terminology in Table 1. Several errors in the original table have been corrected.

Wilks' Λ criterion is calculated as the ratio of determinants of the two matrices,

$$\Lambda = \frac{|SS_W|}{|SS_T|}$$

Λ may be considered a multivariate extension of the F-test of analysis of variance. The test statistic,

$$V = -m \ln \Lambda$$

is distributed as χ^2 with $i(j-1)$ degrees of freedom. In this equation, j is the number of groups, i is the number of variables, and

$$m = (\Sigma n_j - 1) - \frac{i+j}{2}$$

The hypothesis tested is one of equality of multivariate means or

$$H_0: \mu_{i1} = \mu_{i2} = \ldots = \mu_{ij}$$

against the alternate

$$H_1: \text{Not all } \mu_{ij} \text{ are equal.}$$

PROGRAM DESCRIPTION

Earlier versions of the program followed the outline given by Miller and Kahn (1962); it has been revised subsequently into the present form. Up to 12 data groups may be used, each containing up to 10 variables and any number of observations. The number of samples need not be equal in each group. Data are not stored, but sums, squares, and cross-products are accumulated. Each card is counted as read in and the total number of observations accumulated as COUNT1. Register COUNT2 totals

Table 1.- Computational forms for sums of squares and products matrices on i observations in j groups. All possible combinations are computed except $j = j'$.

Sums of Squares	Symbol	Computational form
Between Groups	$[SS_B] =$	$(\Sigma x_{ij}\, \Sigma x_{ij}/n_j + \Sigma x_{ij'}\, \Sigma x_{ij'}/n_{j'}) - [(\Sigma x_{ij} + \Sigma x_{ij'})(\Sigma x_{ij} + \Sigma x_{ij'})/n_j + n_{j'}]$
Within Groups	$[SS_W] =$	$(\Sigma x_{ij} x_{ij} + \Sigma x_{ij'} \Sigma x_{ij'}) - (\Sigma x_{ij} \Sigma x_{ij}/n_j + \Sigma x_{ij'} \Sigma x_{ij'}/n_{j'})$
Total	$[SS_T] =$	$(\Sigma x_{ij} x_{ij} + \Sigma x_{ij'} x_{ij'}) - \dfrac{(\Sigma x_{ij} + \Sigma x_{ij'})(\Sigma x_{ij} + \Sigma x_{ij'})}{n_j + n_{j'}}$

$SS_T = SS_W + SS_B$

observations in each successive group. As soon as a group is read in, the means and variances are calculated. After all groups have entered the computer, the sums of sums are calculated and the matrices [SS_W] and [SS_T] established, the determinants are obtained using Crout's[1/] method (Nielson, 1956, p. 181). Then, Λ and V are calculated and printed, with the accompanying degrees of freedom and a set of $\chi^2_{.95}$ values (Krumbein and Graybill, 1965, p. 418).

This program is designed for use on an IBM 1620 Model 2 computer with 40K core storage, automatic divide, and indirect addressing, using a PDQ FORTRAN compiler. Simple modifications should allow the program to be run on other computers of comparable size having FORTRAN compilers.

Only one control card is required to operate the program. This card is the first entered into the computer after the program object decks and subroutines have been loaded, and contains the following information:

[1/] Crout's method may not return determinants from certain ill-conditioned matrices. In such cases, the order of variables may be changed and the program rerun, or the [SS_W] and [SS_T] matrices may be punched out and the determinants found by another procedure.

Col. 1- 5	N, an integer indicating the number of groups
Col. 6-10	NVAR, an integer indicating the number of variables.

This card is followed by data for the first group, punched in the format specified by statement 818 (as listed, this is 4F15.0). The data set is terminated by a card containing the number 999. punched in the field of the first variable. The second group follows with a similar termination card, and so on through all N groups.

The program test data listed in Appendix II was taken from an example published by Miller and Kahn (1962, p. 257). Output from this program may be compared with their Tables 12.3 and 12.4. A test statistic of 39.622 was obtained for testing the difference of multivariate means of Merychoidoden culbertsoni and Prodesmatochoerus meeki. Miller and Kahn obtained a value of 39.402. The difference probably is due to the additional significant digits carried by the computer.

ACKNOWLEDGMENTS

I wish to thank Dr. T. C. Bagchi, Head of the Department of Geology and Geophysics, Indian Institute of Technology, for his encouragement of computer studies in the earth sciences. Mr. B. S. R. Rao, Department of Civil Engineering, IIT, contributed to the writing of this program. The program was developed and tested at the IIT Computer Centre through the cooperation of Mr. Asish Maity and A. Roy.

REFERENCES

Krumbein, W. C., and Graybill, A. F., 1965, An introduction to statistical models in geology: McGraw-Hill Book Co., New York, 475 p.

Miller, R. L., and Kahn, J. S., 1962, Statistical analysis in the geological sciences: John Wiley & Sons, New York, 483 p.

Nielsen, K. L., 1956, Methods in numerical analysis: Macmillan Co., New York, 382 p.

Rao, C. R., 1952, Advanced statistical methods in biometrical research: John Wiley & Sons, New York, 390 p.

Wilks, S. S., 1932, Certain generalizations in the analysis of variance: Biometrika, v. 24, p. 471-494.

```
C         MULTIVARIATE ANALYSES
C         FORTRAN II VERSION
C         TESTED ON IBM 1620 UNIT WITH 40 CORE MEMORY
C         MAIN PROGRAM OCCUPIES 19034 STORAGE LOCATIONS
C         12 SETS OF DATA EACH WITH 1 - 10 VARIABLES
C         CAN BE PROCESSED. IN EACH SET, THERE IS NO LIMIT
C         TO THE NUMBER OF OBSERVATIONS TO BE INCLUDED.
C         FIRST CARD IN INPUT IS N - NVAR CARD IN I5
C         FORMAT. N = NO. OF SETS AND NVAR = NO. OF VARIABLES
C         DATA FOR EACH SET FOLLOWS WITH A 999. TERMINATION CARD
C         AFTER EACH SET
C
C         PROGRAMMER = S.V.L.N.RAO,ASST.PROFESSOR,
C         IND.INST.OF TECHNOLOGY,KHARAGPUR,INDIA
C
      DIMENSION X(10),P(12,10),XY(55),SXSY(55),VXY(55),VAR(12,10)
      DIMENSION    SP(10),W(10,11),A(10,11),S(10,11),DET(2)
      DIMENSION   PM(12,10)
C
C          SUMS,CROSS PRODUCT SUMS ARE COMPUTED
C
 1234 COUNT1 = 0.
      SUM = .00626454
      SMEN = .54454155
      VARR = .00654159
      READ 826,N,NVAR
      DO 47 J = 1,NVAR
   47 SP(J) = 0.
  826 FORMAT (2I5)
      AN = NVAR
      JJ = (AN*(AN+1.))/2.
      DO 4311 J = 1,JJ
      SXSY (J ) = 0.
      XY(J ) = 0.
 4311 VXY(J ) = 0.
      DO 88 K = 1,N
      NP = 1
      DO 78 J = 1,NVAR
   78 P(K,J) = 0.
      COUNT2 = 0.
   68 READ 818,(X(J),J= 1,NVAR)
  818 FORMAT(4F15.0)
      IF(X(1) - 999.) 178,198,178
  178 DO 40 J = 1,NVAR
   40 P(K,J) = P(K,J) + X(J)
 1709 MC = 1
      MB = 0
      NB = NVAR
      GO TO (408,508,608),NP
  408 DO 34 J = 1,NB
      JM = J +MB
      XY(MC)= XY(MC) +X(J) *X(JM)
      VXY(MC)= VXY(MC) + X(J) * X(JM)
   34 MC = MC+1
  708 MB = MB +1
      NB = NB-1
      GO TO (1409,1509,1609),NP
 1409 IF(NB) 408,409,408
  409 COUNT1= COUNT1 + 1.
      COUNT2= COUNT2 + 1.
      GO TO 68
```

```
C
C         MEAN, VARIANCE ARE COMPUTED
C
  198 DO 108 J = 1,NVAR
      PM(K,J) = P(K,J)/COUNT2
      TTP = P(K,J)**2 / COUNT2
  108 VAR(K,J) =(VXY(J) - TTP   )/(COUNT2 - 1.)
      NP = 2
      GO TO 1709
  508 DO 365 J = 1,NB
      JM = J+MB
      SXSY(MC) = SXSY(MC)  + PM(K,J) * PM(K,JM) * COUNT2
  365 MC= MC+1
      GO TO 708
 1509 IF(NB) 508,509,508
  509 DO 2 J = 1,NVAR
    2 SP(J) = SP(J) + P(K,J)
      DO 77 J = 1,JJ
   77 VXY( J) = 0.
   88 CONTINUE
C
C         WITHIN AND TOTAL VARIANCE MATRICES ARE COMPUTED
C
      NP = 3
      GO TO 1709
  608 DO 784 J = 1,NB
      JM = J + MB
      W(J,JM) = XY(MC) - SXSY(MC)
      W(JM,J) = W(J,JM)
      TNN =(SP(J)*SP(JM))/COUNT1
      S(J,JM) = XY(MC) - TNN
      S(JM,J) = S(J,JM)
  784 MC = MC+1
      GO TO 708
 1609 IF(NB) 608,698,608
C
C         PUNCHES OUT THE MEAN SUM AND VARIANCE FOR THE SETS
C
  698 PUNCH 1113
 1113 FORMAT(25X,30H MULTIVARIATE ANALYSES RESULTS,//)
      DO 1123 K = 1,N
      PUNCH 1115,K
 1115 FORMAT(37X,7HSET NO=I3)
      PUNCH 1117,SUM,( P(K,J),J = 1,NVAR)
      PUNCH 1117,SMEN,(PM(K,J) , J= 1,NVAR)
 1123 PUNCH 1117,VARR,(VAR(K,J),J = 1,NVAR)
      PUNCH 1141
 1141 FORMAT(/,30X,22HWITHIN VARIANCE MATRIX,/)
      DO 4412 J = 1,NVAR
 4412 PUNCH 1145,(W(I,J) ,I = 1,NVAR)
 1145 FORMAT ( 5(2X,E14.8))
      PUNCH 1151
 1151 FORMAT(/,30X,22HTOTAL  VARIANCE MATRIX,/)
      DO 4414 J = 1,NVAR
 4414 PUNCH 1145,(S(I,J) ,I = 1,NVAR)
 1117 FORMAT(10X,1A4/5(2X,E14.8) )
C
C         DETERMINANTS ARE COMPUTED BY CROUT-S METHOD
C
      DO 101 IA=1,2
      IF (IA-1) 116,116,118
```

```
  116 DO 122 I=1,NVAR
      DO 122 J=1,NVAR
  122 A(I,J)=S(I,J)
      GO TO 202
  118 DO 124 I=1,NVAR
      DO 124 J=1,NVAR
  124 A(I,J)=W(I,J)
  202 DO 252 J=1,NVAR
      DO 222 I=J,NVAR
      IF(J-1)210,210,211
  210 GS=0.
      GO TO 222
  211 M=J-1
      GS=0.
      DO 221 K=1,M
  221 GS=GS+A(I,K)*A(K,J)
  222 A(I,J)=A(I,J)-GS
      I=J
      KA=I+1
      IF(KA - NVAR) 357,357,359
  357 DO 252 JB = KA,NVAR
      IF(I-1)220,220,232
  220 GS=0.
      GO TO 252
  232 KB=I-1
      GS=0.
      DO 242 NA=1,KB
  242 GS=GS+A(I,NA)*A(NA,JB)
  252 A(I,JB) = ( A(I,JB) - GS) / A(I,I)
  359 DET(IA)=1.
      DO 101 I=1,NVAR
  101 DET(IA)=DET(IA)*A(I,I)
      VIN=DET(2)/DET(1)
C
C        TEST STATISTIC COMPUTED
C
      TK = N
      PP = NVAR
      TM = (COUNT1-1.)  -((PP+(TK-1.) + 1.)/  TK)
      IDEGF=PP * (TK - 1.)
      VV= -TM * LOGF(VIN)
      PUNCH 1161, VV ,  IDEGF
 1161 FORMAT(/,10X,11HTEST VALUE=F8.3,14X,12HDEG.FREEDOM=I3,//)
      PUNCH 832
  832 FORMAT(5X,22HX2 DISTRIBUTION VALUES,/)
      PUNCH 834
  834 FORMAT(5X,12HDF( 4)= 9.49,2X,12HDF( 6)=12.59,2X,
     112HDF( 8)=15.01,2X,12HDF(10)=18.31)
      PUNCH 837
  837 FORMAT(5X,12HDF(12)=21.03,2X,12HDF(14)=23.68,2X,
     112HDF(16)=26.30,2X,12HDF(18)=28.87)
      PUNCH 839
  839 FORMAT(5X,12HDF(20)=31.41,2X,12HDF(24)=36.42,2X,
     112HDF(30)=43.77,2X,12HDF(40)=55.76)
      PRINT 1712
 1712 FORMAT(5X,20HPLEASE PUT NEXT SET.)
      GO TO 1234
      END
```

```
                    MULTIVARIATE ANALYSES RESULTS
                                        SET NO=  1
      SUM
.66800000E+03    .12710000E+04    .24600000E+03    .92500000E+02
      MEAN
.47714285E+02    .90785714E+02    .17571428E+02    .66071428E+01
      VAR
.32967692E+01    .44892307E+01    .17637384E+01    .78686846E+00
                                        SET NO=  2
      SUM
.47200000E+03    .99900000E+03    .18800000E+03    .53000000E+02
      MEAN
.39333333E+02    .83250000E+02    .15666666E+02    .44166666E+01
      VAR
.15151545E+02    .13659090E+02    .42424272E+01    .17969727E+00

                              WITHIN VARIANCE MATRIX

.20952800E+03    .72150000E+02    .31621000E+02    .95623000E+01
.72150000E+02    .20861000E+03    .34718000E+02   -.19280000E+01
.31621000E+02    .34718000E+02    .69595900E+02    .45098000E+01
.95623000E+01   -.19280000E+01    .45098000E+01    .12206000E+02

                              TOTAL VARIANCE MATRIX

.66338500E+03    .48024000E+03    .13477000E+03    .12818470E+03
.48024000E+03    .57554000E+03    .12746200E+03    .10473100E+03
.13477000E+03    .12746200E+03    .93038500E+02    .31469300E+02
.12818470E+03    .10473100E+03    .31469300E+02    .43209620E+02

      TEST VALUE=   39.622                DEG.FREEDOM=   4

   X2* DISTRIBUTION VALUES

   DF**( 4) =  9.49  DF( 6) = 12.59  DF( 8) = 15.01  DF(10) = 18.31
   DF  (12) = 21.03  DF(14) = 23.68  DF(16) = 26.30  DF(18) = 28.87
   DF  (20) = 31.41  DF(24) = 36.42  DF(30) = 43.77  DF(40) = 55.76
```

*Symbol for X^2

**Degrees of freedom

INPUT DETAILS

	2 4				N, NVAR card
	45.0	91.0	16.0	07.5	
	46.0	93.0	17.0	06.5	
	48.0	92.0	19.0	05.0	
	46.0	91.0	19.0	06.0	
	45.0	86.0	15.0	06.5	
	51.0	93.0	19.0	07.5	
Data for Set 1	47.0	92.0	16.0	05.0	
	48.0	89.0	18.0	06.5	
	47.0	91.0	17.5	06.0	
	50.0	91.0	17.0	07.2	
	48.0	91.0	19.0	07.6	
	49.0	93.0	17.5	07.0	
	49.0	87.0	17.0	06.5	
	49.0	91.0	19.0	07.7	
	999.				Termination card
	37.0	88.0	17.0	03.9	
	43.0	84.0	19.0	04.2	
	43.0	79.0	14.0	04.0	
	42.0	80.0	17.0	05.2	
	39.0	83.0	12.0	04.5	
Data for Set 2	39.0	87.0	15.0	04.5	
	40.0	86.0	18.0	04.5	
	34.0	77.0	16.0	04.8	
	35.0	82.0	15.0	04.6	
	45.0	88.0	17.0	04.9	
	33.0	80.0	15.0	03.9	
	42.0	85.0	13.0	04.0	
	999.				Termination card

KANSAS GEOLOGICAL SURVEY COMPUTER PROGRAM
THE UNIVERSITY OF KANSAS, LAWRENCE

PROGRAM ABSTRACT

Title (If subroutine state in title):

FORTRAN II program for the calculation of Wilks' Λ using an IBM 1620 computer.

Date:

Author, organization: S.V.L.N. Rao, Dept. Geology, Indian Institute of Technology, Kharagpur, India

Direct inquiries to: Author, or

Name: D.F. Merriam

Address: Kansas Geological Survey
Lawrence, Kansas 66044

Purpose/description: Computes Wilks' criterion for testing equality of multivariate means (multivariate analysis of variance or analysis of dispersion).

Mathematical method: Λ is calculated as the ratio of the determinants of two matrices. Determinants are found by Crout's method.

Restrictions, range: Up to 12 data sets may be used, each having up to 10 variables. Any number of samples may be used in the groups.

Computer manufacturer: IBM

Model: 1620 Model 2

Programming language: PDQ FORTRAN

Memory required: 40 K

Approximate running time:

Special peripheral equipment required: None

Remarks (special compilers or operating systems, required word lengths, number of successful runs, other machine versions, additional information useful for operation or modification of program)

Requires PDQ FORTRAN Processor C2 (IBM User's Group Program 2.0.031).

COMPUTER CONTRIBUTIONS

Kansas Geological Survey
University of Kansas
Lawrence, Kansas

Computer Contribution

1. Mathematical simulation of marine sedimentation with IBM 7090/7094 computers, by J.W. Harbaugh, 1966 . . . . $1.00
2. A generalized two-dimensional regression procedure, by J.R. Dempsey, 1966 . . . . $0.50
3. FORTRAN IV and MAP program for computation and plotting of trend surfaces for degrees 1 through 6, by Mont O'Leary, R.H. Lippert, and O.T. Spitz, 1966 . . . . . . $0.75
4. FORTRAN II program for multivariate discriminant analysis using an IBM 1620 computer, by J.C. Davis and R.J. Sampson, 1966 . . . . $0.50
5. FORTRAN IV program using double Fourier series for surface fitting of irregularly spaced data, by W.R. James, 1966 . . . . $0.75
6. FORTRAN IV program for estimation of cladistic relationships using the IBM 7040, by R.L. Bartcher, 1966 . . . . $1.00
7. Computer applications in the earth sciences: Colloquium on classification procedures, edited by D.F. Merriam, 1966 . . . . $1.00
8. Prediction of the performance of a solution gas drive reservoir by Muskat's Equation, by Apolonio Baca, 1967 . . . . $1.00
9. FORTRAN IV program for mathematical simulation of marine sedimentation with IBM 7040 or 7094 computers, by J.W. Harbaugh and W.J. Wahlstedt, 1967 . . . . . . $1.00
10. Three-dimensional response surface program in FORTRAN II for the IBM 1620 computer, by R.J. Sampson and J.C. Davis, 1967 . . . . $0.75
11. FORTRAN IV program for vector trend analyses of directional data, by W.T. Fox, 1967 . $1.00
12. Computer applications in the earth sciences: Colloquium on trend analysis, edited by D.F. Merriam and N.C. Cocke, 1967. . . . . $1.00
13. FORTRAN IV computer programs for Markov chain experiments in geology, by W.C. Krumbein, 1967 . . . . $1.00
14. FORTRAN IV programs to determine surface roughness in topography for the CDC 3400 computer, by R.D. Hobson, 1967 . . . . $1.00
15. FORTRAN II program for progressive linear fit of surfaces on a quadratic base using an IBM 1620 computer, by A.J. Cole, C. Jordan, and D.F. Merriam, 1967 . . . . . $1.00
16. FORTRAN IV program for the GE 625 to compute the power spectrum of geological surfaces, by J.E. Esler and F.W. Preston, 1967. . . . . $0.75
17. FORTRAN IV program for Q-mode cluster analysis of nonquantitative data using IBM 7090/7094 computers, by G.F. Bonham-Carter, 1967 . . . . $1.00
18. Computer applications in the earth sciences: Colloquium on time-series analysis, D.F. Merriam, editor, 1967. . . . . $1.00
19. FORTRAN II time-trend package for the IBM 1620 computer, by J.C. Davis and R.J. Sampson, 1967 . . . . $1.00
20. Computer programs for multivariate analysis in geology, D.F. Merriam, editor, 1968 . . . $1.00

NOTE: Decks for the programs are available at the following prices for a limited time.

Wolleben, J.A., Pauken, R.J., and Dearien, J.A., FORTRAN IV program for multivariate paleontologic analysis using an IBM System/360 Model 40 computer $10.00

Sampson, R.J., R-mode factor analysis program in FORTRAN II for the IBM 1620 computer. . $15.00

Klovan, J.E., Q-mode factor analysis program in FORTRAN IV for small computers . . . $10.00

Rao, S.V.L.N., FORTRAN II program for the calculation of Wilks' Λ using an IBM 1620 computer . . . . $ 5.00