VARIANCE-COVARIANCE MATRIX

PURPOSE
Compute the variance-covariance matrix of a matrix.

DESCRIPTION
The variance-covariance matrix computes the covariance between each of the columns of a data matrix. That is, row i and column j of the variance-covariance matrix is the covariance between column i and column j of the original matrix. The diagonal elements (i.e., i=j) are the variances of the columns. The variance-covariance matrix is symmetric (since the variance-covariance of column i with column j is the same as the variance-covariance of column j with column i).

SYNTAX
LET <mat2> = VARIANCE-COVARIANCE MATRIX <mat1> <SUBSET/EXCEPT/FOR qualification>

where <mat1> is a data matrix;

<mat2> is a matrix where the resulting variance-covariances are saved;

and where the <SUBSET/EXCEPT/FOR qualification> is optional and rarely used in this context.

EXAMPLES
LET C = VARIANCE-COVARIANCE MATRIX A

NOTE
Covariance and correlation matrices are the building blocks for many multivariate techniques. Many of these can supported with DATAPLOT macros (in conjunction with various matrix commands). For example, the program example for CHOLESKY DECOMPOSITION demonstrates a canonical correlation analysis and the second program example for this command demonstrates Fisher’s discriminant analysis for more than 2 populations.

DEFAULT
None

SYNONYMS
None

RELATED COMMANDS

CORRELATION MATRIX	=	Compute the correlation matrix of a matrix.
PRINCIPAL COMPONENTS	=	Compute the principal components of a matrix.
MATRIX EIGENVALUES	=	Compute the matrix eigenvalues.
MATRIX EIGENVECTORS	=	Compute the matrix eigenvectors.
MATRIX INVERSE	=	Compute a matrix inverse.
MATRIX MULTIPLICATION	=	Perform a matrix multiplication.
MATRIX SOLUTION	=	Solve a system of linear equations.
SINGULAR VALUES	=	Compute the singular values of a matrix.

APPLICATIONS
Linear Algebra

IMPLEMENTATION DATE
87/10

PROGRAM 1

```
DIMENSION 20 COLUMNS
READ MATRIX X
16 16 19 21 20 23
14 17 15 22 18 22
24 23 21 24 20 23
18 17 16 15 20 19
18 11 9 18 7 14
END OF DATA
LET C = VARIANCE-COVARIANCE MATRIX X
PRINT C
```

The following output is generated.

```
         MATRIX C       --           6 ROWS
                        --           6 COLUMNS

VARIABLES--C1          C2          C3          C4          C5          C6

  0.1400E+02  0.9500E+01  0.7000E+01  0.3500E+01  0.2000E+01  0.1000E+01
  0.9500E+01  0.1820E+02  0.1725E+02  0.8750E+01  0.1875E+02  0.1280E+02
  0.7000E+01  0.1725E+02  0.2100E+02  0.8750E+01  0.2325E+02  0.1600E+02
  0.3500E+01  0.8750E+01  0.8750E+01  0.1250E+02  0.5500E+01  0.9000E+01
  0.2000E+01  0.1875E+02  0.2325E+02  0.5500E+01  0.3200E+02  0.1925E+02
  0.1000E+01  0.1280E+02  0.1600E+02  0.9000E+01  0.1925E+02  0.1470E+02
```

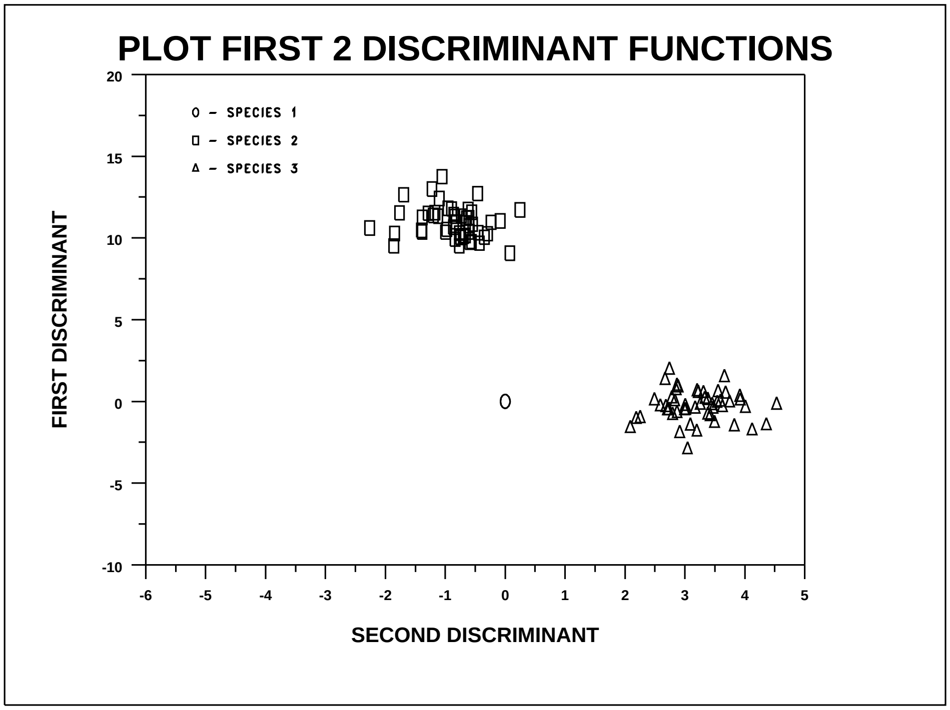
PROGRAM 2

```
. Perform a Fisher’s discriminant analysis (use Fisher’s Iris data)
FEEDBACK OFF
DIMENSION 200 COLUMNS
SKIP 25
READ IRIS.DAT SEPLENG SEPWIDTH PETLENG PETWIDTH TAG
SKIP 0
LET NTOT = SIZE SEPLENG
LET X = MATRIX DEFINITION SEPLENG NTOT 4
LET P = MATRIX NUMBER OF COLUMNS X
LET GROUPID = DISTINCT TAG
LET NG = SIZE GROUPID
LOOP FOR K = 1 1 NG
    LET N(K) = SIZE TAG SUBSET TAG = K
END OF LOOP
LOOP FOR K = 1 1 P
    LET XMGRAND(K) = MEAN X^K
END OF LOOP
. CALCULATE B0 = SUM (I=1,NG) (XBARi - XBARALL)(XBARi-XBARALL)’
LET JUNK = 0 FOR I = 1 1 P
LET B0 = DIAGONAL MATRIX JUNK
LOOP FOR K = 1 1 G
    LOOP FOR L = 1 1 P
            LET XMEANI(L) = MEAN X^L SUBSET TAG = K
    END OF LOOP
    LET TEMP = XMEANI - XMGRAND
    LET XM = MATRIX DEFINITION TEMP P 1
    LET XMT = MATRIX TRANSPOSE XM
    LET B0TEMP = MATRIX MULTIPLY XM XMT
    LET B0 = MATRIX ADDITION B0 B0TEMP
END OF LOOP
PRINT “ “; PRINT “ “; PRINT “B0 MATRIX:”; PRINT B0
. CALCULATE Spooled = (N1-1)S1 + ... + (Ng-1)Sg)/(N1+...+ Ng - g)
. W = Spooled WITHOUT THE DENOMINATOR TERM
LET JUNK2 = 0 FOR I = 1 1 G
LET W = DIAGONAL MATRIX JUNK2
LOOP FOR K = 1 1 G
    LET ROWLAST = SUM N FOR I = 1 1 K
    LET ATEMP = N(K)
    LET ROWFIRST = ROWLAST - ATEMP + 1
    LET XTEMP = MATRIX DEFINITION X ROWLAST P ROWFIRST
    LET SI = VARIANCE-COVARIANCE MATRIX XTEMP
    LET ATEMP = ATEMP - 1
    LET SI = MATRIX MULTIPLICATION SI ATEMP
    LET W = MATRIX ADDITION W SI
END OF LOOP
LET DENOM = 1/(NTOT - NG)
LET SPOOL = MATRIX MULTIPLY W DENOM
PRINT “ “; PRINT “ “; PRINT “POOLED COVARIANCE MATRIX:”; PRINT SPOOL
LET WINV = MATRIX INVERSE W
LET WINVB = MATRIX MULTIPLICATION WINV B0
. COMPUTE EIGENVALUES AND SORT IN DECREASING ORDER
. COMPUTE EIGENVECTORS, ONLY KEEP REAL COMPONENT, SORT
LET INDX = SEQUENCE 1 1 P
LET E = MATRIX EIGENVALUES WINVB
RETAIN E FOR I = 1 1 P
LET ESORT = SORTC E INDX
```

```
LET REVERSE = SEQUENCE P 1 1
LET REVERSE = SORTC REVERSE ESORT INDX
LET EV = MATRIX EIGENVECTORS WINVB
LET DIAG = 0 FOR I = 1 1 P
LET EVECT = DIAGONAL MATRIX DIAG
LOOP FOR K = 1 1 P
    LET LTAG = INDX(K)
    LET EVECT^LTAG = EV^K FOR I = 1 1 P
END OF LOOP
PRINT " "; PRINT " "; PRINT "EIGENVALUES:"; PRINT ESORT
. NORMALIZE L'SpooledL =1
. DIST = L'SpooledL, MULTIPLY EIGENVECTOR BY 1/SQRT(DIST)
LOOP FOR K = 1 1 P
    LET LTEMP = MATRIX DEFINITION EVECT^K P 1
    LET LTEMPT = MATRIX TRANSPOSE LTEMP
    LET MATTMP = MATRIX MULTIPLY LTEMPT SPOOL
    LET MATTMZ = MATRIX MULTIPLY MATTMP LTEMP
    LET DIST = MATTMZ1(1)
    LET FACT = 1/SQRT(DIST)
    LET EVECT^K = FACT*EVECT^K
END OF LOOP
PRINT " "; PRINT " "
PRINT "COLUMNS ARE THE DISCRIMINANT FUNCTIONS:"; PRINT EVECT
. PLOT FIRST 2 DISCRIMINANTS
LET ZY = 0 FOR I = 1 1 NTOT
LOOP FOR K = 1 1 P
    LET FACT = EVECT1(K)
    LET ZY = ZY + FACT*X^K
END OF LOOP
LET ZX = 0 FOR I = 1 1 NTOT
LOOP FOR K = 1 1 P
    LET FACT = EVECT2(K)
    LET ZX = ZX + FACT*X^K
END OF LOOP
LOOP FOR K = 1 1 NG
    LET AY = MEAN ZY SUBSET TAG = K
    LET AX = MEAN ZX SUBSET TAG = K
    LET GMEANX(K) = AX
    LET GMEANY(K) = AY
END OF LOOP
PRINT " "; PRINT " "; PRINT "GROUP MEANS:"; PRINT GMEANX GMEANY
Y1LABEL FIRST DISCRIMINANT
X1LABEL SECOND DISCRIMINANT
CHARACTER CIRCLE SQUARE TRIANGLE
LINE BLANK ALL
LEGEND 1 CIRC() - SPECIES 1
LEGEND 2 SQUA() - SPECIES 2
LEGEND 3 TRIA() - SPECIES 3
LEGEND FONT DUPLEX
LEGEND SIZE 1.2
PLOT ZY ZX TAG AND
PLOT GMEANY GMEANX
```

The following output is generated.

```
B0 MATRIX:

VARIABLES--B01 B02 B03 B04

 0.1264242E+01 -0.3990533E+00 0.4142301E+01 0.1332920E+01
 -0.3990533E+00 0.2268987E+00 -0.1515458E+01 -0.1713200E+00
 0.4142301E+01 -0.1515458E+01 0.1400072E+02 0.3853480E+01
 0.1332920E+01 -0.1713200E+00 0.3853480E+01 0.2021600E+01


POOLED COVARIANCE MATRIX:

VARIABLES--SPOOL1 SPOOL2 SPOOL3 SPOOL4

 0.2650082E+00 0.9272107E-01 0.1675143E+00 0.3840136E-01
 0.9272107E-01 0.1153878E+00 0.5524353E-01 0.3271021E-01
 0.1675143E+00 0.5524353E-01 0.1851877E+00 0.4266530E-01
 0.3840136E-01 0.3271021E-01 0.4266530E-01 0.4188164E-01


EIGENVALUES:

VARIABLES--ESORT

 0.8901302E+00
 0.2225433E+00
 0.8941539E-09
 -0.7700800E-07


COLUMNS ARE THE DISCRIMINANT FUNCTIONS:

VARIABLES--EVECT1 EVECT2 EVECT3 EVECT4

 0.1369849E+01 0.8875510E+00 0.5686842E-01 0.2306328E+01
 0.9835789E+00 -0.9720638E+00 -0.2889864E+01 0.3226218E-01
 -0.3014444E+01 -0.2081233E+01 -0.5299196E+00 -0.5492300E+00
 -0.1221005E+01 0.5894730E+01 0.7277119E+00 -0.4710006E+00


GROUP MEANS:

VARIABLES--GMEANX GMEANY

 0.1599419E+01 0.8536134E+01
 -0.4368486E+01 -0.2383641E+01
 0.3343984E+01 -0.7260220E+01
```



PLOT FIRST 2 DISCRIMINANT FUNCTIONS
O - SPECIES 1
□ - SPECIES 2
Δ - SPECIES 3
FIRST DISCRIMINANT
SECOND DISCRIMINANT
20
15
10
5
0
-5
-10
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5