NAG Fortran Library Routine Document

G02BWF

Note: before using this routine, please read the Users' Note for your implementation to check the interpretation of ***bold italicised*** terms and other implementation-dependent details.

1 Purpose

G02BWF calculates a matrix of Pearson product-moment correlation coefficients from sums of squares and cross-products of deviations about the mean.

2 Specification

```
SUBROUTINE G02BWF(M, R, IFAIL)
INTEGER           M, IFAIL
real              R((M*M+M)/2)
```

3 Description

G02BWF calculates a matrix of Pearson product-moment correlation coefficients from sums of squares and cross-products about the mean for observations on m variables which can be computed by a single call to G02BUF or a series of calls to G02BTF. The sums of squares and cross-products are stored in an array packed by column and are overwritten by the correlation coefficients.

Let c_{jk} be the cross-product of deviations from the mean for variables $j = 1, 2, \ldots, m$; $k = j, j+1, \ldots, m$, then the product-moment correlation coefficient, r_{jk} is given by

$$r_{jk} = \frac{c_{jk}}{\sqrt{c_{jj}c_{kk}}}.$$

4 References

None.

5 Parameters

1: M – INTEGER *Input*

On entry: the number, m, of variables.

Constraint: M ≥ 1.

2: R((M∗M+M)/2) – ***real*** array *Input/Output*

On entry: R contains the upper triangular part of the sums of squares and cross-products matrix of deviations from the mean. These are stored packed by column, i.e., the cross-product between variable j and k, $k \geq j$, is stored in R$(k \times (k-1)/2 + j)$.

On exit: Pearson product-moment correlation coefficients.

These are stored packed by column corresponding to the input cross-products.

3: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. Users who are unfamiliar with this parameter should refer to Chapter P01 for details.

On exit: IFAIL $= 0$ unless the routine detects an error (see Section 6).

For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, because for this routine the values of the output parameters may be useful even if IFAIL $\neq 0$ on exit, the recommended value is -1. **When the value -1 or 1 is used it is essential to test the value of IFAIL on exit.**

6 Error Indicators and Warnings

If on entry IFAIL $= 0$ or -1, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

IFAIL $= 1$

On entry, M < 1.

IFAIL $= 2$

A variable has a zero variance. All correlations involving the variable with zero variance will be returned as zero.

7 Accuracy

The accuracy of this routine is entirely dependent upon the accuracy of the elements of array R.

8 Further Comments

G02BWF may also be used to calculate the correlations between parameter estimates from the variance-covariance matrix of the parameter estimates as is given by several routines in this chapter.

9 Example

A program to calculate the correlation matrix from raw data. The sum of squares and cross-products about the mean are calculated from the raw data by a call to G02BUF. The correlation matrix is then calculated from these values.

9.1 Program Text

Note: the listing of the example program presented below uses ***bold italicised*** terms to denote precision-dependent details. Please read the Users' Note for your implementation to check the interpretation of these terms. As explained in the Essential Introduction to this manual, the results produced may not be identical for all implementations.

```
*     G02BWF Example Program Text
*     Mark 14 Release.  NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          LDX, MMAX, MP
      PARAMETER        (LDX=12,MMAX=12,MP=(MMAX*(MMAX+1))/2)
*     .. Local Scalars ..
      real             SW
      INTEGER          IFAIL, J, K, M, N
      CHARACTER        MEAN, WEIGHT
*     .. Local Arrays ..
      real             C(MP), WMEAN(MMAX), WT(LDX), X(LDX,MMAX)
*     .. External Subroutines ..
      EXTERNAL         G02BUF, G02BWF, X04CCF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G02BWF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*,END=20) MEAN, WEIGHT, M, N
      IF (M.LE.MMAX .AND. N.LE.LDX) THEN
```

```
         READ (NIN,*) (WT(J),J=1,N)
         READ (NIN,*) ((X(J,K),K=1,M),J=1,N)
         IFAIL = 0
*
*        Calculate the sums of squares and cross-products matrix
         CALL G02BUF(MEAN,WEIGHT,N,M,X,LDX,WT,SW,WMEAN,C,IFAIL)
*
         IFAIL = -1
*
*        Calculate the correlation matrix
         CALL G02BWF(M,C,IFAIL)
*
*        Print the correlation matrix
         IF (IFAIL.EQ.0) THEN
            WRITE (NOUT,*)
            CALL X04CCF('Upper','Non-unit',M,C,'Correlation matrix',
     +                  IFAIL)
         ELSE IF (IFAIL.EQ.2) THEN
            WRITE (NOUT,*)
            WRITE (NOUT,*) ' NOTE: some variances are zero'
            WRITE (NOUT,*)
            CALL X04CCF('Upper','Non-unit',M,C,'Correlation matrix',
     +                  IFAIL)
         END IF
       ELSE
         WRITE (NOUT,99999) 'M or N is too large. M =', M, ', N =', N
       END IF
   20 STOP
*
99999 FORMAT (1X,A,I6,A,I6)
      END
```

9.2 Program Data

```
G02BWF Example Program Data
 'M'     'W'      3     3
  0.1300   1.3070   0.3700
  9.1231   3.7011   4.5230
  0.9310   0.0900   0.8870
  0.0009   0.0099   0.0999
```

9.3 Program Results

```
 G02BWF Example Program Results

 Correlation matrix
          1       2       3
 1   1.0000  0.9908  0.9903
 2           1.0000  0.9624
 3                   1.0000
```