NAG Fortran Library Routine Document

G05QBF

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

G05QBF generates a random correlation matrix with given eigenvalues.

2 Specification

```
SUBROUTINE G05QBF(N, D, C, LDC, EPS, IGEN, ISEED, WK, IFAIL)
INTEGER           N, LDC, IGEN, ISEED(4), IFAIL
real              D(N), C(LDC,N), EPS, WK(2*N)
```

3 Description

Given n eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_n$, such that

$$\sum_{i=1}^{n} \lambda_i = n$$

and

$$\lambda_i \geq 0, \quad i = 1, 2, \ldots, n,$$

G05QBF will generate a random correlation matrix, C, of dimension n, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

The method used is based on that described by Lin and Bendel (1985). Let D be the diagonal matrix with values $\lambda_1, \lambda_2, \ldots, \lambda_n$ and let A be a random orthogonal matrix generated by G05QAF then the matrix $C_0 = ADA^T$ is a random covariance matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. The matrix C_0 is transformed into a correlation matrix by means of $n-1$ elementary rotation matrices P_i such that $C = P_{n-1}P_{n-2}\ldots P_1 C_0 P_1^T \ldots P_{n-2}^T P_{n-1}^T$. The restriction on the sum of eigenvalues implies that for any diagonal element of $C_0 > 1$, there is another diagonal element < 1. The P_i are constructed from such pairs, chosen at random, to produce a unit diagonal element corresponding to the first element. This is repeated until all diagonal elements are 1 to within a given tolerance ϵ.

The randomness of C should be interpreted only to the extent that A is a random orthogonal matrix and C is computed from A using the P_i which are chosen as arbitrarily as possible.

One of the initialisation routines G05KBF (for a repeatable sequence if computed sequentially) or G05KCF (for a non-repeatable sequence) must be called prior to the first call to G05QBF.

4 References

Lin S P and Bendel R B (1985) Algorithm AS213: Generation of population correlation on matrices with specified eigenvalues *Appl. Statist.* **34** 193–198

5 Parameters

1: N – INTEGER *Input*

On entry: the dimension of the correlation matrix to be generated, n.

Constraint: N ≥ 1.

2: D(N) – ***real*** array *Input*

On entry: the n eigenvalues, λ_i, for $i = 1, 2, \ldots, n$.

Constraints:

$$\mathrm{D}(i) \geq 0.0, \text{ for } i = 1, 2, \ldots, n, \text{ and } \sum_{i=1}^{n} \mathrm{D}(i) = n \text{ to within EPS.}$$

3: C(LDC,N) – ***real*** array *Output*

On exit: a random correlation matrix, C, of dimension n.

4: LDC – INTEGER *Input*

On entry: the first dimension of the array C as declared in the (sub)program from which G05QBF is called.

Constraint: LDC $\geq$ N.

5: EPS – ***real*** *Input*

On entry: the maximum acceptable error in the diagonal elements, ϵ.

Constraint: EPS $\geq$ N $\times$ ***machine precision*** (see Chapter X02).

Suggested value: EPS=0.00001.

6: IGEN – INTEGER *Input*

On entry: must contain the identification number for the generator to be used to return a pseudo-random number and should remain unchanged following initialisation by a prior call to one of the routines G05KBF or G05KCF.

7: ISEED(4) – INTEGER array *Input/Output*

On entry: contains values which define the current state of the selected generator.

On exit: contains updated values defining the new state of the selected generator.

8: WK(2*N) – ***real*** array *Workspace*

9: 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, for users not familiar with this parameter the recommended value is 0. **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, N < 1,
or LDC $<$ N.
or EPS $<$ N $\times$ ***machine precision***.

IFAIL = 2

On entry, $\mathrm{D}(i) < 0.0$ for some i,

or $\sum_{i=1}^{n} \mathrm{D}(i) \neq n$ to within EPS.

IFAIL = 3

The error in a diagonal element is greater than EPS. The value of EPS should be increased. Otherwise the program could be re-run with a different value used for the seed of the random number generator, see G05KBF or G05KCF.

7 Accuracy

The maximum error in a diagonal element is given by EPS.

8 Further Comments

The time taken by the routine is approximately proportional to n^2.

9 Example

Following initialisation of the pseudo random number generator by a call to G05KBF, a 3 by 3 correlation matrix with eigenvalues of 0.7, 0.9 and 1.4 is generated and printed.

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.

```
*     G05QBF Example Program Text
*     Mark 20 Release. NAG Copyright 2001.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          NMAX
      PARAMETER        (NMAX=10)
*     .. Local Scalars ..
      real             EPS
      INTEGER          I, IFAIL, IGEN, J, LDC, N
*     .. Local Arrays ..
      real             C(NMAX,NMAX), D(NMAX), WK(2*NMAX)
      INTEGER          ISEED(4)
*     .. External Subroutines ..
      EXTERNAL         G05CBF, G05KBF, G05QBF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G05QBF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) N
      IF (N.LE.NMAX) THEN
         READ (NIN,*) (D(I),I=1,N)
*
         WRITE (NOUT,*)
*
         LDC = NMAX
         CALL G05CBF(0)
         EPS = 0.0001e0
*
*        IGEN identifies the stream.
         IGEN = 1
*        Initialise the seed to a repeatable sequence
         ISEED(1) = 1762543
         ISEED(2) = 9324783
```

```
         ISEED(3) = 423446
         ISEED(4) = 742355
         CALL G05KBF(IGEN,ISEED)
*
         IFAIL = 0
*
         CALL G05QBF(N,D,C,LDC,EPS,IGEN,ISEED,WK,IFAIL)
*
         DO 20 I = 1, N
            WRITE (NOUT,99999) (C(I,J),J=1,N)
   20    CONTINUE
      END IF
      STOP
*
99999 FORMAT (1X,3F9.3)
      END
```

9.2 Program Data

None.

9.3 Program Results

```
 G05QBF Example Program Results

     1.000    0.204   -0.106
     0.204    1.000   -0.278
    -0.106   -0.278    1.000
```