NAG Library Routine Document

G05PYF

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

G05PYF generates a random correlation matrix with given eigenvalues.

2 Specification

```
SUBROUTINE G05PYF (N, D, EPS, STATE, C, LDC, IFAIL)

INTEGER             N, STATE(*), LDC, IFAIL
REAL (KIND=nag_wp) D(N), EPS, C(LDC,N)
```

3 Description

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

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

and

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

G05PYF 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 G05PXF then the matrix $C_0 = ADA^{\mathrm{T}}$ 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 $C = P_{n-1} P_{n-2} \ldots P_1 C_0 P_1^{\mathrm{T}} \ldots P_{n-2}^{\mathrm{T}} P_{n-1}^{\mathrm{T}}$. 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 initialization routines G05KFF (for a repeatable sequence if computed sequentially) or G05KGF (for a non-repeatable sequence) must be called prior to the first call to G05PYF.

4 References

Lin S P and Bendel R B (1985) Algorithm AS 213: Generation of population correlation on matrices with specified eigenvalues *Appl. Statist.* **34** 193–198

5 Parameters

1: N – INTEGER *Input*

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

Constraint: $\mathrm{N} \geq 1$.

2: D(N) – REAL (KIND=nag_wp) array *Input*

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

Constraints:

$\mathrm{D}(i) \geq 0.0$, for $i = 1, 2, \ldots, n$;

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

3: EPS – REAL (KIND=nag_wp) *Input*

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

Suggested value: EPS = 0.00001.

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

4: STATE($*$) – INTEGER array *Communication Array*

Note: the actual argument supplied must be the array STATE supplied to the initialization routines G05KFF or G05KGF.

On entry: contains information on the selected base generator and its current state.

On exit: contains updated information on the state of the generator.

5: C(LDC,N) – REAL (KIND=nag_wp) array *Output*

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

6: LDC – INTEGER *Input*

On entry: the first dimension of the array C as declared in the (sub)program from which G05PYF is called.

Constraint: LDC $\geq$ N.

7: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction for details.

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, if you are 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.**

On exit: IFAIL = 0 unless the routine detects an error or a warning has been flagged (see Section 6).

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.

IFAIL = 2

On entry, $\mathrm{D}(i) < 0.0$ for some i, or $$\sum_{i=1}^{n} \mathrm{D}(i) \neq n.$$

IFAIL = 3

On entry, EPS < N × ***machine precision***.

IFAIL = 4

On entry, STATE vector was not initialized or has been corrupted.

IFAIL = 5

The error in a diagonal element is greater than EPS. The value of EPS should be increased. Otherwise the program could be rerun with a different value used for the seed of the random number generator, see G05KFF or G05KGF.

IFAIL = 6

On entry, LDC < N.

7 Accuracy

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

8 Further Comments

The time taken by G05PYF is approximately proportional to n^2.

9 Example

Following initialization of the pseudorandom number generator by a call to G05KFF, a 3 by 3 correlation matrix with eigenvalues of 0.7, 0.9 and 1.4 is generated and printed.

9.1 Program Text

```
    Program g05pyfe

!     G05PYF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: g05kff, g05pyf, nag_wp, x04caf
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: lseed = 1, nin = 5, nout = 6
!     .. Local Scalars ..
      Real (Kind=nag_wp)               :: eps
      Integer                          :: genid, ifail, ldc, lstate, n, subid
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: c(:,:), d(:)
      Integer                          :: seed(lseed)
      Integer, Allocatable             :: state(:)
!     .. Executable Statements ..
      Write (nout,*) 'G05PYF Example Program Results'
      Write (nout,*)
      Flush (nout)

!     Skip heading in data file
      Read (nin,*)

!     Read in the base generator information and seed
```

```
      Read (nin,*) genid, subid, seed(1)

!     Initial call to initialiser to get size of STATE array
      lstate = 0
      Allocate (state(lstate))
      ifail = 0
      Call g05kff(genid,subid,seed,lseed,state,lstate,ifail)

!     Reallocate STATE
      Deallocate (state)
      Allocate (state(lstate))

!     Initialize the generator to a repeatable sequence
      ifail = 0
      Call g05kff(genid,subid,seed,lseed,state,lstate,ifail)

!     Read in the problem size and tolerance
      Read (nin,*) n, eps

      ldc = n
      Allocate (c(ldc,n),d(n))

!     Read in the eigenvalues
      Read (nin,*) d(1:n)

!     Generate the correlation matrix
      ifail = 0
      Call g05pyf(n,d,eps,state,c,ldc,ifail)

!     Display the results
      ifail = 0
      Call x04caf('General',' ',n,n,c,ldc,'Correlation Matrix',ifail)

    End Program g05pyfe
```

9.2 Program Data

```
G05PYF Example Program Data
1  1  1762543      :: GENID,SUBID,SEED(1)
3   1.0E-5         :: N,EPS
0.7 0.9 1.4        :: D
```

9.3 Program Results

```
 G05PYF Example Program Results

 Correlation Matrix
          1       2       3
 1   1.0000 -0.2549 -0.1004
 2  -0.2549  1.0000  0.2343
 3  -0.1004  0.2343  1.0000
```