NAG Library Routine Document

G05RDF

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

G05RDF sets up a reference vector and generates an array of pseudorandom numbers from a Normal (Gaussian) copula with covariance matrix C.

2 Specification

```
SUBROUTINE G05RDF(MODE, N, M, C, LDC, R, LR, STATE, X, LDX, IFAIL)
INTEGER           MODE, N, M, LDC, LR, STATE(lstate), LDX, IFAIL
double precision  C(LDC,M), R(LR), X(LDX,M)
```

3 Description

The Gaussian copula, G, is defined by

$$G(u_1, u_2, \ldots, u_m; C) = \Phi_C\left(\phi_{C_{11}}^{-1}(u_1), \phi_{C_{22}}^{-1}(u_2), \ldots, \phi_{C_{mm}}^{-1}(u_m)\right)$$

where m is the number of dimensions, Φ_C is the multivariate Normal density function with mean zero and covariance matrix C and $\phi_{C_{ii}}^{-1}$ is the inverse of the univariate Normal density function with mean zero and variance C_{ii}.

G05RZF is used to generate a vector from a multivariate Normal distribution and G01EAF is used to convert each element of that vector into a uniformly distributed value between zero and one.

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 G05RDF.

4 References

Nelsen R B (1998) *An Introduction to Copulas. Lecture Notes in Statistics 139* Springer

Sklar A (1973) Random Variables: Joint Distribution Functions and Copulas *Kybernetika* **9** 499–460

5 Parameters

1: MODE – INTEGER *Input*

On entry: a code for selecting the operation to be performed by the routine.

MODE $= 0$
: Set up reference vector only.

MODE $= 1$
: Generate variates using reference vector set up in a prior call to G05RDF.

MODE $= 2$
: Set up reference vector and generate two-way table.

Constraint: MODE $= 0$, 1 or 2.

2: N – INTEGER *Input*

On entry: n, the number of random variates required.

Constraint: N ≥ 0.

3: M – INTEGER *Input*

On entry: m, the number of dimensions of the distribution.

Constraint: M > 0.

4: C(LDC,M) – ***double precision*** array *Input*

On entry: the covariance matrix of the distribution. Only the upper triangle need be set.

Constraint: C must be positive semi-definite to ***machine precision***

5: LDC – INTEGER *Input*

On entry: the first dimension of the array C as declared in the (sub)program from which G05RDF is called.

Constraint: LDC $\geq$ M.

6: R(LR) – ***double precision*** array *Communication Array*

On entry: if MODE $= 1$, the reference vector as set up by G05RDF in a previous call with MODE $= 0$ or 2.

On exit: if MODE $= 0$ or 2, the reference vector that can be used in subsequent calls to G05RDF with MODE $= 1$.

7: LR – INTEGER *Input*

On entry: the dimension of the array R as declared in the (sub)program from which G05RDF is called. If MODE $= 1$, it must be the same as the value of LR specified in the prior call to G05RDF with MODE $= 0$ or 2.

Constraint: LR $\geq$ M $\times$ (M $+ 1) + 1$.

8: STATE($lstate$) – INTEGER array *Communication Array*

Note: $lstate$ = LSTATE, the dimension of STATE as supplied to the initialization routine 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.

9: X(LDX,M) – ***double precision*** array *Output*

On exit: the array of values from a multivariate Gaussian copula, with X(i, j) holding the jth dimension for the ith variate.

10: LDX – INTEGER *Input*

On entry: the first dimension of the array X as declared in the (sub)program from which G05RDF is called.

Constraint: LDX $\geq$ N.

11: 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.

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, 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.**

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, MODE $\neq 0$, 1 or 2.

IFAIL $= 2$

On entry, N < 1.

IFAIL $= 3$

On entry, M < 1.

IFAIL $= 4$

The covariance matrix C is not positive semi-definite to ***machine precision***.

IFAIL $= 5$

On entry, LDC $<$ M.

IFAIL $= 6$

The reference vector R has been corrupted or M has changed since R was set up in a previous call to G05RDF with MODE $= 0$ or 2.

IFAIL $= 7$

On entry, LR $\leq$ M $\times$ (M $+ 1$).

IFAIL $= 8$

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

IFAIL $= 10$

On entry, LDX $<$ N.

7 Accuracy

See Section 7 in G05RZF for an indication of the accuracy of the underlying multivariate Normal distribution.

8 Further Comments

The time taken by G05RDF is of order nm^3.

It is recommended that the diagonal elements of C should not differ too widely in order of magnitude. This may be achieved by scaling the variables if necessary. The actual matrix decomposed is $C + E = LL^{\mathrm{T}}$, where E is a diagonal matrix with small positive diagonal elements. This ensures that,

even when C is singular, or nearly singular, the Cholesky factor L corresponds to a positive-definite covariance matrix that agrees with C within ***machine precision***.

9 Example

This example prints ten pseudorandom observations from a Normal copula with covariance matrix

$$\begin{bmatrix} 1.69 & 0.39 & -1.86 & 0.07 \\ 0.39 & 98.01 & -7.07 & -0.71 \\ -1.86 & -7.07 & 11.56 & 0.03 \\ 0.07 & -0.71 & 0.03 & 0.01 \end{bmatrix},$$

generated by G05RDF. All ten observations are generated by a single call to G05RDF with MODE $= 2$. The random number generator is initialized by G05KFF.

9.1 Program Text

```
*     G05RDF Example Program Text
*     Mark 22 Release. NAG Copyright 2006.
*     .. Parameters ..
      INTEGER          NOUT
      PARAMETER        (NOUT=6)
      INTEGER          MSEED, MSTATE, LDC, LDX, LR
      PARAMETER        (MSEED=1,MSTATE=633,LDC=5,LDX=100,
     +                 LR=LDC*LDC+LDC+1)
*     .. Local Scalars ..
      INTEGER          GENID, I, IFAIL, J, LSEED, LSTATE, M, MODE, N,
     +                 SUBID
*     .. Local Arrays ..
      DOUBLE PRECISION C(LDC,LDC), R(LR), X(LDX,LDC)
      INTEGER          SEED(MSEED), STATE(MSTATE)
*     .. External Subroutines ..
      EXTERNAL         G05KFF, G05RDF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G05RDF Example Program Results'
      WRITE (NOUT,*)
*     Set the number of variables and variates
      M = 4
      N = 10
*     Input the upper triangle portion of the covariance matrix
      C(1,1) = 1.69D0
      C(1,2) = 0.39D0
      C(1,3) = -1.86D0
      C(1,4) = 0.07D0
      C(2,2) = 98.01D0
      C(2,3) = -7.07D0
      C(2,4) = -0.71D0
      C(3,3) = 11.56D0
      C(3,4) = 0.03D0
      C(4,4) = 0.01D0
*     Initialize the seed
      SEED(1) = 1762543
*     Choose the random generator to use
      GENID = 1
      SUBID = 1
*     Initialize the generator to a repeatable sequence
      LSTATE = MSTATE
      LSEED = MSEED
      IFAIL = 1
      CALL G05KFF(GENID,SUBID,SEED,LSEED,STATE,LSTATE,IFAIL)
      IF (IFAIL.NE.0) THEN
         WRITE (NOUT,99998) IFAIL
         GO TO 40
      END IF
*     Set up reference vector and generate variates
*     Choose MODE = 2
      MODE = 2
      IFAIL = 1
```

```
      CALL G05RDF(MODE,N,M,C,LDC,R,LR,STATE,X,LDX,IFAIL)
      IF (IFAIL.NE.0) THEN
         WRITE (NOUT,99997) IFAIL
         GO TO 40
      END IF
*     Display the results
      DO 20 I = 1, N
         WRITE (NOUT,99999) (X(I,J),J=1,M)
   20 CONTINUE
*
   40 CONTINUE
*
99999 FORMAT (1X,10F10.4)
99998 FORMAT (1X,' ** G05KFF returned with IFAIL = ',I5)
99997 FORMAT (1X,' ** G05RDF returned with IFAIL = ',I5)
      END
```

9.2 Program Data

None.

9.3 Program Results

```
 G05RDF Example Program Results

     0.6364    0.0517    0.4137    0.8817
     0.1065    0.2461    0.7993    0.3806
     0.7460    0.6313    0.2708    0.5421
     0.7983    0.0564    0.6868    0.9234
     0.1046    0.5790    0.8533    0.2208
     0.4925    0.2784    0.3513    0.5158
     0.3843    0.2349    0.9472    0.7801
     0.7871    0.9941    0.9403    0.2044
     0.4982    0.9015    0.7176    0.2914
     0.6717    0.5359    0.5961    0.4487
```