NAG Fortran Library Routine Document

G05LZF

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

G05LZF sets up a reference vector and generates a vector of pseudo-random numbers from a multivariate Normal distribution with mean vector a and covariance matrix C.

2 Specification

```
SUBROUTINE G05LZF(MODE, N, XMU, C, IC, X, IGEN, ISEED, R, NR, IFAIL)
INTEGER            MODE, N, IC, IGEN, ISEED(4), NR, IFAIL
real               XMU(N), C(IC,N), X(N), R(NR)
```

3 Description

When the covariance matrix is non-singular (i.e., strictly positive-definite), the distribution has probability density function

$$f(x) = \sqrt{\frac{|C^{-1}|}{(2\pi)^n}} \exp\left\{-(x-a)^T C^{-1}(x-a)\right\}$$

where n is the number of dimensions, C is the covariance matrix, a is the vector of means and x is the vector of positions.

Covariance matrices are symmetric and positive semi-definite. Given such a matrix C, there exists a lower triangular matrix L such that $LL^T = C$. L is not unique, if C is singular.

G05LZF decomposes C to find such an L. It then stores n, a and L in the reference vector r which is used to generate a vector x of independent standard Normal pseudo-random numbers. It then returns the vector $a + Lx$, which has the required multivariate Normal distribution.

It should be noted that this routine will work with a singular covariance matrix C, provided C is positive semi-definite, despite the fact that the above formula for the probability density function is not valid in that case. Wilkinson (1965) should be consulted if further information is required.

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 G05LZF.

4 References

Knuth D E (1981) *The Art of Computer Programming (Volume 2)* (2nd Edition) Addison-Wesley

Wilkinson J H (1965) *The Algebraic Eigenvalue Problem* Oxford University Press, Oxford

5 Parameters

1: MODE – INTEGER *Input*

On entry: selects the operation to be performed:

MODE = 0

Initialise **and** generate random numbers.

MODE = 1

Initialise only (i.e., set up reference vector);

MODE = 2

Generate random numbers using previously set up reference vector;

Constraint: MODE = 0, 1 or 2.

2: N – INTEGER *Input*

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

Constraint: N > 0.

3: XMU(N) – ***real*** array *Input*

On entry: the vector of means, a, of the distribution.

4: C(IC,N) – ***real*** array *Input*

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

5: IC – INTEGER *Input*

On entry: the first dimension of the array C as declared in the (sub)program from which G05LZF is called.

Constraint: IC $\geq$ N.

6: X(N) – ***real*** array *Output*

On exit: the pseudo-random multivariate Normal vector generated by the routine.

7: 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.

8: 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.

9: R(NR) – ***real*** array *Input/Output*

On entry: if MODE = 2, the reference vector as set up by G05LZF in a previous call with MODE = 0 or 1.

On exit: if MODE = 0 or 1, the reference vector that can be used in subsequent calls to G05LZF with MODE = 2.

10: NR – INTEGER *Input*

On entry: the dimension of the array R as declared in the (sub)program from which G05LZF is called. If MODE = 2, it must be the same as the value of NR specified in the prior call to G05LZF with MODE = 0 or 1.

Constraint: NR $\geq$ (N + 1)(N + 2)/2.

11: 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.

IFAIL $= 2$

On entry, NR $< (\mathrm{N}+1) * (\mathrm{N}+2)/2$.

IFAIL $= 3$

On entry, IC $<$ N.

IFAIL $= 4$

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

IFAIL $= 5$

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

IFAIL $= 6$

The reference vector R has been corrupted or N has changed since R was set up in a previous call with MODE $= 0$ or 1.

7 Accuracy

The maximum absolute error in LL^T, and hence in the covariance matrix of the resulting vectors, is less than $(n \times \epsilon + (n+3)\epsilon/2)$ times the maximum element of C, where ϵ is the ***machine precision***. Under normal circumstances, the above will be small compared to sampling error.

8 Further Comments

The time taken by the routine is of order n^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 $C + E = LL^T$, 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

The example program prints two pseudo-random observations from a bivariate Normal distribution with means vector

$$\begin{bmatrix} 1.0 \\ 2.0 \end{bmatrix}$$

and covariance matrix

$$\begin{bmatrix} 2.0 & 1.0 \\ 1.0 & 3.0 \end{bmatrix},$$

generated by G05LZF. The first observation is generated by a single call to G05LZF with MODE = 0, and the second observation is generated using the same reference vector a call to G05LZF with MODE = 2. The random number generator is initialised by G05KBF.

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.

```
*     G05LZF Example Program Text
*     Mark 20 Release. NAG Copyright 2001.
*     .. Parameters ..
      INTEGER          NOUT, M, NR
      PARAMETER        (NOUT=6,M=2,NR=((M+1)*(M+2))/2)
*     .. Local Scalars ..
      INTEGER          IFAIL, IGEN
*     .. Local Arrays ..
      real             C(M,M), R(NR), X(M), XMU(M)
      INTEGER          ISEED(4)
*     .. External Subroutines ..
      EXTERNAL         G05KBF, G05LZF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G05LZF Example Program Results'
      WRITE (NOUT,*)
*     Initialise the seed to a repeatable sequence
      ISEED(1) = 1762543
      ISEED(2) = 9324783
      ISEED(3) = 42344
      ISEED(4) = 742355
*     IGEN identifies the stream.
      IGEN = 1
      CALL G05KBF(IGEN,ISEED)
*
      C(1,1) = 2.0e0
      C(2,1) = 1.0e0
      C(1,2) = 1.0e0
      C(2,2) = 3.0e0
*
      XMU(1) = 1.0e0
      XMU(2) = 2.0e0

      IFAIL = 0

*     Set up reference vector and generate numbers
      CALL G05LZF(0,M,XMU,C,M,X,IGEN,ISEED,R,NR,IFAIL)
      WRITE (NOUT,99999) X

*     Generate numbers
      CALL G05LZF(2,M,XMU,C,M,X,IGEN,ISEED,R,NR,IFAIL)
      WRITE (NOUT,99999) X

      STOP
*
99999 FORMAT (1X,F10.4)
      END
```

9.2 Program Data

None.

9.3 Program Results

```
 G05LZF Example Program Results

     3.9620
     2.4272
     0.9189
    -0.1605
```