# NAG Library Routine Document

# G05RZF

**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

G05RZF sets up a reference vector and generates an array of pseudorandom numbers from a multivariate Normal distribution with mean vector $a$ and covariance matrix $C$.

## 2 Specification

```
SUBROUTINE G05RZF(MODE, N, M, XMU, C, LDC, R, LR, STATE, X, LDX, IFAIL)
INTEGER            MODE, N, M, LDC, LR, STATE(lstate), LDX, IFAIL
double precision   XMU(M), C(LDC,M), R(LR), X(LDX,M)
```

## 3 Description

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

$$f(x) = \sqrt{\frac{|C^{-1}|}{(2\pi)^m}} \exp\Big(-(x-a)^{\mathrm{T}} C^{-1} (x-a)\Big)$$

where $m$ 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 $LL^{\mathrm{T}} = C$. $L$ is not unique, if $C$ is singular.

G05RZF decomposes $C$ to find such an $L$. It then stores $m$, $a$ and $L$ in the reference vector $r$ which is used to generate a vector $x$ of independent standard Normal pseudorandom 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 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 G05RZF.

## 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*: 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 G05RZF.

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 $\geq 0$.

3: M – INTEGER *Input*

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

*Constraint*: M $> 0$.

4: XMU(M) – ***double precision*** array *Input*

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

5: 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***

6: LDC – INTEGER *Input*

*On entry*: the first dimension of the array C as declared in the (sub)program from which G05RZF is called.

*Constraint*: LDC $\geq$ M.

7: R(LR) – ***double precision*** array *Input/Output*

*On entry*: if MODE $= 1$, the reference vector as set up by G05RZF 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 G05RZF with MODE $= 1$.

8: LR – INTEGER *Input*

*On entry*: the dimension of the array R as declared in the (sub)program from which G05RZF is called. If MODE $= 1$, it must be the same as the value of LR specified in the prior call to G05RZF with MODE $= 0$ or $2$.

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

9: 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.

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

*On exit*: the array of pseudorandom multivariate Normal vectors generated by the routine, with X($i, j$) holding the $j$th dimension for the $i$th variate.

11: LDX – INTEGER *Input*

*On entry*: the first dimension of the array X as declared in the (sub)program from which G05RZF is called.

*Constraint*: LDX $\geq$ N.

12: 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 $= 5$

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

IFAIL $= 6$

On entry, LDC $<$ M.

IFAIL $= 7$

The reference vector R has been corrupted or M has changed since R was set up in a previous call to G05RZF with MODE $= 0$ or 2.

IFAIL $= 8$

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

IFAIL $= 9$

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

IFAIL $= 11$

On entry, LDX $<$ N.

# 7 Accuracy

Not applicable.

# 8 Further Comments

The time taken by G05RZF 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 multivariate Normal distribution with means vector

$$\begin{bmatrix} 1.0 \\ 2.0 \\ -3.0 \\ 0.0 \end{bmatrix}$$

and 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 G05RZF. All ten observations are generated by a single call to G05RZF with MODE $= 2$. The random number generator is initialized by G05KFF.

## 9.1 Program Text

```
*     G05RZF 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), XMU(LDC)
      INTEGER          SEED(MSEED), STATE(MSTATE)
*     .. External Subroutines ..
      EXTERNAL         G05KFF, G05RZF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G05RZF 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
*     Input the means
      XMU(1) = 1.0D0
      XMU(2) = 2.0D0
      XMU(3) = -3.0D0
      XMU(4) = 0.0D0
*     Initialize the seed
      SEED(1) = 1762543
*     GENID and SUBID identify the base generator
      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 N numbers
*     Choose MODE = 2
      MODE = 2
      IFAIL = 1
      CALL G05RZF(MODE,N,M,XMU,C,LDC,R,LR,STATE,X,LDX,IFAIL)
      IF (IFAIL.NE.0) THEN
         WRITE (NOUT,99997) IFAIL
         GO TO 40
      END IF
*     Display the variates
      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,' ** G05RZF returned with IFAIL = ',I5)
      END
```

## 9.2 Program Data

None.

## 9.3 Program Results

```
 G05RZF Example Program Results

    1.4534  -14.1206   -3.7410    0.1184
   -0.6191   -4.8000   -0.1473   -0.0304
    1.8607    5.3206   -5.0753    0.0106
    2.0861  -13.6996   -1.3451    0.1428
   -0.6326    3.9729    0.5721   -0.0770
    0.9754   -3.8162   -4.2978    0.0040
    0.6174   -5.1573    2.5037    0.0772
    2.0352   26.9359    2.2939   -0.0826
    0.9941   14.7700   -1.0421   -0.0549
    1.5780    2.8916   -2.1725   -0.0129
```