NAG Toolbox for MATLAB

g05lz

1 Purpose

g05lz 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 Syntax

```
[x, iseed, r, ifail] = g05lz(mode, xmu, c, igen, iseed, r, 'n', n)
```

3 Description

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

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

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

g05lz 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 function 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 functions g05kb (for a repeatable sequence if computed sequentially) or g05kc (for a non-repeatable sequence) must be called prior to the first call to g05lz.

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

5.1 Compulsory Input Parameters

1: **mode – int32 scalar**

Selects the operation to be performed:

mode $= 0$

Initialize and generate random numbers.

mode $= 1$

Initialize only (i.e., set up reference vector).

mode $= 2$

Generate random numbers using previously set up reference vector.

Constraint: $0 \leq$ **mode** ≤ 2.

2: **xmu**(**n**) – **double array**

a, the vector of means of the distribution.

3: **c**(**ldc**,**n**) – **double array**

ldc, the first dimension of the array, must be at least **n**.

The covariance matrix of the distribution. Only the upper triangle need be set.

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

4: **igen** – **int32 scalar**

Must contain the identification number for the generator to be used to return a pseudo-random number and should remain unchanged following initialization by a prior call to g05kb or g05kc.

5: **iseed**(4) – **int32 array**

Contains values which define the current state of the selected generator.

6: **r**(**nr**) – **double array**

If **mode** $= 2$, the reference vector as set up by g05lz in a previous call with **mode** $= 0$ or 1.

5.2 Optional Input Parameters

1: **n** – **int32 scalar**

Default: The dimension of the arrays **xmu**, **c**, **x**. (An error is raised if these dimensions are not equal.)

n, the number of dimensions of the distribution.

Constraint: **n** > 0.

5.3 Input Parameters Omitted from the MATLAB Interface

ldc, nr

5.4 Output Parameters

1: **x**(**n**) – **double array**

The pseudo-random multivariate Normal vector generated by the function.

2: **iseed**(4) – **int32 array**

Contains updated values defining the new state of the selected generator.

3: **r**(**nr**) – **double array**

If **mode** $= 0$ or 1, the reference vector that can be used in subsequent calls to g05lz with **mode** $= 2$.

4: **ifail – int32 scalar**

0 unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail $= 1$

On entry, $\mathbf{n} < 1$.

ifail $= 2$

On entry, $\mathbf{nr} < (\mathbf{n} + 1) \times (\mathbf{n} + 2)/2$.

ifail $= 3$

On entry, $\mathbf{ldc} < \mathbf{n}$.

ifail $= 4$

On entry, $\mathbf{mode} \neq 0$, 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 $\mathbf{mode} = 0$ 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 g05lz 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 $C + E = LL^{\mathrm{T}}$, 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

```
mode = int32(0);
xmu = [1;
      2];
c = [2, 1;
      1, 3];
igen = int32(1);
iseed = [int32(1762543);
      int32(9324783);
      int32(42344);
      int32(742355)];
```

```
r = [0;
     0;
     0;
     0;
     0;
     -0.04044723527840609];
[x, iseedOut, rOut, ifail] = g05lz(mode, xmu, c, igen, iseed, r)
```

```
x =
    3.9620
    2.4272
iseedOut =
     2332001
     5590614
    12141456
    12654846
rOut =
    2.5000
    1.0000
    2.0000
    1.4142
    0.7071
    1.5811
ifail =
           0
```