NAG Toolbox for MATLAB

g05eg

1 Purpose

g05eg sets up a reference vector for an autoregressive moving-average (ARMA) time series model with Normally distributed errors, so that g05ew may be used to generate successive terms. It also initializes the series to a stationary position.

2 Syntax

```
[r, var, ifail] = g05eg(e, a, b, nr, 'na', na, 'nb', nb)
```

3 Description

The ARMA model of such a time series in discrete time is

$$(x_n - E) = A_1(x_{n-1} - E) + \cdots + A_{\mathrm{NA}}(x_{n-NA} - E) + B_1\epsilon_n + \cdots + B_{\mathrm{NB}}\epsilon_{n-NB+1}$$

where x_n is the value of the series at time n, and ϵ_n is a series of independent random Standard Normal perturbations.

g05eg copies **a**, **e** and **b** to the reference vector so that g05ew can generate the terms of the series. It sets up initial values corresponding to a stationary position using the method described in Tunnicliffe–Wilson 1979.

4 References

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

Tunnicliffe–Wilson G 1979 Some efficient computational procedures for high order ARMA models *J. Statist. Comput. Simulation* **8** 301–309

5 Parameters

5.1 Compulsory Input Parameters

1: **e – double scalar**

The mean of the time series.

2: **a**$(*)$ **– double array**

Note: the dimension of the array **a** must be at least $\max(1, \mathbf{na})$.

The autoregressive coefficients of the model, if any.

3: **b**(**nb**) **– double array**

The moving-average coefficients of the model.

4: **nr – int32 scalar**

Constraint: $\mathbf{nr} \geq \mathbf{na} + \mathbf{nb} + 4 + \max(\mathbf{na}, \mathbf{nb})$.

5.2 Optional Input Parameters

1: **na – int32 scalar**

Default: The dimension of the array **a**.

The number of autoregressive coefficients supplied.

Constraint: $\mathbf{na} \geq 0$.

2: **nb – int32 scalar**

Default: The dimension of the array **b**.

the number of moving-average coefficients supplied.

Constraint: $\mathbf{nb} \geq 1$.

5.3 Input Parameters Omitted from the MATLAB Interface

None.

5.4 Output Parameters

1: **r(nr) – double array**

The reference vector and the recent history of the series.

2: **var – double scalar**

The proportion of the variance of a term in the series that is due to the moving-average (error) terms in the model. The smaller this is, the nearer is the model to non-stationarity.

3: **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:

$\mathbf{ifail} = 1$

On entry, $\mathbf{na} < 0$

$\mathbf{ifail} = 2$

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

$\mathbf{ifail} = 3$

On entry, $\mathbf{nr} < \mathbf{na} + \mathbf{nb} + 4 + \max(\mathbf{na}, \mathbf{nb})$.

$\mathbf{ifail} = 4$

The array **a** does not define a stationary autoregressive process.

7 Accuracy

The errors in the initialization process should be very much smaller than the error term; see Tunnicliffe–Wilson 1979.

8 Further Comments

The time taken by g05eg is essentially of order $(\mathbf{na})^2$.

Note: g05cb, g05cc, g05cf and g05cg must be used with care if this function is used as well. The reference vector, as mentioned before, contains a copy of the recent history of the series. This will not be altered properly by calls to any of the above functions. A call to g05cb or g05cc should be followed by calls to g05eg to re-initialize all time series reference vectors in use. To maintain repeatability with g05cb, the calls to g05eg should be performed in the same order and at the same point or points in the simulation every time g05cb is used. When functions g05cf and g05cg are used to save and restore the generator state, all the time series reference vectors in use must be saved and restored as well.

The ARMA model for a time series can also be written as:

$$(x_t - c) = \phi_1(x_{t-1} - c) + \cdots + \phi_p\left(x_{t-p} - c\right) + a_t - \theta_1 a_{t-1} \ldots - \theta_q a_{t-q}$$

where x_t is the observed value of the time series at time t,
p is the number of autoregressive parameters, ϕ_i,
q is the number of moving average parameters, θ_i,
c is the mean of the time series
and a_t is a series of independent random Normal perturbations with variance σ^2.

This is the form used in the G13 Chapter Introduction. This is related to the form given in Section 3 by:

$B_1^2 = \sigma^2$,

$B_{i+1} = -\theta_i\sigma = -\theta_i B_1, \qquad i = 1, 2, \ldots, q$,

$\mathbf{nb} = q + 1$,

$E = c$,

$A_i = \phi_i, \qquad i = 1, 2, \ldots, p$,

$\mathbf{na} = p$.

9 Example

```
e = 0;
a = [0.4;
     0.2];
b = [1];
nr = int32(9);
g05za('O');
g05cb(int32(0));
[r, var, ifail] = g05eg(e, a, b, nr)
```

```
r =
    2.5000
    1.5000
    9.5000
         0
    0.4000
    0.2000
    1.0000
    1.5236
    1.3092
var =
    0.7200
ifail =
           0
```