NAG Toolbox

nag_tsa_uni_arima_forcecast (g13aj)

1 Purpose

nag_tsa_uni_arima_forcecast (g13aj) applies a fully specified seasonal ARIMA model to an observed time series, generates the state set for forecasting and (optionally) derives a specified number of forecasts together with their standard deviations.

2 Syntax

```
[rms, st, nst, fva, fsd, isf, ifail] = nag_tsa_uni_arima_forcecast(mr, par, c,
kfc, x, ist, nfv, ifv, 'npar', npar, 'nx', nx)
```

```
[rms, st, nst, fva, fsd, isf, ifail] = g13aj(mr, par, c, kfc, x, ist, nfv, ifv,
'npar', npar, 'nx', nx)
```

3 Description

The time series $x_1, x_2, \ldots, x_n$ supplied to the function is assumed to follow a seasonal autoregressive integrated moving average (ARIMA) model with known parameters.

The model is defined by the following relations.

(a) $\nabla^d \nabla_s^D x_t - c = w_t$ where $\nabla^d \nabla_s^D x_t$ is the result of applying non-seasonal differencing of order d and seasonal differencing of seasonality s and order D to the series x_t, and c is a constant.

(b) $w_t = \Phi_1 w_{t-s} + \Phi_2 w_{t-2\times s} + \cdots + \Phi_P w_{t-P\times s} + e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2\times s} - \cdots - \Theta_Q e_{t-Q\times s}$.

This equation describes the seasonal structure with seasonal period s; in the absence of seasonality it reduces to $w_t = e_t$.

(c) $e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \cdots + \phi_p e_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$.

This equation describes the non-seasonal structure.

Given the series, the constant c, and the model parameters Φ, Θ, ϕ, θ, the function computes the following.

(a) The state set required for forecasting. This contains the minimum amount of information required for forecasting and comprises:

(i) the differenced series w_t, for $(N - s \times P) \leq t \leq N$;

(ii) the $(d + D \times s)$ values required to reconstitute the original series x_t from the differenced series w_t;

(iii) the intermediate series e_t, for $N - \max(p, Q \times s) < t \leq N$;

(iv) the residual series a_t, for $(N - q) < t \leq N$, where $N = n - (d + D \times s)$.

(b) A set of L forecasts of x_t and their estimated standard errors, s_t, for $t = n + 1, \ldots, n + L$ (L may be zero).

The forecasts and estimated standard errors are generated from the state set, and are identical to those that would be produced from the same state set by nag_tsa_uni_arima_forecast_state (g13ah).

Use of nag_tsa_uni_arima_forcecast (g13aj) should be confined to situations in which the state set for forecasting is unknown. Forecasting from the series requires recalculation of the state set and this is relatively expensive.

4 References

Box G E P and Jenkins G M (1976) *Time Series Analysis: Forecasting and Control* (Revised Edition) Holden–Day

5 Parameters

5.1 Compulsory Input Parameters

1: **mr**(**7**) – INTEGER array

The orders vector (p, d, q, P, D, Q, s) of the ARIMA model, in the usual notation.

Constraints:

$p, d, q, P, D, Q, s \geq 0$;
$p + q + P + Q > 0$;
$s \neq 1$;
if $s = 0$, $P + D + Q = 0$;
if $s > 1$, $P + D + Q > 0$;
$d + s \times (P + D) \leq n$;
$p + d - q + s \times (P + D - Q) \leq n$.

2: **par**(**npar**) – REAL (KIND=nag_wp) array

The p values of the ϕ parameters, the q values of the θ parameters, the P values of the Φ parameters, and the Q values of the Θ parameters, in that order.

3: **c** – REAL (KIND=nag_wp)

c, the expected value of the differenced series (i.e., c is the constant correction). Where there is no constant term, **c** must be set to 0.0.

4: **kfc** – INTEGER

Must be set to 0 if **c** was not estimated, and 1 if **c** was estimated. This is irrespective of whether or not $\mathbf{c} = 0.0$. The only effect is that the residual degrees of freedom are one greater when $\mathbf{kfc} = 0$. Assuming the supplied time series to be the same as that to which the model was originally fitted, this ensures an unbiased estimate of the residual mean-square.

Constraint: $\mathbf{kfc} = 0$ or 1.

5: **x**(**nx**) – REAL (KIND=nag_wp) array

The n values of the original undifferenced time series.

6: **ist** – INTEGER

The dimension of the array **st**.

Constraint: $\mathbf{ist} \geq (P \times s) + d + (D \times s) + q + \max(p, Q \times s)$. The expression on the right-hand side of the inequality is returned in **nst**.

7: **nfv** – INTEGER

The required number of forecasts. If $\mathbf{nfv} \leq 0$, no forecasts will be computed.

8: **ifv** – INTEGER

The dimension of the arrays **fva** and **fsd**.

Constraint: $\mathbf{ifv} \geq \max(1, \mathbf{nfv})$.

5.2 Optional Input Parameters

1: **npar** – INTEGER

Default: the dimension of the array **par**.

The exact number of ϕ, θ, Φ and Θ parameters.

Constraint: **npar** $= p + q + P + Q$.

2: **nx** – INTEGER

Default: the dimension of the array **x**.

n, the length of the original undifferenced time series.

5.3 Output Parameters

1: **rms** – REAL (KIND=nag_wp)

The residual variance (mean square) associated with the model.

2: **st**(**ist**) – REAL (KIND=nag_wp) array

The **nst** values of the state set.

3: **nst** – INTEGER

The number of values in the state set array **st**.

4: **fva**(**ifv**) – REAL (KIND=nag_wp) array

If **nfv** > 0, **fva** contains the **nfv** forecast values relating to the original undifferenced time series.

5: **fsd**(**ifv**) – REAL (KIND=nag_wp) array

If **nfv** > 0, **fsd** contains the estimated standard errors of the **nfv** forecast values.

6: **isf**(**4**) – INTEGER array

Contains validity indicators, one for each of the four possible parameter types in the model (autoregressive, moving average, seasonal autoregressive, seasonal moving average), in that order.

Each indicator has the interpretation:

-1 On entry the set of parameter values of this type does not satisfy the stationarity or invertibility test conditions.

0 No parameter of this type is in the model.

1 Valid parameter values of this type have been supplied.

7: **ifail** – INTEGER

ifail $= 0$ unless the function detects an error (see Section 5).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail $= 1$

On entry, **npar** $\neq p + q + P + Q$,
or the orders vector **mr** is invalid (check the constraints in Section 5),
or **kfc** $\neq 0$ or 1.

ifail $= 2$

On entry, **nx** $- d - D \times s \leq$ **npar** $+$ **kfc**, i.e., the number of terms in the differenced series is not greater than the number of parameters in the model. The model is over-parameterised.

ifail $= 3$

On entry, the workspace array w is too small.

ifail $= 4$

On entry, the state set array **st** is too small. It must be at least as large as the exit value of **nst**.

ifail $= 5$

This indicates a failure in nag_linsys_real_posdef_solve_1rhs (f04as) which is used to solve the equations giving estimates of the backforecasts.

ifail $= 6$

On entry, valid values were not supplied for all parameter types in the model. Inspect array **isf** for further information on the parameter type(s) in error.

ifail $= 7$

On entry, **ifv** $< \max(1,$ **nfv**$)$.

ifail $= -99$

An unexpected error has been triggered by this routine. Please contact NAG.

ifail $= -399$

Your licence key may have expired or may not have been installed correctly.

ifail $= -999$

Dynamic memory allocation failed.

7 Accuracy

The computations are believed to be stable.

8 Further Comments

The time taken by nag_tsa_uni_arima_forcecast (g13aj) is approximately proportional to n and the square of the number of backforecasts derived.

9 Example

The data is that used in the example program for nag_tsa_uni_arima_estim_easy (g13af). Five forecast values and their standard errors, together with the state set, are computed and printed.

9.1 Program Text

```
function g13aj_example

fprintf('g13aj example results\n\n');

% orders
mr  = [nag_int(1);  1;  2;  0;  0;  0;  0];
% Parameter estimates
par = [-0.0547;     -0.5568;     -0.6636];
```

```
% data
x = [-217; -177; -166; -136; -110;  -95;  -64;  -37;  -14;  -25;
      -51;  -62;  -73;  -88; -113; -120;  -83;  -33;  -19;   21;
       17;   44;   44;   78;   88;  122;  126;  114;   85;   64];

% From mr ...
ist = nag_int(4);

% Problem sizes
kfc = nag_int(1);
nfv = nag_int(5);
ifv = nag_int(nfv);
c   = 9.9807;

% Apply ARIMA model
[rms, st, nst, fva, fsd, isf, ifail] = ...
  g13aj( ...
         mr, par, c, kfc, x, ist, nfv, ifv);

%   Display results
fprintf('The residual mean square is %9.2f\n\n', rms);
fprintf('The state set consists of %4d values\n', nst);
for j = 1:6:nst
  fprintf('%11.4f', st(j:min(j+5,nst)));
  fprintf('\n');
end
fprintf('\nThe %4d forecast values and standard errors are -\n', nfv);
fprintf('%10.2f%10.2f\n', [fva fsd]');
```

9.2 Program Results

```
     g13aj example results

The residual mean square is    375.91

The state set consists of    4 values
    64.0000   -30.9807   -20.4495    -2.7212

The    5 forecast values and standard errors are -
     60.59     19.39
     69.50     34.99
     79.54     54.25
     89.51     67.87
     99.50     79.20
```