G13ADF – NAG Fortran Library Routine Document

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

G13ADF calculates preliminary estimates of the parameters of an autoregressive integrated moving average (ARIMA) model from the autocorrelation function of the appropriately differenced times series.

2 Specification

```
SUBROUTINE G13ADF(MR, R, NK, XV, NPAR, WA, NWA, PAR, RV, ISF, IFAIL)
INTEGER           MR(7), NK, NPAR, NWA, ISF(4), IFAIL
real              R(NK), XV, WA(NWA), PAR(NPAR), RV
```

3 Description

Preliminary estimates of the p non-seasonal autoregressive parameters $\phi_1, \phi_2, \ldots, \phi_p$ and the q non-seasonal moving average parameters $\theta_1, \theta_2, \ldots, \theta_q$ may be obtained from the sample autocorrelations relating to lags 1 to $p+q$, i.e., $r_1, \ldots, r_{p+q}$, of the differenced $\nabla^d \nabla_s^D x_t$ where x_t is assumed to follow a (possibly) seasonal ARIMA model (see Section 3 of the document for G13AEF for the specification of an ARIMA model).

Taking $r_0 = 1$ and $r_{-k} = r_k$, the ϕ_i, for $i = 1, 2, \ldots, p$ are the solutions to the equations

$$r_{q+i-1}\phi_1 + r_{q+i+2}\phi_2 + \ldots + r_{q+i-p}\phi_p = r_{q+i}, \quad \text{for } i = 1, 2, \ldots, p.$$

The θ_j, for $j = 1, 2, \ldots, q$ are obtained from the solutions to the equations

$$c_j = \tau_0\tau_j + \tau_1\tau_{j+1} + \ldots + \tau_{q+j}\tau_q, \quad \text{for } j = 0, 1, \ldots, q$$

(Cramer Wold-factorization) by setting

$$\theta_j = -\frac{\tau_j}{\tau_0}$$

where c_j are the 'covariances' modified in a 2-stage process by the autoregressive parameters.

Stage 1:

$$\begin{aligned} d_j &= r_j - \phi_1 r_{j-1} - \ldots - \phi_p r_{j-p}, && \text{for } j = 0, 1, \ldots, q; \\ d_j &= 0, && \text{for } j = q+1, q+2, \ldots, p+q. \end{aligned}$$

Stage 2:

$$c_j = d_j - \phi_1 d_{j+1} - \phi_2 d_{j+2} - \ldots - \phi_p d_{j+p}, \quad \text{for } j = 0, 1, \ldots, q.$$

The P seasonal autoregressive parameters $\Phi_1, \Phi_2, \ldots, \Phi_P$ and the Q seasonal moving average parameters $\Theta_1, \Theta_2, \ldots, \Theta_Q$ are estimated in the same way as the non-seasonal parameters, but each r_j is replaced in the calculation by $r_{s \times j}$, where s is the seasonal period.

An estimate of the residual variance is obtained by successively reducing the sample variance, first for non-seasonal, and then for seasonal, parameter estimates. If moving average parameters are estimated, the variance is reduced by a multiplying factor of τ_0^2, but otherwise by c_0.

4 References

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

5 Parameters

1: MR(7) — INTEGER array *Input*

On entry: the orders vector (p, d, q, P, D, Q, s) of the ARIMA model whose parameters are to be estimated. p, q, P and Q refer respectively to the number of autoregressive (ϕ), moving average (θ), seasonal autoregressive (Φ) and seasonal moving average (Θ) parameters. d, D and s refer respectively to the order of non-seasonal differencing, the order of seasonal differencing and the seasonal period.

Constraints:

$p, d, q, P, D, Q, s \geq 0$,
$p + q + P + Q > 0$,
$s \neq 1$,
if $s = 0$, then $P + D + Q = 0$,
if $s > 1$, then $P + D + Q > 0$.

2: R(NK) — ***real*** array *Input*

On entry: the autocorrelations (starting at lag 1), which must have been calculated after the time series has been appropriately differenced.

Constraint: $-1.0 \leq \mathrm{R}(i) \leq 1.0$, for $i = 1, 2, \ldots, \mathrm{NK}$.

3: NK — INTEGER *Input*

On entry: the maximum lag of the autocorrelations in array R.

Constraint: $\mathrm{NK} \geq \max(p + q, s \times (P + Q))$.

4: XV — ***real*** *Input*

On entry: the series sample variance, calculated after appropriate differencing has been applied to the series.

Constraint: $\mathrm{XV} > 0.0$.

5: NPAR — INTEGER *Input*

On entry: the exact number of parameters specified in the model by array MR.

Constraint: $\mathrm{NPAR} = p + q + P + Q$.

6: WA(NWA) — ***real*** array *Workspace*

7: NWA — INTEGER *Input*

On entry: the amount of workspace available.

Constraint: if $\mathrm{MR} = (p, d, q, P, D, Q, s)$ and $p' = \max(p, P)$ and $q' = \max(q, Q)$ then $\mathrm{NWA} \geq \max(p'^2 + p', 4(q' + 1))$.

8: PAR(NPAR) — ***real*** array *Output*

On exit: the first NPAR elements of PAR contain the preliminary estimates of the ARIMA model parameters, in standard order.

9: RV — ***real*** *Output*

On exit: an estimate of the residual variance of the preliminarily estimated model.

10: ISF(4) — INTEGER array *Output*

On exit: contains success/failure indicators, one for each of the 4 types of parameter (autoregressive, moving average, seasonal autoregressive, seasonal moving average).

The indicator has the interpretation:

0 no parameter of this type is in the model.

1 parameters of this type appear in the model and satisfactory preliminary estimates of this type were obtained.

−1 parameters of this type appear in the model but satisfactory preliminary estimates of this type were not obtainable. The estimates of this type of parameter were set to 0.0 in array PAR.

11: IFAIL — INTEGER *Input/Output*

On entry: IFAIL must be set to 0, −1 or 1. For users not familiar with this parameter (described in Chapter P01) the recommended value is 0.

On exit: IFAIL = 0 unless the routine detects an error (see Section 6).

6 Error Indicators and Warnings

Errors detected by the routine:

IFAIL = 1

On entry, the orders vector MR is invalid. One of the constraints in Section 5 has been violated.

IFAIL = 2

On entry, NK $< \max(p+q, s\times(P+Q))$. There are not enough autocorrelations to enable the required model to be estimated.

IFAIL = 3

On entry, at least one element of R lies outside the range $[-1.0, 1.0]$.

IFAIL = 4

On entry, XV ≤ 0.0.

IFAIL = 5

On entry, NPAR $\neq p+q+P+Q$.

IFAIL = 6

On entry, the workspace array WA is too small. See Section 5 for the minimum size formula.

IFAIL = 7

Satisfactory parameter estimates could not be obtained for all parameter types in the model. Inspect array ISF for indicators of the parameter type(s) which could not be estimated.

7 Accuracy

The performance of the algorithm is conditioned by the roots of the autoregressive and moving average operators. If these are not close to unity in modulus, the errors, e, should satisfy $e < 100\epsilon$ where ϵ is ***machine precision***.

8 Further Comments

The time taken by the routine is approximately proportional to $(p^3+q^2+P^3+Q^2)$.

9 Example

This example reads the sample autocorrelations to lag 40 and the sample variance of the lagged and doubly differenced series of airline passenger totals (Box and Jenkins example series G [1]). Preliminary estimates of the parameters of the (0,1,1,0,1,1,12) model are obtained by a call to G13ADF.

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.

```
*     G13ADF Example Program Text
*     Mark 14 Revised.  NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          NPMAX, NWA, NLMAX
      PARAMETER        (NPMAX=10,NWA=200,NLMAX=50)
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
*     .. Local Scalars ..
      real             RV, YV
      INTEGER          I, IFAIL, NL, NPAR
*     .. Local Arrays ..
      real             PAR(NPMAX), R(NLMAX), WA(NWA)
      INTEGER          ISF(4), MR(7)
*     .. External Subroutines ..
      EXTERNAL         G13ADF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G13ADF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) NL
      READ (NIN,*) YV
      WRITE (NOUT,*)
      IF (NL.GT.0 .AND. NL.LE.NLMAX) THEN
         READ (NIN,*) (R(I),I=1,NL)
         READ (NIN,*) MR
         NPAR = MR(1) + MR(3) + MR(4) + MR(6)
         IF (NL.GT.0 .AND. NPAR.LE.NPMAX) THEN
            IFAIL = 1
*
            CALL G13ADF(MR,R,NL,YV,NPAR,WA,NWA,PAR,RV,ISF,IFAIL)
*
            IF (IFAIL.NE.0) THEN
               WRITE (NOUT,99999) 'G13ADF fails. IFAIL = ', IFAIL
               WRITE (NOUT,*)
            END IF
            IF (IFAIL.EQ.0 .OR. IFAIL.GE.7) THEN
               WRITE (NOUT,99998)
     +           'Parameter estimation success/failure indicator',
     +           (ISF(I),I=1,4)
               WRITE (NOUT,*)
               WRITE (NOUT,99997) 'ARIMA model parameter values ',
     +           (PAR(I),I=1,NPAR)
               WRITE (NOUT,*)
               WRITE (NOUT,99997) 'Residual variance', RV
            END IF
         END IF
      END IF
```

```
      STOP
*
99999 FORMAT (1X,A,I1)
99998 FORMAT (1X,A,4I4)
99997 FORMAT (1X,A,5F10.5)
      END
```

9.2 Program Data

```
G13ADF Example Program Data
  40 200
  0.00213
 -0.32804   0.09850  -0.21854   0.05585   0.04679   0.04135
 -0.07989   0.00335   0.13973  -0.04022   0.07618  -0.40583
  0.18239  -0.05057   0.16094  -0.15900   0.09152  -0.03474
  0.05195  -0.14417   0.04264  -0.08170   0.23389  -0.02828
 -0.09001   0.03050  -0.02046   0.05522  -0.02048  -0.06651
 -0.02940   0.20204  -0.13953   0.10098  -0.20849   0.03338
  0.00829   0.07082  -0.04457  -0.01216
  0  1  1  0  1  1 12
```

9.3 Program Results

```
 G13ADF Example Program Results

 Parameter estimation success/failure indicator   0   1   0   1

 ARIMA model parameter values    0.37390   0.51237

 Residual variance   0.00148
```