NAG Toolbox for MATLAB

g13as

1 Purpose

g13as is a diagnostic checking function suitable for use after fitting a Box–Jenkins ARMA model to a univariate time series using g13ae or g13af. The residual autocorrelation function is returned along with an estimate of its asymptotic standard errors and correlations. Also, g13as calculates the Box–Ljung portmanteau statistic and its significance level for testing model adequacy.

2 Syntax

```
[r, rcm, chi, idf, siglev, ifail] = g13as(v, mr, m, par, ishow, 'n', n,
'npar', npar)
```

3 Description

Consider the univariate multiplicative autoregressive-moving average model

$$\phi(B)\Phi(B^s)(W_t - \mu) = \theta(B)\Theta(B^s)\epsilon_t \quad (1)$$

where W_t, for $t = 1, 2, \ldots, n$, denotes a time series and ϵ_t, for $t = 1, 2, \ldots, n$, is a residual series assumed to be normally distributed with zero mean and variance σ^2 (> 0). The ϵ_t's are also assumed to be uncorrelated. Here μ is the overall mean term, s is the seasonal period and B is the backward shift operator such that $B^r W_t = W_{t-r}$. The polynomials in (1) are defined as follows:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$$

is the non-seasonal autoregressive (AR) operator;

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

is the non-seasonal moving average (MA) operator;

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps}$$

is the seasonal AR operator; and

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs}$$

is the seasonal MA operator. The model (1) is assumed to be stationary, that is the zeros of $\phi(B)$ and $\Phi(B^s)$ are assumed to lie outside the unit circle. The model (1) is also assumed to be invertible, that is the zeros of $\theta(B)$ and $\Theta(B^s)$ are assumed to lie outside the unit circle. When both $\Phi(B^s)$ and $\Theta(B^s)$ are absent from the model, that is when $P = Q = 0$, then the model is said to be non-seasonal.

The estimated residual autocorrelation coefficient at lag l, $\hat{r}_l$, is computed as:

$$\hat{r}_l = \frac{\sum_{t=l+1}^{n} (\hat{\epsilon}_{t-l} - \bar{\epsilon})(\hat{\epsilon}_t - \bar{\epsilon})}{\sum_{t=1}^{n} (\hat{\epsilon}_t - \bar{\epsilon})^2}, \qquad l = 1, 2, \ldots$$

where $\hat{\epsilon}_t$ denotes an estimate of the tth residual, ϵ_t, and $\bar{\epsilon} = \sum_{t=1}^{n} \hat{\epsilon}_t / n$. A portmanteau statistic, $Q_{(m)}$, is calculated from the formula (see Box and Ljung 1978):

$$Q_{(m)} = n(n+2) \sum_{l=1}^{m} \hat{r}_l^2 / (n - l)$$

where m denotes the number of residual autocorrelations computed. (Advice on the choice of m is given in Section 8.2.) Under the hypothesis of model adequacy, $Q_{(m)}$ has an asymptotic χ^2-distribution on $m - p - q - P - Q$ degrees of freedom. Let $\hat{r}^{\mathrm{T}} = (\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_m)$ then the variance-covariance matrix of $\hat{r}$ is given by:

$$\mathrm{Var}(\hat{r}) = \left[I_m - X\left(X^{\mathrm{T}}X\right)^{-1}X^{\mathrm{T}}\right]/n.$$

The construction of the matrix X is discussed in McLeod 1978. (Note that the mean, μ, and the residual variance, σ^2, play no part in calculating $\mathrm{Var}(\hat{r})$ and therefore are not required as input to g13as.)

Note: for additive models with fixed parameter values (i.e., fitted by g13dc) g13ds should be used instead of g13as.

4 References

Box G E P and Ljung G M 1978 On a measure of lack of fit in time series models *Biometrika* **65** 297–303

McLeod A I 1978 On the distribution of the residual autocorrelations in Box–Jenkins models *J. Roy. Statist. Soc. Ser. B* **40** 296–302

5 Parameters

5.1 Compulsory Input Parameters

1: **v(n) – double array**

$\mathbf{v}(t)$ must contain an estimate of ϵ_t, for $t = 1, 2, \ldots, n$.

If g13as is used following a call to g13ae then the actual parameter **v** must be **exr**(**icount**$(1) + 1$) as returned by g13ae.

If g13as is used following a call to g13af then the actual parameter **v** must be **res** as returned by g13af.

Constraint: **v** must contain at least two distinct elements.

2: **mr(7) – int32 array**

The orders vector (p, d, q, P, D, Q, s) as supplied to g13ae or g13af.

Constraints:

$p, q, P, Q, s \geq 0$;
$p + q + P + Q > 0$;
if $s = 0$, then $P = 0$ and $Q = 0$.

3: **m – int32 scalar**

the value of m, the number of residual autocorrelations to be computed. See Section 8.2 for advice on the value of **m**.

Constraint: **npar** < **m** < **n**.

4: **par(npar) – double array**

The parameter estimates in the order $\phi_1, \phi_2, \ldots, \phi_p$, $\theta_1, \theta_2, \ldots, \theta_q$, $\Phi_1, \Phi_2, \ldots, \Phi_P$, $\Theta_1, \Theta_2, \ldots, \Theta_Q$ only.

Constraint: the elements in **par** must satisfy the stationarity and invertibility conditions.

5: **ishow – int32 scalar**

Must be nonzero if the residual autocorrelations, their standard errors and the portmanteau statistics are to be printed and zero otherwise.

These quantities are available also as output variables in **r**, **rcm**, **chi**, **idf** and **siglev**.

5.2 Optional Input Parameters

1: **n – int32 scalar**

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

n, the number of observations in the residual series.

If g13as is used following a call to g13ae, then **n** must be the value **icount**(2) returned by g13ae.

If g13as is used following a call to g13af, then **n** must be the value **nres** returned by g13af.

Constraint: $\mathbf{n} \geq 3$.

2: **npar – int32 scalar**

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

The total number of ϕ, θ, Φ and Θ parameters, i.e., $\mathbf{npar} = p + q + P + Q$.

Constraint: $\mathbf{npar} = \mathbf{mr}(1) + \mathbf{mr}(3) + \mathbf{mr}(4) + \mathbf{mr}(6)$.

5.3 Input Parameters Omitted from the MATLAB Interface

ldrcm, iw, liw, work, lwork

5.4 Output Parameters

1: **r(m) – double array**

An estimate of the residual autocorrelation coefficient at lag l, for $l = 1, 2, \ldots, m$. If $\mathbf{ifail} = 3$ on exit then all elements of **r** are set to zero.

2: **rcm(ldrcm,m) – double array**

The estimated standard errors and correlations of the elements in the array **r**. The correlation between $\mathbf{r}(i)$ and $\mathbf{r}(j)$ is returned as $\mathbf{rcm}(ij)$ except that if $i = j$ then $\mathbf{rcm}(i,j)$ contains the standard error of $\mathbf{r}(i)$. If on exit, $\mathbf{ifail} \geq 5$, then all off-diagonal elements of **rcm** are set to zero and all diagonal elements are set to $1/\sqrt{n}$.

3: **chi – double scalar**

The value of the portmanteau statistic, $Q_{(m)}$. If $\mathbf{ifail} = 3$ on exit then **chi** is returned as zero.

4: **idf – int32 scalar**

The number of degrees of freedom of **chi**.

5: **siglev – double scalar**

The significance level of **chi** based on **idf** degrees of freedom. If $\mathbf{ifail} = 3$ on exit, **siglev** is returned as one.

6: **ifail – int32 scalar**

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

6 Error Indicators and Warnings

Note: g13as may return useful information for one or more of the following detected errors or warnings.

$\mathbf{ifail} = 1$

On entry, $\mathbf{mr}(1) < 0$,
or $\mathbf{mr}(3) < 0$,

or **mr**$(4) < 0$,
or **mr**$(6) < 0$,
or **mr**$(7) < 0$,
or **mr**$(7) = 0$ and either **mr**$(4) > 0$ or **mr**$(6) > 0$,
or **mr**$(1) =$ **mr**$(3) =$ **mr**$(4) =$ **mr**$(6) = 0$,
or **m** $\leq$ **npar**,
or **m** $\geq$ **n**,
or **n** < 3,
or **npar** $\neq$ **mr**$(1) +$ **mr**$(3) +$ **mr**$(4) +$ **mr**(6),
or **ldrcm** $<$ **m**,
or **liw** is too small,
or **lwork** is too small.

ifail = 2

On entry, the autoregressive (or moving average) parameters are extremely close to or outside the stationarity (or invertibility) region. To proceed, you must supply different parameter estimates in the array **par**.

ifail = 3

On entry, the residuals are practically identical giving zero (or near zero) variance. In this case **chi** is set to zero and **siglev** to one and all the elements of **r** are set to zero.

ifail = 4

This is an unlikely exit brought about by an excessive number of iterations being needed to evaluate the zeros of the AR or MA polynomials. All output parameters are undefined.

ifail = 5

On entry, one or more of the AR operators has a factor in common with one or more of the MA operators. To proceed, this common factor must be deleted from the model. In this case, the off-diagonal elements of **rcm** are returned as zero and the diagonal elements set to $1/\sqrt{n}$. All other output quantities will be correct.

ifail = 6

This is an unlikely exit. At least one of the diagonal elements of **rcm** was found to be either negative or zero. In this case all off-diagonal elements of **rcm** are returned as zero and all diagonal elements of **rcm** set to $1/\sqrt{n}$.

7 Accuracy

The computations are believed to be stable.

8 Further Comments

8.1 Timing

The time taken by g13as depends upon the number of residual autocorrelations to be computed, m.

8.2 Choice of *m*

The number of residual autocorrelations to be computed, m should be chosen to ensure that when the ARMA model (1) is written as either an infinite order autoregressive process:

$$W_t - \mu = \sum_{j=1}^{\infty} \pi_j \left(W_{t-j} - \mu\right) + \epsilon_t$$

or as an infinite order moving average process:

$$W_t - \mu = \sum_{j=1}^{\infty} \psi_j \epsilon_{t-j} + \epsilon_t$$

then the two sequences $\{\pi_1, \pi_2, \ldots\}$ and $\{\psi_1, \psi_2, \ldots\}$ are such that π_j and ψ_j are approximately zero for $j > m$. An over-estimate of m is therefore preferable to an under-estimate of m. In many instances the choice $m = 10$ will suffice. In practice, to be on the safe side, you should try setting $m = 20$.

8.3 Approximate Standard Errors

When **ifail** $= 5$ or 6 all the standard errors in **rcm** are set to $1/\sqrt{n}$. This is the asymptotic standard error of $\hat{r}_l$ when all the autoregressive and moving average parameters are assumed to be known rather than estimated.

8.4 Alternative Applications

g13as may be used for diagnostic checking of suitable univariate ARMA models, as described in Section 3, fitted by g13be or g13dc. Great care must be taken in obtaining the input values for g13as from the output values from g13be or g13dc.

9 Example

```
v = [19.62746933753217;
     -5.309287748702281;
     9.798293298952524;
     15.24123763096101;
     -9.169280254619032;
     16.11067630246432;
     15.39291433217872;
     -5.450004312837414;
     -27.6205284052683;
     -18.13061645026666;
     5.720197603730954;
     -13.08812421913039;
     -22.71514540998739;
     -14.92555865550874;
     4.693009266513346;
     33.54056385487841;
     19.71375888111006;
     -27.33603163811242;
     32.12310786592178;
     -11.76808232386082;
     1.152380086683287;
     -1.775570325352465;
     23.68209327485152;
     -10.62375411730706;
     13.96192179999122;
     -5.272654385984258;
     -28.78678842222793;
     -20.65731709475202;
     -2.255525918186007];
mr = [int32(1);
     int32(1);
     int32(2);
     int32(0);
     int32(0);
     int32(0);
     int32(0)];
m = int32(10);
par = [-0.05429075588805302;
     -0.5547824600332715;
     -0.6734171925737445];
ishow = int32(1);
```

```
[r, rcm, chi, idf, siglev, ifail] = g13as(v, mr, m, par, ishow)
```

```
RESIDUAL AUTOCORRELATION FUNCTION
 ---------------------------------

 LAG   K       1      2      3      4      5      6      7
 R(K)       0.020 -0.040 -0.019  0.068 -0.143 -0.046 -0.205
 ST.ERROR  0.007  0.125  0.128  0.150  0.168  0.168  0.178
 ----------------------------------------------------------
 LAG   K       8      9     10
 R(K)      -0.108 -0.001 -0.058
 ST.ERROR  0.179  0.181  0.183
 ----------------------------------------------------------

 BOX - LJUNG PORTMANTEAU STATISTIC =      3.465
               SIGNIFICANCE LEVEL =      0.839
 (BASED ON   7 DEGREES OF FREEDOM)

 VALUE OF IFAIL PARAMETER ON EXIT FROM G13ASF =   0
r =
    0.0199
   -0.0401
   -0.0191
    0.0683
   -0.1427
   -0.0456
   -0.2048
   -0.1082
   -0.0007
   -0.0581
rcm =
  Columns 1 through 7
    0.0067    0.9986    0.2512    0.5337   -0.4515   -0.0672    0.3230
    0.9986    0.1252    0.3019    0.5451   -0.4251   -0.0920    0.3193
    0.2512    0.3019    0.1277    0.3989    0.3692   -0.4447   -0.0019
    0.5337    0.5451    0.3989    0.1503    0.0587    0.2843   -0.1867
   -0.4515   -0.4251    0.3692    0.0587    0.1683    0.0854    0.0939
   -0.0672   -0.0920   -0.4447    0.2843    0.0854    0.1682    0.0602
    0.3230    0.3193   -0.0019   -0.1867    0.0939    0.0602    0.1777
   -0.1352   -0.1175    0.2821   -0.0770   -0.1056    0.1049    0.0123
   -0.1394   -0.1466   -0.1538    0.1658   -0.0041   -0.0975    0.0538
    0.1652    0.1576   -0.1015   -0.0401    0.0717   -0.0157   -0.0375
  Columns 8 through 10
   -0.1352   -0.1394    0.1652
   -0.1175   -0.1466    0.1576
    0.2821   -0.1538   -0.1015
   -0.0770    0.1658   -0.0401
   -0.1056   -0.0041    0.0717
    0.1049   -0.0975   -0.0157
    0.0123    0.0538   -0.0375
    0.1793    0.0320    0.0305
    0.0320    0.1809    0.0080
    0.0305    0.0080    0.1835
chi =
    3.4654
idf =
           7
siglev =
    0.8389
ifail =
           0
```