

NAG Toolbox

nag_tsa_multi_varma_diag (g13ds)

1 Purpose

`nag_tsa_multi_varma_diag` (g13ds) is a diagnostic checking function suitable for use after fitting a vector ARMA model to a multivariate time series using `nag_tsa_multi_varma_estimate` (g13dd). The residual cross-correlation matrices are returned along with an estimate of their asymptotic standard errors and correlations. Also, `nag_tsa_multi_varma_diag` (g13ds) calculates the modified Li–McLeod portmanteau statistic and its significance level for testing model adequacy.

2 Syntax

```
[qq, r0, r, rcm, chi, idf, siglev, ifail] = nag_tsa_multi_varma_diag(v, ip, iq, m, par, parhld, qq, ishow, 'k', k, 'n', n)
```

```
[qq, r0, r, rcm, chi, idf, siglev, ifail] = g13ds(v, ip, iq, m, par, parhld, qq, ishow, 'k', k, 'n', n)
```

3 Description

Let $W_t = (w_{1t}, w_{2t}, \dots, w_{kt})^T$, for $t = 1, 2, \dots, n$, denote a vector of k time series which is assumed to follow a multivariate ARMA model of the form

$$W_t - \mu = \phi_1(W_{t-1} - \mu) + \phi_2(W_{t-2} - \mu) + \dots + \phi_p(W_{t-p} - \mu) + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \dots - \theta_q\epsilon_{t-q}, \quad (1)$$

where $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \dots, \epsilon_{kt})^T$, for $t = 1, 2, \dots, n$, is a vector of k residual series assumed to be Normally distributed with zero mean and positive definite covariance matrix Σ . The components of ϵ_t are assumed to be uncorrelated at non-simultaneous lags. The ϕ_i and θ_j are k by k matrices of parameters. $\{\phi_i\}$, for $i = 1, 2, \dots, p$, are called the autoregressive (AR) parameter matrices, and $\{\theta_i\}$, for $i = 1, 2, \dots, q$, the moving average (MA) parameter matrices. The parameters in the model are thus the p (k by k) ϕ -matrices, the q (k by k) θ -matrices, the mean vector μ and the residual error covariance matrix Σ . Let

$$A(\phi) = \begin{bmatrix} \phi_1 & I & 0 & \cdot & \cdot & \cdot & 0 \\ \phi_2 & 0 & I & 0 & \cdot & \cdot & 0 \\ \cdot & & & \cdot & & & \\ \cdot & & & & & & \\ \phi_{p-1} & 0 & \cdot & \cdot & \cdot & 0 & I \\ \phi_p & 0 & \cdot & \cdot & \cdot & 0 & 0 \end{bmatrix}_{pk \times pk} \quad \text{and} \quad B(\theta) = \begin{bmatrix} \theta_1 & I & 0 & \cdot & \cdot & \cdot & 0 \\ \theta_2 & 0 & I & 0 & \cdot & \cdot & 0 \\ \cdot & & & \cdot & & & \\ \cdot & & & & & & \\ \theta_{q-1} & 0 & \cdot & \cdot & \cdot & \cdot & I \\ \theta_q & 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{bmatrix}_{qk \times qk}$$

where I denotes the k by k identity matrix.

The ARMA model (1) is said to be stationary if the eigenvalues of $A(\phi)$ lie inside the unit circle, and invertible if the eigenvalues of $B(\theta)$ lie inside the unit circle. The ARMA model is assumed to be both stationary and invertible. Note that some of the elements of the ϕ - and/or θ -matrices may have been fixed at pre-specified values (for example by calling `nag_tsa_multi_varma_estimate` (g13dd)).

The estimated residual cross-correlation matrix at lag l is defined to the k by k matrix \hat{R}_l whose (i, j) th element is computed as

$$\hat{r}_{ij}(l) = \frac{\sum_{t=l+1}^n (\hat{\epsilon}_{it-l} - \bar{\epsilon}_i)(\hat{\epsilon}_{jt} - \bar{\epsilon}_j)}{\sqrt{\sum_{t=1}^n (\hat{\epsilon}_{it} - \bar{\epsilon}_i)^2 \sum_{t=1}^n (\hat{\epsilon}_{jt} - \bar{\epsilon}_j)^2}}, \quad l = 0, 1, \dots, i \text{ and } j = 1, 2, \dots, k,$$

where $\hat{\epsilon}_{it}$ denotes an estimate of the t th residual for the i th series ϵ_{it} and $\bar{\epsilon}_i = \sum_{t=1}^n \hat{\epsilon}_{it}/n$. (Note that \hat{R}_l is an estimate of $E(\epsilon_{t-l}\epsilon_t^T)$, where E is the expected value.)

A modified portmanteau statistic, $Q_{(m)}^*$, is calculated from the formula (see Li and McLeod (1981))

$$Q_{(m)}^* = \frac{k^2 m(m+1)}{2n} + n \sum_{l=1}^m \hat{r}(l)^T (\hat{R}_0^{-1} \otimes \hat{R}_0^{-1}) \hat{r}(l),$$

where \otimes denotes Kronecker product, \hat{R}_0 is the estimated residual cross-correlation matrix at lag zero and $\hat{r}(l) = \text{vec}(\hat{R}_l^T)$, where vec of a k by k matrix is a vector with the (i, j) th element in position $(i-1)k + j$. m denotes the number of residual cross-correlation matrices computed. (Advice on the choice of m is given in Section 9.2.) Let l_C denote the total number of ‘free’ parameters in the ARMA model excluding the mean, μ , and the residual error covariance matrix Σ . Then, under the hypothesis of model adequacy, $Q_{(m)}^*$, has an asymptotic χ^2 -distribution on $mk^2 - l_C$ degrees of freedom.

Let $\hat{x} = (\text{vec}(R_1^T), \text{vec}(R_2^T), \dots, \text{vec}(R_m^T))$ then the covariance matrix of \hat{x} is given by

$$\text{Var}(\hat{x}) = [Y - X(X^T G G^T X)^{-1} X^T] / n,$$

where $Y = I_m \otimes (\Delta \otimes \Delta)$ and $G = I_m (G G^T)$. Δ is the dispersion matrix Σ in correlation form and G a nonsingular k by k matrix such that $G G^T = \Delta^{-1}$ and $G \Delta G^T = I_k$. The construction of the matrix X is discussed in Li and McLeod (1981). (Note that the mean, μ , plays no part in calculating $\text{Var}(\hat{x})$ and therefore is not required as input to `nag_tsa_multi_varma_diag` (g13ds).)

4 References

Li W K and McLeod A I (1981) Distribution of the residual autocorrelations in multivariate ARMA time series models *J. Roy. Statist. Soc. Ser. B* **43** 231–239

5 Parameters

The output quantities **k**, **n**, **v**, **kmax**, **ip**, **iq**, **par**, **parhld** and **qq** from `nag_tsa_multi_varma_estimate` (g13dd) are suitable for input to `nag_tsa_multi_varma_diag` (g13ds).

5.1 Compulsory Input Parameters

1: **v**(*kmax*, **n**) – REAL (KIND=nag_wp) array

kmax, the first dimension of the array, must satisfy the constraint $kmax \geq \mathbf{k}$.

v(*i*, *t*) must contain an estimate of the i th component of ϵ_t , for $i = 1, 2, \dots, k$ and $t = 1, 2, \dots, n$.

Constraints:

no two rows of **v** may be identical;

in each row there must be at least two distinct elements.

2: **ip** – INTEGER

p, the number of AR parameter matrices.

Constraint: **ip** ≥ 0 .

3: **iq** – INTEGER

q, the number of MA parameter matrices.

Constraint: **iq** ≥ 0 .

Note: **ip** = **iq** = 0 is **not permitted**.

4: **m** – INTEGER

The value of m , the number of residual cross-correlation matrices to be computed. See Section 9.2 for advice on the choice of **m**.

Constraint: **ip** + **iq** < **m** < **n**.

5: **par**((**ip** + **iq**) × **k** × **k**) – REAL (KIND=nag_wp) array

The parameter estimates read in row by row in the order $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$.

Thus,

if **ip** > 0, **par**(($l-1$) × $k \times k + (i-1) \times k + j$) must be set equal to an estimate of the (i, j)th element of ϕ_l , for $l = 1, 2, \dots, p$ and $i = 1, 2, \dots, k$;

if **iq** ≥ 0, **par**($p \times k \times k + (l-1) \times k \times k + (i-1) \times k + j$) must be set equal to an estimate of the (i, j)th element of θ_l , for $l = 1, 2, \dots, q$ and $i = 1, 2, \dots, k$.

The first $p \times k \times k$ elements of **par** must satisfy the stationarity condition and the next $q \times k \times k$ elements of **par** must satisfy the invertibility condition.

6: **parhld**((**ip** + **iq**) × **k** × **k**) – LOGICAL array

parhld(i) must be set to *true* if **par**(i) has been held constant at a pre-specified value and *false* if **par**(i) is a free parameter, for $i = 1, 2, \dots, (p+q) \times k \times k$.

7: **qq**($kmax, k$) – REAL (KIND=nag_wp) array

$kmax$, the first dimension of the array, must satisfy the constraint $kmax \geq k$.

qq(i, j) is an efficient estimate of the (i, j)th element of Σ . The lower triangle only is needed.

Constraint: **qq** must be positive definite.

8: **ishow** – INTEGER

Must be nonzero if the residual cross-correlation matrices $\{\hat{r}_{ij}(l)\}$ and their standard errors $\{se(\hat{r}_{ij}(l))\}$, the modified portmanteau statistic with its significance and a summary table are to be printed. The summary table indicates which elements of the residual correlation matrices are significant at the 5% level in either a positive or negative direction; i.e., if $\hat{r}_{ij}(l) > 1.96 \times se(\hat{r}_{ij}(l))$ then a '+' is printed, if $\hat{r}_{ij}(l) < -1.96 \times se(\hat{r}_{ij}(l))$ then a '-' is printed, otherwise a fullstop (.) is printed. The summary table is only printed if $k \leq 6$ on entry.

The residual cross-correlation matrices, their standard errors and the modified portmanteau statistic with its significance are available also as output variables in **r**, **rcm**, **chi**, **idf** and **siglev**.

5.2 Optional Input Parameters

1: **k** – INTEGER

Default: the first dimension of the arrays **v**, **qq** and the second dimension of the array **qq**. (An error is raised if these dimensions are not equal.)

k , the number of residual time series.

Constraint: **k** ≥ 1.

2: **n** – INTEGER

Default: the second dimension of the array **v**.

n , the number of observations in each residual series.

5.3 Output Parameters

- 1: **qq**(*kmax*, **k**) – REAL (KIND=nag_wp) array
If **ifail** \neq 1, then the upper triangle is set equal to the lower triangle.
- 2: **r0**(*kmax*, **k**) – REAL (KIND=nag_wp) array
If $i \neq j$, then **r0**(*i*, *j*) contains an estimate of the (*i*, *j*)th element of the residual cross-correlation matrix at lag zero, \hat{R}_0 . When $i = j$, **r0**(*i*, *j*) contains the standard deviation of the *i*th residual series. If **ifail** = 3 on exit then the first **k** rows and columns of **r0** are set to zero.
- 3: **r**(*kmax*, *kmax*, **m**) – REAL (KIND=nag_wp) array
r(*l*, *i*, *j*) is an estimate of the (*i*, *j*)th element of the residual cross-correlation matrix at lag *l*, for $i = 1, 2, \dots, k$, $j = 1, 2, \dots, k$ and $l = 1, 2, \dots, m$. If **ifail** = 3 on exit then all elements of **r** are set to zero.
- 4: **rcm**(*ldrcm*, **m** \times **k** \times **k**) – REAL (KIND=nag_wp) array
The estimated standard errors and correlations of the elements in the array **r**. The correlation between **r**(*l*, *i*, *j*) and **r**(*l*₂, *i*₂, *j*₂) is returned as **rcm**(*s*, *t*) where $s = (l - 1) \times k \times k + (j - 1) \times k + i$ and $t = (l_2 - 1) \times k \times k + (j_2 - 1) \times k + i_2$ except that if $s = t$, then **rcm**(*s*, *t*) contains the standard error of **r**(*l*, *i*, *j*). If on exit, **ifail** \geq 5, then all off-diagonal elements of **rcm** are set to zero and all diagonal elements are set to $1/\sqrt{n}$.
- 5: **chi** – REAL (KIND=nag_wp)
The value of the modified portmanteau statistic, $Q_{(m)}^*$. If **ifail** = 3 on exit then **chi** is returned as zero.
- 6: **idf** – INTEGER
The number of degrees of freedom of **chi**.
- 7: **siglev** – REAL (KIND=nag_wp)
The significance level of **chi** based on **idf** degrees of freedom. If **ifail** = 3 on exit, **siglev** is returned as one.
- 8: **ifail** – INTEGER
ifail = 0 unless the function detects an error (see Section 5).

6 Error Indicators and Warnings

Note: nag_tsa_multi_varma_diag (g13ds) may return useful information for one or more of the following detected errors or warnings.

Errors or warnings detected by the function:

ifail = 1

- On entry, **k** < 1,
- or $kmax < k$,
- or **ip** < 0,
- or **iq** < 0,
- or **ip** = **iq** = 0,
- or **m** \leq **ip** + **iq**,
- or **m** \geq **n**,
- or $ldrcm < m \times k \times k$,
- or *liw* is too small,

or $lwork$ is too small.

ifail = 2

On entry, either **qq** is not positive definite or the autoregressive parameter matrices are extremely close to or outside the stationarity region, or the moving average parameter matrices are extremely close to or outside the invertibility region. To proceed, you must supply different parameter estimates in the arrays **par** and **qq**.

ifail = 3 (*warning*)

On entry, at least one of the k residual series is such that all its elements are practically identical giving zero (or near zero) variance or at least two of the residual series are identical. In this case **chi** is set to zero, **siglev** to one and all the elements of **r0** and **r** 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 determinantal polynomials $\det(A(\phi))$ and $\det(B(\theta))$. All output arguments are undefined.

ifail = 5

On entry, either the eigenvalues and eigenvectors of Δ (the matrix **qq** in correlation form) could not be computed or the determinantal polynomials $\det(A(\phi))$ and $\det(B(\theta))$ have a factor in common. To proceed, you must either supply different parameter estimates in the array **qq** or delete this common factor 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 $1/\sqrt{n}$.

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

8.1 Timing

The time taken by `nag_tsa_multi_varma_diag` (g13ds) depends upon the number of residual cross-correlation matrices to be computed, m , and the number of time series, k .

8.2 Choice of m

The number of residual cross-correlation matrices to be computed, m , should be chosen to ensure that when the ARMA model (1) is written as either an infinite order autoregressive process, i.e.,

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

or as an infinite order moving average process, i.e.,

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

then the two sequences of k by k matrices $\{\pi_1, \pi_2, \dots\}$ and $\{\psi_1, \psi_2, \dots\}$ are such that π_j and ψ_j are approximately zero for $j > m$. An overestimate 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 Checking a ‘White Noise’ Model

If you have fitted the ‘white noise’ model

$$W_t - \mu = \epsilon_t$$

then `nag_tsa_multi_varma_diag` (g13ds) should be entered with $p = 1$, $q = 0$, and the first k^2 elements of `par` and `parhld` set to zero and `true` respectively.

8.4 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 $\hat{r}_{ij}(l)$ when all the autoregressive and moving average parameters are assumed to be known rather than estimated.

8.5 Alternative Tests

\hat{R}_0 is useful in testing for instantaneous causality. If you wish to carry out a likelihood ratio test then the covariance matrix at lag zero (\hat{C}_0) can be used. It can be recovered from \hat{R}_0 by setting

$$\begin{aligned} \hat{C}_0(i, j) &= \hat{R}_0(i, j) \times \hat{R}_0(i, i) \times \hat{R}_0(j, j), & \text{for } i \neq j \\ &= \hat{R}_0(i, j) \times \hat{R}_0(i, j), & \text{for } i = j \end{aligned}$$

9 Example

This example fits a bivariate AR(1) model to two series each of length 48. μ has been estimated but $\phi_1(2, 1)$ has been constrained to be zero. Ten residual cross-correlation matrices are to be computed.

9.1 Program Text

```
function g13ds_example

fprintf('g13ds example results\n\n');

w = [ -1.49 -1.62  5.20  6.23  6.21  5.86  4.09  3.18  2.62  1.49 ...
      1.17  0.85 -0.35  0.24  2.44  2.58  2.04  0.40  2.26  3.34 ...
      5.09  5.00  4.78  4.11  3.45  1.65  1.29  4.09  6.32  7.50 ...
      3.89  1.58  5.21  5.25  4.93  7.38  5.87  5.81  9.68  9.07 ...
      7.29  7.84  7.55  7.32  7.97  7.76  7.00  8.350;
      7.34  6.35  6.96  8.54  6.62  4.97  4.55  4.81  4.75  4.76 ...
     10.88 10.01 11.62 10.36  6.40  6.24  7.93  4.04  3.73  5.60 ...
      5.35  6.81  8.27  7.68  6.65  6.08 10.25  9.14 17.75 13.30 ...
      9.63  6.80  4.08  5.06  4.94  6.65  7.94 10.76 11.89  5.85 ...
```

```

    9.01  7.50 10.02 10.38  8.15  8.37 10.73 12.14];

[k, n] = size(w);
ip = nag_int(1);
iq = nag_int(0);
mean_p = true;
m = nag_int(10);

% Control parameters
iprint = nag_int(-1);
cgetol = 0.0001;
dishow = nag_int(2);
ishow = nag_int(1);
exact = true;

% Initial parameters estimates and hold flags
par = zeros(6,1);
parhld = [false; false; true; false; false; false];

% Initial covariance matrix
qq = zeros(k,k);

% Fit VARMA model
[par, qq, niter, rlogl, v, g, cm, ifail] = ...
    g13dd( ...
        ip, iq, mean_p, par, qq, w, parhld, exact, iprint, cgetol, dishow);

fprintf('\nOutput from g13ds\n\n');

% Display diagnostics
[qq, r0, r, rcm, chi, idf, siglev, ifail] = ...
    g13ds( ...
        v, ip, iq, m, par, parhld, qq, ishow);

```

9.2 Program Results

g13ds example results

VALUE OF LOG LIKELIHOOD FUNCTION ON EXIT = -0.20280E+03

MAXIMUM LIKELIHOOD ESTIMATES OF AR PARAMETER MATRICES

```

-----
PHI(1)    ROW-WISE :    0.802   0.065
                  ( 0.091)( 0.102)

                  0.000   0.575
                  ( 0.000)( 0.121)

```

MAXIMUM LIKELIHOOD ESTIMATE OF PROCESS MEAN

```

-----
                4.271   7.825
                ( 1.219)( 0.776)

```

MAXIMUM LIKELIHOOD ESTIMATE OF SIGMA MATRIX

```

-----
                2.964

                0.637   5.380

```

RESIDUAL SERIES NUMBER 1

```

-----
T      1      2      3      4      5      6      7      8
V(T) -3.33 -1.24  5.75  1.27  0.32  0.11 -1.27 -0.73

T      9     10     11     12     13     14     15     16
V(T) -0.58 -1.26 -0.67 -1.13 -2.02 -0.57  1.24 -0.13

```

T	17	18	19	20	21	22	23	24
V(T)	-0.77	-2.09	1.34	0.95	1.71	0.23	-0.01	-0.60
T	25	26	27	28	29	30	31	32
V(T)	-0.68	-1.89	-0.77	2.05	2.11	0.94	-3.32	-2.50
T	33	34	35	36	37	38	39	40
V(T)	3.16	0.47	0.05	2.77	-0.82	0.25	3.99	0.20
T	41	42	43	44	45	46	47	48
V(T)	-0.70	1.07	0.44	0.28	1.09	0.50	-0.10	1.70

RESIDUAL SERIES NUMBER 2

T	1	2	3	4	5	6	7	8
V(T)	-0.19	-1.20	-0.02	1.21	-1.62	-2.16	-1.63	-1.13
T	9	10	11	12	13	14	15	16
V(T)	-1.34	-1.30	4.82	0.43	2.54	0.35	-2.88	-0.77
T	17	18	19	20	21	22	23	24
V(T)	1.02	-3.85	-1.92	0.13	-1.20	0.41	1.03	-0.40
T	25	26	27	28	29	30	31	32
V(T)	-1.09	-1.07	3.43	-0.08	9.17	-0.23	-1.34	-2.06
T	33	34	35	36	37	38	39	40
V(T)	-3.16	-0.61	-1.30	0.48	0.79	2.87	2.38	-4.31
T	41	42	43	44	45	46	47	48
V(T)	2.32	-1.01	2.38	1.29	-1.14	0.36	2.59	2.64

Output from g13ds

RESIDUAL CROSS-CORRELATION MATRICES

LAG	1	:	0.130	0.112
			(0.119)	(0.143)
			0.094	0.043
			(0.069)	(0.102)
LAG	2	:	-0.312	0.021
			(0.128)	(0.144)
			-0.162	0.098
			(0.125)	(0.132)
LAG	3	:	0.004	-0.176
			(0.134)	(0.144)
			-0.168	-0.091
			(0.139)	(0.140)
LAG	4	:	-0.090	-0.120
			(0.137)	(0.144)
			0.099	-0.232
			(0.142)	(0.143)
LAG	5	:	0.041	0.093
			(0.140)	(0.144)
			-0.009	-0.089
			(0.144)	(0.144)
LAG	6	:	0.234	-0.008
			(0.141)	(0.144)
			0.069	-0.103
			(0.144)	(0.144)
LAG	7	:	-0.076	0.007
			(0.142)	(0.144)


```

                0.168   0.000
                ( 0.144)( 0.144)

LAG      8      :   -0.074   0.559
                  ( 0.143)( 0.144)
                  0.008   -0.101
                  ( 0.144)( 0.144)

LAG      9      :    0.091   0.193
                  ( 0.144)( 0.144)
                  0.055   0.170
                  ( 0.144)( 0.144)

LAG     10     :   -0.060   0.061
                  ( 0.144)( 0.144)
                  0.191   0.089
                  ( 0.144)( 0.144)

```

SUMMARY TABLE

LAGS 1 - 10

```

*****
*           *           *
*  .-..... * .....+. *
*           *           *
*****
*           *           *
*  ..... * ..... *
*           *           *
*****

```

LI-MCLEOD PORTMANTEAU STATISTIC = 49.234
SIGNIFICANCE LEVEL = 0.086
(BASED ON 37 DEGREES OF FREEDOM)
