

NAG Toolbox

nag_tsa_multi_regmat_partial (g13dp)

1 Purpose

`nag_tsa_multi_regmat_partial` (g13dp) calculates the sample partial autoregression matrices of a multivariate time series. A set of likelihood ratio statistics and their significance levels are also returned. These quantities are useful for determining whether the series follows an autoregressive model and, if so, of what order.

2 Syntax

```
[maxlag, parlag, se, qq, x, pvalue, loglhd, ifail] =  
nag_tsa_multi_regmat_partial(k, z, m, 'n', n)  
  
[maxlag, parlag, se, qq, x, pvalue, loglhd, ifail] = g13dp(k, z, m, '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. The partial autoregression matrix at lag l , P_l , is defined to be the last matrix coefficient when a vector autoregressive model of order l is fitted to the series. P_l has the property that if W_t follows a vector autoregressive model of order p then $P_l = 0$ for $l > p$.

Sample estimates of the partial autoregression matrices may be obtained by fitting autoregressive models of successively higher orders by multivariate least squares; see Tiao and Box (1981) and Wei (1990). These models are fitted using a QR algorithm based on the functions `nag_correg_linregm_obs_edit` (g02dc) and `nag_correg_linregm_var_del` (g02df). They are calculated up to lag m , which is usually taken to be at most $n/4$.

The function also returns the asymptotic standard errors of the elements of \hat{P}_l and an estimate of the residual variance-covariance matrix $\hat{\Sigma}_l$, for $l = 1, 2, \dots, m$. If S_l denotes the residual sum of squares and cross-products matrix after fitting an $AR(l)$ model to the series then under the null hypothesis $H_0 : P_l = 0$ the test statistic

$$X_l = -\left((n - m - 1) - \frac{1}{2} - lk\right) \log \left(\frac{|S_l|}{|S_{l-1}|} \right)$$

is asymptotically distributed as χ^2 with k^2 degrees of freedom. X_l provides a useful diagnostic aid in determining the order of an autoregressive model. (Note that $\hat{\Sigma}_l = S_l / (n - l)$.) The function also returns an estimate of the maximum of the log-likelihood function for each AR model that has been fitted.

4 References

Tiao G C and Box G E P (1981) Modelling multiple time series with applications *J. Am. Stat. Assoc.* **76** 802–816

Wei W W S (1990) *Time Series Analysis: Univariate and Multivariate Methods* Addison–Wesley

5 Parameters

5.1 Compulsory Input Parameters

- 1: **k** – INTEGER
 k , the number of time series.
Constraint: $k \geq 1$.
- 2: **z(kmax, n)** – REAL (KIND=nag_wp) array
 $kmax$, the first dimension of the array, must satisfy the constraint $kmax \geq k$.
 $z(i, t)$ must contain the observation w_{it} , for $i = 1, 2, \dots, k$ and $t = 1, 2, \dots, n$.
- 3: **m** – INTEGER
 m , the number of partial autoregression matrices to be computed. If in doubt set $m = 10$.
Constraint: $m \geq 1$ and $n - m - (k \times m + 1) \geq k$.

5.2 Optional Input Parameters

- 1: **n** – INTEGER
Default: the second dimension of the array **z**.
 n , the number of observations in the time series.
Constraint: $n \geq 4$.

5.3 Output Parameters

- 1: **maxlag** – INTEGER
The maximum lag up to which partial autoregression matrices (along with their likelihood ratio statistics and their significance levels) have been successfully computed. On a successful exit **maxlag** will equal **m**. If **ifail** = 2 on exit then **maxlag** will be less than **m**.
- 2: **parlag(kmax, kmax, m)** – REAL (KIND=nag_wp) array
 $kmax = k$.
parlag(i, j, l) contains an estimate of the (i, j)th element of the partial autoregression matrix at lag l , $\hat{P}_l(ij)$, for $l = 1, 2, \dots, \mathbf{maxlag}$, $i = 1, 2, \dots, k$ and $j = 1, 2, \dots, k$.
- 3: **se(kmax, kmax, m)** – REAL (KIND=nag_wp) array
 $kmax = k$.
se(i, j, l) contains an estimate of the standard error of the corresponding element in the array **parlag**.
- 4: **qq(kmax, kmax, m)** – REAL (KIND=nag_wp) array
 $kmax = k$.
qq(i, j, l) contains an estimate of the (i, j)th element of the corresponding variance-covariance matrix $\hat{\Sigma}_l$, for $l = 1, 2, \dots, \mathbf{maxlag}$, $i = 1, 2, \dots, k$ and $j = 1, 2, \dots, k$.
- 5: **x(m)** – REAL (KIND=nag_wp) array
x(l) contains X_l , the likelihood ratio statistic at lag l , for $l = 1, 2, \dots, \mathbf{maxlag}$.

- 6: **pvalue(m)** – REAL (KIND=nag_wp) array
pvalue(l) contains the significance level of the statistic in the corresponding element of **x**.
- 7: **loglhd(m)** – REAL (KIND=nag_wp) array
loglhd(l) contains an estimate of the maximum of the log-likelihood function when an AR(*l*) model has been fitted to the series, for $l = 1, 2, \dots, \mathbf{maxlag}$.
- 8: **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, **k** < 1,
or **n** < 4,
or $kmax < \mathbf{k}$,
or **m** < 1,
or $\mathbf{n} - \mathbf{m} - (\mathbf{k} \times \mathbf{m} + 1) < \mathbf{k}$,
or *lwork* is too small.

ifail = 2 (*warning*)

The recursive equations used to compute the sample partial autoregression matrices have broken down at lag **maxlag** + 1. This exit could occur if the regression model is overparameterised. For your settings of *k* and *n* the value returned by **maxlag** is the largest permissible value of *m* for which the model is not overparameterised. All output quantities in the arrays **parlag**, **se**, **qq**, **x**, **pvalue** and **loglhd** up to and including lag **maxlag** will be correct.

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 is roughly proportional to *nmk*.

For each order of autoregressive model that has been estimated, nag_tsa_multi_regmat_partial (g13dp) returns the maximum of the log-likelihood function. An alternative means of choosing the order of a vector AR process is to choose the order for which Akaike's information criterion is smallest. That is, choose the value of *l* for which $-2 \times \mathbf{loglhd}(l) + 2lk^2$ is smallest. You should be warned that this does not always lead to the same choice of *l* as indicated by the sample partial autoregression matrices and the likelihood ratio statistics.

9 Example

This example computes the sample partial autoregression matrices of two time series of length 48 up to lag 10.

9.1 Program Text

```
function g13dp_example

fprintf('g13dp example results\n\n');

z = [-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, 8.35;
      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(z);
k = nag_int(k);
m = nag_int(10);

[maxlag, parlag, se, qq, x, pvalue, loglhd, ifail] = ...
    g13dp(k, z, m);

fprintf('%31s%13s%10s%13s%8s\n','Partial Autoregression Matrices', ...
        'Indicator', 'Residual', 'Chi-Square', 'Pvalue');
fprintf('%43s%12s%12s\n', 'Symbols', 'Variances', 'Statistic');
fprintf('%31s%13s%10s%13s%8s\n','-----', ...
        '-----', '-----', '-----')
for l = 1:maxlag
    for j = 1:k
        sum = parlag(l,j,l);
        st(j) = '.';
        if sum>1.96*se(l,j,l)
            st(j) = '+';
        end
        if sum<-1.96*se(l,j,l)
            st(j) = '-';
        end
    end
    fprintf('\n Lag %2d :%8.3f%8.3f%15s%13.3f%13.3f%9.3f\n', ...
            l, parlag(l,1:k,l), st(1:k), qq(l,1,l), x(l), pvalue(l));
    fprintf('          (%6.3f)(%6.3f)\n', se(l,1:k,l));
    for i = 2:k
        for j = 1:k
            sum = parlag(i,j,l);
            st(j) = '.';
            if sum>1.96*se(i,j,l)
                st(j) = '+';
            end
            if sum<-1.96*se(i,j,l)
                st(j) = '-';
            end
        end
        fprintf('%17.3f%8.3f%15s%13.3f\n', parlag(i,1:k,l), st(1:k), qq(i,i,l));
        fprintf('          (%6.3f)(%6.3f)\n', se(i,1:k,l));
    end
end
end
```

9.2 Program Results

g13dp example results

Partial Autoregression Matrices	Indicator Symbols	Residual Variances	Chi-Square Statistic	Pvalue
-----	-----	-----	-----	-----
Lag 1 : 0.757 0.062 (0.092)(0.092) 0.061 0.570 (0.129)(0.130)	+. .+	2.731 5.440	49.884	0.000
Lag 2 : -0.161 -0.135 (0.145)(0.109) -0.093 -0.065 (0.213)(0.160)	2.530 5.486	3.347	0.502
Lag 3 : 0.237 0.044 (0.128)(0.095) 0.047 -0.248 (0.222)(0.165)	1.755 5.291	13.962	0.007
Lag 4 : -0.098 0.152 (0.134)(0.099) 0.402 -0.194 (0.228)(0.168)	1.661 4.786	7.071	0.132
Lag 5 : 0.257 -0.026 (0.141)(0.106) 0.400 -0.021 (0.242)(0.183)	1.504 4.447	5.184	0.269
Lag 6 : -0.075 0.112 (0.156)(0.111) 0.196 -0.106 (0.269)(0.192)	1.480 4.425	2.083	0.721
Lag 7 : -0.054 0.097 (0.166)(0.121) 0.574 -0.080 (0.267)(0.195)	.. +.	1.478 3.838	5.074	0.280
Lag 8 : 0.147 0.041 (0.188)(0.128) 0.916 -0.242 (0.246)(0.167)	.. +.	1.415 2.415	10.991	0.027
Lag 9 : -0.039 0.099 (0.251)(0.140) -0.500 0.173 (0.324)(0.181)	1.322 2.196	3.936	0.415
Lag 10 : 0.189 0.131 (0.275)(0.157) -0.183 -0.040 (0.371)(0.212)	1.206 2.201	3.175	0.529
