NAG Fortran Library Routine Document

G13DSF

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

G13DSF is a diagnostic checking routine suitable for use after fitting a vector ARMA model to a multivariate time series using G13DCF. The residual cross-correlation matrices are returned along with an estimate of their asymptotic standard errors and correlations. Also, G13DSF calculates the modified Li–McLeod portmanteau statistic and its significance level for testing model adequacy.

2 Specification

```
SUBROUTINE G13DSF(K, N, V, IK, IP, IQ, M, PAR, PARHLD, QQ, ISHOW, R0, R,
1 RCM, IRCM, CHI, IDF, SIGLEV, IW, LIW, WORK, LWORK,
2 IFAIL)
INTEGER K, N, IK, IP, IQ, M, ISHOW, IRCM, IDF, IW(LIW), LIW, 1LWORK, IFAIL
real V(IK,N), PAR((IP+IQ)*K*K), QQ(IK,K), RO(IK,K),
1 R(IK,IK,M), RCM(IRCM,M*K*K), CHI, SIGLEV, WORK(LWORK)<br>LOGICAL PARHLD((IP+IO)*K*K)
                 PARTL D (ID+IO) * K*K)
```
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

$$
W_t - \mu = \phi_1(W_{t-1} - \mu) + \phi_2(W_{t-2} - \mu) + \dots + \phi_p(W_{t-p} - \mu)
$$

$$
+ \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q},
$$
 (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 $i = 1, 2, \ldots, p$, are called the autoregressive (AR) parameter matrices, and $\{\theta_i\}$, for $i = 1, 2, \ldots, q$, 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

where I denotes the k by k identity matrix.

The 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 G13DCF).

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

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

where $\hat{\epsilon}_{it}$, denotes an estimate of the tth 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))

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

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 8.2.\) Let](#page-5-0) 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 $\hat{\mathbf{r}} = (\text{vec}(R_1^T), \text{vec}(R_2^T), \dots, \text{vec}(R_m^T))$ then the covariance matrix of $\hat{\mathbf{r}}$ is given by

$$
Var(\underline{\hat{r}}) = [Y - X(X^TGG^T X)^{-1} X^T]/n,
$$

where $Y = I_m \otimes (\Delta \otimes \Delta)$ and $G = I_m(GG^T)$. Δ is the dispersion matrix Σ in correlation form and G a non-singular k by k matrix such that $GG^T = \Delta^{-1}$ 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 var (\hat{r}) and therefore is not required as input to G13DSF.)

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, IK, IP, IQ, PAR, PARHLD and QQ fro](#page-2-0)m G13DCF are suitable for input to G13DSF.

1: K – INTEGER *Input*

On entry: the number of residual time series, k .

Constraint: $K > 1$.

2: N – INTEGER *Input*

On entry: the number of observations in each residual series, n .

Constraint: $N > 3$.

 $3:$ V(IK,N) – real array Input

On entry: an estimate of the *i*th component of ϵ_t , for $i = 1, 2, \ldots, k; t = 1, 2, \ldots, n$.

Constraints:

no two rows of V may be identical, in each row there must be at least two distinct elements.

4: IK – INTEGER Input

On entry: the first dimension of the arrays V, QQ and R0 as declared in the (sub)program from which G13DSF is called.

Constraint: $IK \geq K$.

5: IP – INTEGER *Input*

On entry: the number of AR parameter matrices, p .

Constraint: $IP > 0$.

6: IQ – INTEGER *Input*

On entry: the number of MA parameter matrices, q.

Constraint: $IO > 0$.

Note: $IP = IQ = 0$ is not permitted.

7: M – INTEGER *Input*

On entry: the value of m , the number of residual cross-correlation matrices to be computed. See [Section 8.2 for advice](#page-5-0) on the choice of M.

Constraint: $IP + IO < M < N$.

8: $PAR((IP+IQ)*K*K) - real$ array Input

On entry: the parameter estimates read in row by row in the order $\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q$. Thus, if IP > 0 then $PAR((l-1) \times k \times k + (i-1) \times k + j)$ must be set equal to an estimate of the (i, j) th element of ϕ_l for $l = 1, 2, \ldots, p;$ $i, j = 1, 2, \ldots, k$. If $IQ \ge 0$ then $PAR(p \times k \times k + (l-1) \times k \times k + (i-1) \times k + j)$ must be set equal to an estimate of the (i, j) th element of θ_l for $l = 1, 2, \ldots, q; i, j = 1, 2, \ldots, k$.

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.

9: PARHLD((IP+IQ)*K*K) – LOGICAL array Input

On entry: 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 $i = 1, 2, \ldots, (p + q) \times k \times k$.

10: $QQ(IK,K)$ – real array Input/Output

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

Constraint: QQ must be positive-definite.

On exit: if I[FAIL](#page-3-0) \neq 1, then the upper triangle is set equal to the lower triangle.

11: ISHOW – INTEGER *Input*

On entry: ISHOW must be non-zero if the residual cross-correlation matrices $\{\hat{r}_{i,j}(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 and zero otherwise. 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 $\hat{r}_{ii}(l) > 1.96 \times \text{se}(\hat{r}_{ii}(l))$ then a '+' is printed, if $\hat{r}_{ii}(l) < -1.96 \times \text{se}(\hat{r}_{ii}(l))$ then a '-' is printed, otherwise a '.' is printed. The summary table is only printed if $k \le 6$ on entry.

The residual cross-correlation matrices, their standard errors and the modified portmanteau statistic with its significance are available also as output variable[s in R, RCM, CHI, IDF and SIGLEV.](#page-3-0)

12: $R0(IK,K)$ – real array $Output$

On exit: if $i \neq j$, then RO (i, j) contains an estimate of the (i, j) th element of the residual crosscorrelation 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 ro](#page-1-0)ws and columns of R0 are set to zero.

13: R(IK,IK,M) – real array Output

On exit: an estimate of the (i, j) th element of the residual cross-correlation matrix at lag l, for $l = 1, 2, \ldots, m; i, j = 1, 2, \ldots, k$. If IFAIL = 3 on exit then all elements of R are set to zero.

14: RCM(IRCM, M^{*}K^{*}K) – *real* array *Output*

On exit: the estimated standard errors and correlations of the elements in the array R. The correlation between $R(i, j, l)$ and $R(i_2, j_2, l_2)$ 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 (i, j, l) . If on exit, IFAIL ≥ 5 , then all off $s = t$, then KCM(s, t) contains the standard error of K(t, f, t). If on exit, fractize 2, s, diagonal elements of RCM are set to zero and all diagonal elements are set to $1/\sqrt{n}$.

15: IRCM – INTEGER *Input*

On entry: the first dimension of the array RCM as declared in the (sub)program from which G13DSF is called.

Constraint: IRC[M](#page-2-0) $> M \times K \times K$ $> M \times K \times K$ $> M \times K \times K$.

16: CHI – real *Output*

On exit: the value of the modified portmanteau statistic, $Q_{(m)}^*$. If IFAIL = 3 on exit then CHI is returned as zero.

17: IDF – INTEGER Output

On exit: the number of degrees of freedom of CHI.

18: SIGLEV – real Output

On exit: the significance level of CHI based on IDF degrees of freedom. If IFAIL $=$ 3 on exit then SIGLEV is returned as one.

- 19: IW(LIW) INTEGER array Workspace Workspace
- 20: LIW INTEGER *Input*

On entry: the dimension of the array IW as declared in the (sub)program from which G13DSF is called.

Constraint: $LIW > K \times max(IP, IO)$ $LIW > K \times max(IP, IO)$.

21: WORK(LWORK) – real array Workspace Workspace

22: LWORK – INTEGER Input

On entry: the dimension of the array WORK as declared in the (sub)program from which G13DSF is called.

Constraint: if NPAR = $(p+q)k^2$, then LWORK $\geq k(n+IK+2) + mk^2(NPAR + mk^2 + 1) +$ $\geq k(n+IK+2) + mk^2(NPAR + mk^2 + 1) +$ $\geq k(n+IK+2) + mk^2(NPAR + mk^2 + 1) +$ $3k^2 + (NPAR + 1)NPAR$.

23: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to $0, -1$ or 1. Users who are unfamiliar with this parameter should refer to Chapter P01 for details.

On exit: IFAIL $= 0$ unless the routine detects an [error \(see Section 6\).](#page-4-0)

For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the

value 1 is recommended. Otherwise, because for this routine the values of the output parameters may be useful even if IFAIL $\neq 0$ on exit, the recommended value is -1 . When the value -1 or 1 is used it is essential to test the value of IFAIL on exit.

6 Error Indicators and Warnings

If on entry IFAIL $= 0$ or -1 , explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

 $IFAIL = 1$

On entry, $K < 1$, or $IK < K$, or $IP < 0$, or $IQ < 0$, or $IP = IO = 0,$ or $M \le IP + IQ$ $M \le IP + IQ$ $M \le IP + IQ$, or $M > N$, or $IRCM < M \times K \times K$ $IRCM < M \times K \times K$. or [LIW is to](#page-3-0)o small, o[r LWORK is too sm](#page-3-0)all.

$IFAIL = 2$

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, the user must supply different parameter estimates in the arrays PAR and QQ.

 $IFAIL = 3$

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 an](#page-3-0)d all the elemen[ts of R0 and R se](#page-3-0)t 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 parameters are undefined.

```
IFAIL = 5
```
On entry, either the eigenvalues and eigenvectors of Δ (the [matrix QQ in c](#page-2-0)orrelation form) could not be computed or the determinantal polynomials $det(A(\phi))$ and $det(B(\theta))$ have a factor in common. To proceed, the user must either supply different parameter estimates in the [array QQ or](#page-2-0) delete this common factor from the model. In this case, the off-diagonal elem[ents of RCM are](#page-3-0) 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 elem[ents of RCM was f](#page-3-0)ound to be either negative or zero. In this case all off-diagonal elem[ents of RCM are re](#page-3-0)turned as zero and all diagonal elem[ents of RCM set to](#page-3-0) $1/\sqrt{n}$.

7 Accuracy

The computations are believed to be stable.

8 Further Comments

8.1 Timing

The time taken by the routine 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 m[odel \(1\) is w](#page-0-0)ritten 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 $\{\pi_1, \pi_2, ...\}$ and $\{\psi_1, \psi_2, ...\}$ are such that π_i and ψ_i 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, the user should try setting $m = 20$.

8.3 Checking a White Noise Model

If the user has fitted the 'white noise' model

$$
W_t - \mu = \epsilon_t
$$

then G13DSF should be entered with $p = 1$, $q = 0$, and the first k^2 elem[ents of PAR and PARHLD set to](#page-2-0) zero and .TRUE. respectively.

8.4 Approximate Standard Errors

When IFAIL is returned as 5 or 6 all the standard e[rrors in RCM are se](#page-3-0)t 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

 R_0 is useful in testing for instantaneous causality. If the user wishes 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

$$
\hat{C}_0(i,j) = \hat{R}_0(i,j) \times \hat{R}_0(i,i) \times \hat{R}_0(j,j), \quad \text{for } i \neq j
$$

$$
= \hat{R}_0(i,j) \times \hat{R}_0(i,j), \qquad \text{for } i = j
$$

9 Example

A program to fit 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

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.

```
G13DSF Example Program Text
* Mark 15 Revised. NAG Copyright 1991.
     .. Parameters ..
     INTEGER NIN, NOUT
     PARAMETER (NIN=5, NOUT=6)
     INTEGER KMAX, IK, IPMAX, IQMAX, NMAX, NPARMX, LWORK, LIW,
    + MMAX, IRCM, ICM
     PARAMETER (KMAX=3,IK=KMAX,IPMAX=3,IQMAX=3,NMAX=100,
    + NPARMX=(IPMAX+IQMAX)*KMAX*KMAX+KMAX,LWORK=2000,
                    LIW=100,MMAX=20, IRCM=MMAX*KMAX*KMAX, ICM=NPARMX)
* .. Local Scalars ..
     real CGETOL, CHI, RLOGL, SIGLEV
     INTEGER I, IDF, IFAIL, IP, IPRINT, IQ, ISHOW, J, K, M,
    + MAXCAL, N, NITER, NPAR
     LOGICAL EXACT, MEAN
* .. Local Arrays ..
                    CM(ICM,NPARMX), G(NPARMX), PAR(NPARMX),
    + QQ(IK,KMAX), R(IK,IK,MMAX), R0(IK,KMAX),
    + RCM(IRCM,MMAX*KMAX*KMAX), V(IK,NMAX), W(IK,NMAX),<br>+ WORK(LWORK)
                    WORK(LWORK)
     INTEGER IW(LIW)
     LOGICAL PARHLD(NPARMX)
* .. External Subroutines ..
                   G13DCF, G13DSF, X04ABF
     .. Executable Statements ..
     WRITE (NOUT,*) 'G13DSF Example Program Results'
     Skip heading in data file
     READ (NIN,*)
     READ (NIN,*) K, N
*
     CALL X04ABF(1,NOUT)
*
     IF (K.GT.0 .AND. K.LE.KMAX .AND. N.GE.3 .AND. N.LE.NMAX) THEN
       DO 20 I = 1, K
          READ (NIN,*) (W(I,J),J=1,N)20 CONTINUE
       READ (NIN,*) IP, IQ, MEAN, M
       IF (IP.GE.0 .AND. IP.LE.IPMAX .AND. IQ.GE.0 .AND. IQ.LE.IQMAX)
    + THEN
          NPAR = (IP+IQ)*K*KIF (MEAN) NPAR = NPAR + K
          IF (NPAR.LE.NPARMX) THEN
             DO 40 I = 1, NPAR
               PAR(I) = 0.0e0PARHLD(I) = .FALSE.40 CONTINUE
             DO 80 J = 1, KDO 60 I = J, K
                  QQ(I,J) = 0.0e060 CONTINUE
  80 CONTINUE
             PARHLD(3) = .TRUE.EXACT = .TRUE.* ** Set IPRINT > 0 to obtain intermediate output **
             IPRINT = -1CGETOL = 0.0001e0MAXCAL = 40*NPAR*(NPAR+5)ISBNOW = 2IFAIL = 1*
             CALL G13DCF(K,N,IP,IQ,MEAN,PAR,NPAR,QQ,IK,W,PARHLD,EXACT,
    + IPRINT,CGETOL,MAXCAL,ISHOW,NITER,RLOGL,V,G,
    + CM,ICM,WORK,LWORK,IW,LIW,IFAIL)
*
             WRITE (NOUT,*)
```

```
IF (IFAIL.NE.0) THEN
               WRITE (NOUT,99999) 'G13DCF fails. IFAIL =', IFAIL
               WRITE (NOUT,*)
             END IF
             IF ((IFAIL.EQ.0 .OR. IFAIL.GE.4) .AND. M.LE.MMAX) THEN
               WRITE (NOUT,*) 'Output from G13DSF'
               WRITE (NOUT,*)
               ISBNOW = 1IFAIL = -1*
               CALL G13DSF(K,N,V,IK,IP,IQ,M,PAR,PARHLD,QQ,ISHOW,R0,R,
    + RCM,IRCM,CHI,IDF,SIGLEV,IW,LIW,WORK,LWORK,
    + IFAIL)
*
               IF (IFAIL.NE.0) WRITE (NOUT,99999)
    + 'G13DSF fails. IFAIL =', IFAIL
             END IF
          END IF
       END IF
     END IF
     STOP
*
99999 FORMAT (1X,A,I3)
     END
```
9.2 Program Data

```
G13DSF Example Program Data
2 48 : K, no. of time series, N, no. of obs in each time series
-1.490 -1.620 5.200 6.230 6.210 5.860
4.090 3.180 2.620 1.490 1.170 0.850
-0.350 0.240 2.440 2.580 2.040 0.400
 2.260 3.340 5.090 5.000 4.780 4.110
 3.450 1.650 1.290 4.090 6.320 7.500
3.890 1.580 5.210 5.250 4.930 7.380
5.870 5.810 9.680 9.070 7.290 7.840
 7.550 7.320 7.970 7.760 7.000 8.350
 7.340 6.350 6.960 8.540 6.620 4.970
 4.550 4.810 4.750 4.760 10.880 10.010
11.620 10.360 6.400 6.240 7.930 4.040
 3.730 5.600 5.350 6.810 8.270 7.680
 6.650 6.080 10.250 9.140 17.750 13.300
 9.630 6.800 4.080 5.060 4.940 6.650
7.940 10.760 11.890 5.850 9.010 7.500
10.020 10.380 8.150 8.370 10.730 12.140 : End of time series
1 0 T 10 : IP, IQ, MEAN and M
```
9.3 Program Results

G13DSF Example Program Results VALUE OF IFAIL PARAMETER ON EXIT FROM G13DCF = 0 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
