

NAG Fortran Library Routine Document

G13DPF

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

G13DPF 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 Specification

```

SUBROUTINE G13DPF(K, N, Z, IK, M, MAXLAG, PARLAG, SE, QQ, X, PVALUE,
1          LOGLHD, WORK, LWORK, IWORK, IFAIL)
  INTEGER      K, N, IK, M, MAXLAG, LWORK, IWORK(K*M), IFAIL
  real       Z(IK,N), PARLAG(IK,IK,M), SE(IK,IK,M), QQ(IK,IK,M),
1          X(M), PVALUE(M), LOGLHD(M), WORK(LWORK)

```

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 routines G02DCF and G02DFF. They are calculated up to lag m , which is usually taken to be at most $n/4$.

The routine also returns the asymptotic standard errors of the elements of \hat{P}_l and an estimate of the corresponding 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 routine 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

1: K – INTEGER

Input

On entry: the number of time series, k .

Constraint: $K \geq 1$.

- 2: N – INTEGER *Input*
On entry: the number of observations in the time series, n .
Constraint: $N \geq 4$.
- 3: Z(IK,N) – *real* array *Input*
On entry: Z(i, t) must contain the observation w_{it} for $i = 1, 2, \dots, k$; $t = 1, 2, \dots, n$.
- 4: IK – INTEGER *Input*
On entry: the first dimension of the array Z and the first and second dimensions of arrays PARLAG, SE, QQ as declared in the (sub)program from which G13DPF is called.
Constraint: $IK \geq K$.
- 5: M – INTEGER *Input*
On entry: the number, m , of partial autoregression matrices to be computed. If in doubt set $M = 10$.
Constraints: $M \geq 1$ and $N - M - (K \times M + 1) \geq K$.
- 6: MAXLAG – INTEGER *Output*
On exit: 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.
- 7: PARLAG(IK,IK,M) – *real* array *Output*
On exit: 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 $l = 1, 2, \dots, \text{MAXLAG}$; $i = 1, 2, \dots, k$; $j = 1, 2, \dots, k$.
- 8: SE(IK,IK,M) – *real* array *Output*
On exit: SE(i, j, l) contains an estimate of the standard error of the corresponding element in the array PARLAG.
- 9: QQ(IK,IK,M) – *real* array *Output*
On exit: QQ(i, j, l) contains an estimate of the (i, j)th element of the corresponding variance-covariance matrix $\hat{\Sigma}_l$ for $l = 1, 2, \dots, \text{MAXLAG}$; $i = 1, 2, \dots, k$; $j = 1, 2, \dots, k$.
- 10: X(M) – *real* array *Output*
On exit: X(l) contains X_l , the likelihood ratio statistic at lag l for $l = 1, 2, \dots, \text{MAXLAG}$.
- 11: PVALUE(M) – *real* array *Output*
On exit: PVALUE(l) contains the significance level of the statistic in the corresponding element of X.
- 12: LOGLHD(M) – *real* array *Output*
On exit: 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 $l = 1, 2, \dots, \text{MAXLAG}$.
- 13: WORK(LWORK) – *real* array *Workspace*
14: LWORK – INTEGER *Input*
On entry: the dimension of the array WORK as declared in the (sub)program from which G13DPF is called.
Constraint: $LWORK \geq (k + 1)k + l(4 + k) + 2l^2$, where $l = mk + 1$.

15: IWORK(K*M) – INTEGER array *Workspace*

16: 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).

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, for users not familiar with this parameter the recommended value is 0. **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 $N < 4$,
 or $IK < K$,
 or $M < 1$,
 or $N - M - (K \times M + 1) < K$,
 or LWORK is too small.

IFAIL = 2

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 the user's 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 LOGHLD up to and including lag MAXLAG will be correct.

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, G13DPF 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 $-2 \times \text{LOGHLD}(l) + 2lk^2$ is smallest. The user 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

A program to compute the sample partial autoregression matrices of two time series of length 48 up to lag 10.

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.

```

*      G13DPF Example Program Text
*      Mark 16 Release. NAG Copyright 1992.
*      .. Parameters ..
INTEGER          NIN, NOUT
PARAMETER       (NIN=5,NOUT=6)
INTEGER          KMAX, NMAX, MMAX, LWORK
PARAMETER       (KMAX=4,NMAX=50,MMAX=10,LWORK=2081)
*      .. Local Scalars ..
INTEGER          I, IFAIL, J, K, M, MAXLAG, N
*      .. Local Arrays ..
real           LOGLHD(MMAX), PARLAG(KMAX,KMAX,MMAX),
+              PVALUE(MMAX), QQ(KMAX,KMAX,MMAX),
+              SE(KMAX,KMAX,MMAX), W(KMAX,NMAX), WORK(LWORK),
+              X(MMAX)
INTEGER          IWORK(KMAX*MMAX)
*      .. External Subroutines ..
EXTERNAL        G13DPF, ZPRINT
*      .. Executable Statements ..
WRITE (NOUT,*) 'G13DPF Example Program Results'
*      Skip heading in data file
READ (NIN,*)
READ (NIN,*) K, N, M
*
IF (K.GT.0 .AND. K.LE.KMAX .AND. N.GE.1 .AND. N.LE.NMAX .AND.
+  M.GE.1 .AND. M.LE.MMAX) THEN
*
DO 20 I = 1, K
    READ (NIN,*) (W(I,J),J=1,N)
20 CONTINUE
*
IFAIL = 0
*
CALL G13DPF(K,N,W,KMAX,M,MAXLAG,PARLAG,SE,QQ,X,PVALUE,LOGLHD,
+         WORK,LWORK,IWORK,IFAIL)
*
CALL ZPRINT(K,N,M,KMAX,MAXLAG,PARLAG,SE,QQ,X,PVALUE,NOUT,IFAIL)
*
END IF
STOP
END
*
SUBROUTINE ZPRINT(K,N,M,KMAX,MAXLAG,PARLAG,SE,QQ,X,PVALUE,NOUT,
+              IFAIL)
*
.. Scalar Arguments ..
INTEGER          IFAIL, K, KMAX, M, MAXLAG, N, NOUT
*
.. Array Arguments ..
real           PARLAG(KMAX,KMAX,M), PVALUE(M), QQ(KMAX,KMAX,M),
+              SE(KMAX,KMAX,M), X(M)
*
.. Local Scalars ..
real           SUM
INTEGER          I, I2, J, L
*
.. Local Arrays ..
CHARACTER*6      ST(6)
*
.. Executable Statements ..
*
IF (K.GT.1) WRITE (NOUT,99999)
IF (K.EQ.1) WRITE (NOUT,99998)
DO 80 L = 1, MAXLAG
    DO 20 J = 1, K
        SUM = PARLAG(1,J,L)
        ST(J) = '.'
        IF (SUM.GT.1.96e0*SE(1,J,L)) ST(J) = '+'
        IF (SUM.LT.-1.96e0*SE(1,J,L)) ST(J) = '-'
20 CONTINUE

```

```

      IF (K.EQ.1) THEN
        WRITE (NOUT,99997) L, (PARLAG(1,J,L),J=1,K),
+       (ST(I2),I2=1,K), QQ(1,1,L), X(L), PVALUE(L)
        WRITE (NOUT,99996) (SE(1,J,L),J=1,K)
      ELSE IF (K.EQ.2) THEN
        WRITE (NOUT,99995) L, (PARLAG(1,J,L),J=1,K),
+       (ST(I2),I2=1,K), QQ(1,1,L), X(L), PVALUE(L)
        WRITE (NOUT,99994) (SE(1,J,L),J=1,K)
      ELSE IF (K.EQ.3) THEN
        WRITE (NOUT,99993) L, (PARLAG(1,J,L),J=1,K),
+       (ST(I2),I2=1,K), QQ(1,1,L), X(L), PVALUE(L)
        WRITE (NOUT,99992) (SE(1,J,L),J=1,K)
      ELSE IF (K.EQ.4) THEN
        WRITE (NOUT,99991) L
        WRITE (NOUT,99984) (PARLAG(1,J,L),J=1,K), (ST(I2),I2=1,K),
+       QQ(1,1,L), X(L), PVALUE(L)
        WRITE (NOUT,99990) (SE(1,J,L),J=1,K)
      END IF
*
      DO 60 I = 2, K
*
        DO 40 J = 1, K
          SUM = PARLAG(I,J,L)
          ST(J) = '.'
          IF (SUM.GT.1.96E0*SE(I,J,L)) ST(J) = '+'
          IF (SUM.LT.-1.96E0*SE(I,J,L)) ST(J) = '-'
40        CONTINUE
          IF (K.EQ.2) THEN
            WRITE (NOUT,99987) (PARLAG(I,J,L),J=1,K),
+            (ST(I2),I2=1,K), QQ(I,I,L)
            WRITE (NOUT,99994) (SE(I,J,L),J=1,K)
          ELSE IF (K.EQ.3) THEN
            WRITE (NOUT,99986) (PARLAG(I,J,L),J=1,K),
+            (ST(I2),I2=1,K), QQ(I,I,L)
            WRITE (NOUT,99992) (SE(I,J,L),J=1,K)
          ELSE IF (K.EQ.4) THEN
            WRITE (NOUT,99985) (PARLAG(I,J,L),J=1,K),
+            (ST(I2),I2=1,K), QQ(I,I,L)
            WRITE (NOUT,99990) (SE(I,J,L),J=1,K)
          END IF
*
60        CONTINUE
80        CONTINUE
*
          WRITE (NOUT,99983) IFAIL
*
          RETURN
*
99999 FORMAT (' Partial Autoregression Matrices',4X,'Indicator',2X,
+ 'Residual',3X,'Chi-Square',2X,'Pvalue',/37X,'Symbols',3X,
+ 'Variances',3X,'Statistic',/ '-----',
+ '-----',4X,'-----',2X,'-----',2X,'-----',1X,
+ '-----')
99998 FORMAT (' Partial Autoregression Function',4X,'Indicator',2X,
+ 'Residual',3X,'Chi-Square',2X,'Pvalue',/37X,'Symbols',3X,
+ 'Variances',3X,'Statistic',/ '-----',
+ '-----',4X,'-----',2X,'-----',2X,'-----',1X,
+ '-----')
99997 FORMAT (' Lag',I3,1X,':',F7.3,22X,A1,F14.3,3X,F10.3,F9.3)
99996 FORMAT (10X,'( ',F5.3,')')
99995 FORMAT (' Lag',I3,1X,':',2F8.3,14X,2A1,F13.3,3X,F10.3,F9.3)
99994 FORMAT (11X,'( ',F5.3,') ( ',F5.3,')')
99993 FORMAT (' Lag',I3,1X,':',3F8.3,6X,3A1,F12.3,3X,F10.3,F9.3)
99992 FORMAT (11X,'( ',F5.3,') ( ',F5.3,') ( ',F5.3,')')
99991 FORMAT (' Lag',I3)
99990 FORMAT (3X,'( ',F5.3,') ( ',F5.3,') ( ',F5.3,') ( ',F5.3,')')
99987 FORMAT (9X,2F8.3,14X,2A1,F13.3)
99986 FORMAT (9X,3F8.3,6X,3A1,F12.3)
99985 FORMAT (1X,4F8.3,5X,4A1,F12.3)
99984 FORMAT (1X,4F8.3,5X,4A1,F12.3,3X,F10.3,F9.3)
99983 FORMAT (' Value of IFAIL parameter on exit from G13DPF = ',I2)

```

END

9.2 Program Data

G13DPF Example Program Data

```

2 48 10 : K, no. of series, N, no. of obs in each series, M, no. of lags
-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

```

9.3 Program Results

G13DPF Example Program 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	1.322 2.196	3.936	0.415

(0.324) (0.181)

Lag 10 :	0.189	0.131	..	1.206	3.175	0.529
	(0.275)	(0.157)				
	-0.183	-0.040	..	2.201		
	(0.371)	(0.212)				

Value of IFAIL parameter on exit from G13DPF = 0
