NAG Fortran Library Routine Document

G13FGF

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

G13FGF estimates the parameters of a univariate regression-exponential GARCH(p, q) process (see Engle and Ng (1993)).

2 Specification

```
      SUBROUTINE G13FGF(DIST, YT, X, LDX, NUM, IP, IQ, NREG, MN, NPAR, THETA,
     1                  SE, SC, COVAR, LDC, HP, ET, HT, LGF, COPTS, MAXIT,
     2                  TOL, WORK, LWORK, IFAIL)
      INTEGER           LDX, NUM, IP, IQ, NREG, MN, NPAR, LDC, MAXIT, LWORK,
     1                  IFAIL
      real              YT(NUM), X(LDX,*), THETA(NPAR), SE(NPAR), SC(NPAR),
     1                  COVAR(LDC,NPAR), HP, ET(NUM), HT(NUM), LGF, TOL,
     2                  WORK(LWORK)
      LOGICAL           COPTS
      CHARACTER*1       DIST
```

3 Description

A univariate regression-exponential GARCH(p, q) process, with q coefficients α_i, for $i = 1, \ldots, q$, q coefficients ϕ_i, for $i = 1, \ldots, q$, p coefficients, β_i, for $i = 1, \ldots, p$ and k linear regression coefficients b_i, for $i = 1, \ldots, k$, can be represented by:

$$ln(h_t) = \alpha_0 + \sum_{i=1}^{q} \alpha_i z_{t-i} + \sum_{j=1}^{q} \phi_i(|z_{t-j}| - E[|z_{t-i}|]) + \sum_{j=1}^{p} \beta_i ln(h_{t-j}), \quad t = 1, \ldots, T \qquad (1)$$

where $z_t = \frac{\epsilon_t}{\sqrt{h_t}}$, $E[|z_{t-i}|]$ denotes the expected value of $|z_{t-i}|$ and $\epsilon_t|\psi_{t-1} = N(0, h_t)$ or $\epsilon_t|\psi_{t-1} = S_t(df, h_t)$. Here S_t is a standardised Student's t-distribution with df degrees of freedom and variance h_t, T is the number of terms in the sequence, y_t denotes the endogenous variables, x_t the exogenous variables, b_o the regression mean, b the regression coefficients, ϵ_t the residuals, h_t is the conditional variance, df the number of degrees of freedom of the Student's t-distribution, and ψ_t the set of all information up to time t.

G13FGF provides an estimate for $\hat{\theta}$, the vector $\theta = (b_o, b^T, \omega^T)$ where $b^T = (b_1, \ldots, b_k)$, $\omega^T = (\alpha_0, \alpha_1, \ldots, \alpha_q, \phi_1, \ldots, \phi_q, \phi_1, \ldots, \phi_q, \gamma)$ when DIST = 'N', and $\omega^T = (\alpha_0, \alpha_1, \ldots, \alpha_q, \phi_1, \ldots, \phi_q, \phi_1, \ldots, \phi_q, \gamma, df)$ when DIST = 'T'.

MN, NREG can be used to simplify the GARCH(p, q) expression in (1) as follows:

No Regression and No Mean

$y_t = \epsilon_t$,

MN $= 0$,

NREG $= 0$ and

θ is a $(2 \times q + p + 2)$ vector when DIST = 'N', and a $(2 \times q + p + 3)$ vector, when DIST = 'T'.

No Regression

$y_t = b_o + \epsilon_t$,

MN = 1,

NREG = 0 and

θ is a $(2 \times q + p + 3)$ vector when DIST = 'N' and a $(2 \times q + p + 4)$ vector, when DIST = 'T'.

Note: if the $y_t = \mu + \epsilon_t$, where μ is known (not to be estimated by G13FGF) then (1) can be written as $y_t^{\mu} = \epsilon_t$, where $y_t^{\mu} = y_t - \mu$. This corresponds to the case **No Regression and No Mean**, with y_t replaced by $y_t - \mu$.

No Mean

$y_t = x_t^T b + \epsilon_t$,

MN = 0,

NREG = k and

θ is a $(2 \times q + p + 2 + k)$ vector when DIST = 'N' and a $(2 \times q + p + 3 + k)$ vector, when DIST = 'T'.

4 References

Engle R (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation *Econometrica* **50** 987–1008

Bollerslev T (1986) Generalised autoregressive conditional heteroskedasticity *Journal of Econometrics* **31** 307–327

Engle R and Ng V (1993) Measuring and Testing the Impact of News on Volatility *Journal of Finance* **48** 1749–1777

Hamilton J (1994) *Time Series Analysis* Princeton University Press

Glosten L, Jagannathan R and Runkle D (1993) Relationship between the expected value and the volatility of nominal excess return on stocks *Journal of Finance* **48** 1779–1801

5 Parameters

1: DIST – CHARACTER*1 *Input*

On entry: the type of distribution to use for e_t.

If DIST = 'N', a Normal distribution is used.

If DIST = 'T', a Student's t-distribution is used.

Constraint: DIST = 'N' or 'T'.

2: YT(NUM) – ***real*** array *Input*

On entry: the sequence of observations, y_t, for $t = 1, \ldots, T$.

3: X(LDX,*) – ***real*** array *Input*

Note: the second dimension of the array X must be at least $\max(1, \mathrm{NREG} + \mathrm{MN})$.

On entry: row t of X must contain the time dependent exogenous vector x_t, where $x_t^T = (x_t^1, \ldots, x_t^k)$, for $t = 1, \ldots, T$.

4: LDX – INTEGER *Input*

On entry: the first dimension of the array X as declared in the (sub)program from which G13FGF is called.

Constraint: LDX $\geq$ NUM.

5: NUM – INTEGER *Input*

On entry: the number of terms in the sequence, T.

Constraint: NUM $\geq \max(\mathrm{IP}, \mathrm{IQ})$.

6: IP – INTEGER *Input*

On entry: the number of coefficients, β_i, for $i = 1, \ldots, p$.

Constraint: IP ≥ 0 (see also NPAR).

7: IQ – INTEGER *Input*

On entry: the number of coefficients, α_i, for $i = 1, \ldots, q$.

Constraint: IQ ≥ 1 (see also NPAR).

8: NREG – INTEGER *Input*

On entry: the number of regression coefficients, k.

Constraint: NREG ≥ 0 (see also NPAR).

9: MN – INTEGER *Input*

On entry: if MN $= 1$, the mean term b_0 will be included in the model.

Constraint: MN $= 0$ or 1.

10: NPAR – INTEGER *Input*

On entry: the number of parameters to be included in the model. NPAR $= 1 + 2 \times \mathrm{IQ} + \mathrm{IP} + \mathrm{MN} + \mathrm{NREG}$ when DIST = 'N' and NPAR $= 2 + 2 \times \mathrm{IQ} + \mathrm{IP} + \mathrm{MN} + \mathrm{NREG}$ when DIST = 'T'.

Constraint: NPAR < 20.

11: THETA(NPAR) – ***real*** array *Input/Output*

On entry: the initial parameter estimates for the vector θ. The first element must contain the coefficient α_o and the next IQ elements must contain the autoregressive coefficients α_i, for $i = 1, \ldots, q$. The next IQ elements contain the coefficients ϕ_i, for $i = 1, \ldots, q$. The next IP elements must contain the moving average coefficients β_j, for $j = 1, \ldots, p$. If DIST = 'T', then the next element must contain an estimate for df, the number of degrees of freedom of the Student's t-distribution. If MN $= 1$, then the next element must contain the mean term b_o. If COPT = .FALSE., then the remaining NREG elements are taken as initial estimates of the linear regression coefficients b_i, for $i = 1, \ldots, k$.

On exit: the estimated values $\hat{\theta}$ for the vector θ. The first element contains the coefficient α_o, The first element contains the coefficient α_o, next IQ elements contain the coefficients ϕ_i, for $i = 1, \ldots, q$. The next IP elements are the moving average coefficients β_j, for $j = 1, \ldots, p$. If DIST = 'T' then the next element contains an estimate for df, the number of degrees of freedom of the Student's t-distribution. If MN $= 1$ then the next element contains an estimate for the mean term b_o. The final NREG elements are the estimated linear regression coefficients b_i, for $i = 1, \ldots, k$.

12: SE(NPAR) – ***real*** array *Output*

On exit: the standard errors for $\hat{\theta}$. The first element contains the standard error for α_o, the next IQ elements contain the standard errors for α_i, for $i = 1, \ldots, q$, the next IQ elements contain the standard errors for ϕ_i, for $i = 1, \ldots, q$, the next IP elements are the standard errors for β_j, for $j = 1, \ldots, p$. If DIST = 'T' then the next element contains the standard error for df, the number of degrees of freedom of the Student's t-distribution. If MN = 1 then the next element contains the standard error for b_o. The final NREG elements are the standard errors for b_j, for $j = 1, \ldots, k$.

13: SC(NPAR) – ***real*** array *Output*

On exit: the scores for $\hat{\theta}$. The first element contains the scores for α_o, the next IQ elements contain the scores for α_i, for $i = 1, \ldots, q$, the next IQ elements contain the scores for ϕ_i, for $i = 1, \ldots, q$, the next IP elements are the scores for β_j, for $j = 1, \ldots, p$. If DIST = 'T' then the next element contains the scores for df, the number of degrees of freedom of the Student's t-distribution. If MN = 1 then the next element contains the score for b_o. The final NREG elements are the scores for b_j, for $j = 1, \ldots, k$.

14: COVAR(LDC,NPAR) – ***real*** array *Output*

On exit: the covariance matrix of the parameter estimates $\hat{\theta}$, that is the inverse of the Fisher Information Matrix.

15: LDC – INTEGER *Input*

On entry: the first dimension of the array COVAR as declared in the (sub)program from which G13FGF is called.

Constraint: LDC $\geq$ NPAR.

16: HP – ***real*** *Input/Output*

On entry: if COPT = .FALSE. then HP is the value to be used for the pre-observed conditional variance, otherwiseHP is not referenced.

On exit: if COPT = .TRUE. then HP is the estimated value of the pre-observed conditional variance.

17: ET(NUM) – ***real*** array *Output*

On exit: the estimated residuals, ϵ_t, for $t = 1, \ldots, T$.

18: HT(NUM) – ***real*** array *Output*

On exit: the estimated conditional variances, h_t, for $t = 1, \ldots, T$.

19: LGF – ***real*** *Output*

On exit: the value of the log likelihood function at $\hat{\theta}$.

20: COPTS – LOGICAL *Input*

On entry: If COPT = .TRUE. then the routine provides initial parameter estimates of the regression terms, otherwise these are provided by the user.

21: MAXIT – INTEGER *Input*

On entry: the maximum number of iterations to be used by the optimization routine when estimating the GARCH(p, q) parameters. If MAXIT is set to 0 then the standard errors, score vector and variance-covariance are calculated for the input value of θ in THETA; however the value of θ is not updated.

Constraint: MAXIT $>$ 0.

22: TOL – ***real*** *Input*

On entry: the tolerance to be used by the optimization routine when estimating the GARCH(p,q) parameters.

23: WORK(LWORK) – ***real*** array *Workspace*
24: LWORK – INTEGER *Input*

On entry: the dimension of the array WORK as declared in the (sub)program from which G13FGF is called.

Constraint: LWORK $\geq$ (NREG + 3) $\times$ NUM + 3.

25: 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, 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, NREG < 0,
or MN > 1,
or MN < 0,
or IQ < 1,
or IP < 0,
or NPAR $\geq$ 20,
or NPAR has an invalid value,
or LDC < NPAR,
or LDX < NUM,
or DIST $\neq$ 'N',
or DIST $\neq$ 'T',
or MAXIT $\leq$ 0,
or NUM < max(IP, IQ).

IFAIL = 2

On entry, LWORK < (NREG + 3) $\times$ NUM + 3.

IFAIL = 3

The matrix X is not full rank.

IFAIL = 4

The information matrix is not positive definite.

IFAIL = 5

The maximum number of iterations has been reached.

IFAIL = 6

The log-likelihood cannot be optimised any further.

IFAIL = 7

No feasible model parameters could be found.

7 Accuracy

Not applicable.

8 Further Comments

None.

9 Example

This example program uses G05HNF to generate 1500 data points, with known process parameters θ for the following two time-series:

(i) A GARCH(1,1) sequence with normally distributed residuals.

(ii) A GARCH(1,2) sequence with Student's t-distributed residuals.

Here G05HNF is initially called, with the output discarded, to eliminate 'start-up effects' in these sequences. The process parameter estimates, $\hat{\theta}$, are then obtained using G13FGF, and compared with their true values, θ. Finally a four step ahead volatility estimate is computed using G13FHF.

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.

```
*     G13FGF Example Program Text
*     Mark 20 Release. NAG Copyright 2001.
*     .. Parameters ..
      INTEGER          NOUT
      PARAMETER        (NOUT=6)
      INTEGER          NPARMX, NUM
      real             ZERO
      PARAMETER        (NPARMX=10,NUM=1500,ZERO=0.0e0)
      INTEGER          NUM1, NREGMX
      PARAMETER        (NUM1=3000,NREGMX=10)
*     .. Local Scalars ..
      real             DF, FAC1, HP, LGF, MEAN, TOL, XTERM
      INTEGER          I, IFLAG, IGEN, IP, IQ, K, LDX, LWK, MAXIT, MN,
     +                 NPAR, NPAR2, NREG, NT
      LOGICAL          COPT, FCALL
      CHARACTER        DIST
*     .. Local Arrays ..
      real             BX(10), COVAR(NPARMX,NPARMX), CVAR(100),
     +                 ETM(NUM1), HT(NUM1+10), HTM(NUM1), PARAM(NPARMX),
     +                 RVEC(40), RWSAV(9), SC(NPARMX), SE(NPARMX),
     +                 THETA(NPARMX), WK(NUM1*3+NPARMX+NREGMX*NUM1+20*
     +                 20+1), X(NUM1,10), YT(NUM1+10)
      INTEGER          ISEED(4)
*     .. External Subroutines ..
      EXTERNAL         G05HNF, G05KBF, G13FGF, G13FHF
*     .. Intrinsic Functions ..
      INTRINSIC        real, SIN
*     .. Executable Statements ..

      WRITE (NOUT,*) 'G13FGF Example Program Results'

      ISEED(1) = 111
```

```
      IGEN = 0

      LDX = NUM1
      BX(1) = 1.5e0
      BX(2) = 2.5e0
      BX(3) = 3.0e0
      MEAN = 3.0e0

      DO 20 I = 1, NUM
         FAC1 = real(I)*0.01e0
         X(I,1) = 0.01e0 + 0.7e0*SIN(FAC1)
         X(I,2) = 0.5e0 + FAC1*0.1e0
         X(I,3) = 1.0e0
   20 CONTINUE

      NREG = 2
      MN = 1
      IP = 1
      IQ = 1
      NPAR = IP + 2*IQ + 1

      PARAM(1) = 0.1e0
      PARAM(2) = -0.3e0
      PARAM(3) = 0.1e0
      PARAM(4) = 0.9e0

      DF = 5.0e0
      DIST = 'N'

      FCALL = .TRUE.
      CALL G05KBF(IGEN,ISEED)
      CALL G05HNF(DIST,800,IP,IQ,PARAM,DF,HT,YT,FCALL,RVEC,IGEN,ISEED,
     +            RWSAV,IFLAG)
      FCALL = .FALSE.
      CALL G05HNF(DIST,NUM,IP,IQ,PARAM,DF,HT,YT,FCALL,RVEC,IGEN,ISEED,
     +            RWSAV,IFLAG)

      IFLAG = -1
      DO 60 I = 1, NUM
         XTERM = ZERO
         DO 40 K = 1, NREG
            XTERM = XTERM + X(I,K)*BX(K)
   40    CONTINUE
         IF (MN.EQ.1) THEN
            YT(I) = MEAN + XTERM + YT(I)
         ELSE
            YT(I) = XTERM + YT(I)
         END IF
   60 CONTINUE

      COPT = .TRUE.

      MAXIT = 50
      TOL = 1.0e-5

      DO 80 I = 1, NPAR
         THETA(I) = PARAM(I)*0.5e0
   80 CONTINUE

      IF (MN.EQ.1) THEN
         THETA(NPAR+MN) = MEAN*0.5e0
      END IF

      DO 100 I = 1, NREG
         THETA(NPAR+MN+I) = BX(I)*0.5e0
  100 CONTINUE

      LWK = NREG*NUM + 3*NUM + 3
      NPAR2 = 1 + IQ*2 + IP + MN + NREG

      CALL G13FGF(DIST,YT,X,LDX,NUM,IP,IQ,NREG,MN,NPAR2,THETA,SE,SC,
```

```
     +            COVAR,NPARMX,HP,ETM,HTM,LGF,COPT,MAXIT,TOL,WK,LWK,
     +            IFLAG)

      WRITE (NOUT,*)
      WRITE (NOUT,*) 'Normal distribution'
      WRITE (NOUT,*)
      WRITE (NOUT,*) '          Parameter       Standard        Correct'
      WRITE (NOUT,*) '          estimates       errors          values'

      DO 120 I = 1, NPAR
         WRITE (NOUT,99999) THETA(I), SE(I), PARAM(I)
  120 CONTINUE

      IF (MN.EQ.1) THEN
         WRITE (NOUT,99999) THETA(NPAR+1), SE(NPAR+1), MEAN
      END IF

      DO 140 I = 1, NREG
         WRITE (NOUT,99999) THETA(NPAR+MN+I), SE(NPAR+MN+I), BX(I)
  140 CONTINUE

      NT = 4
      CALL G13FHF(NUM,NT,IP,IQ,THETA,CVAR,HTM,ETM,IFLAG)

      WRITE (NOUT,*)
      WRITE (NOUT,99998) 'Volatility forecast = ', CVAR(NT)
      WRITE (NOUT,*)

      NREG = 2
      MN = 1
      IP = 1
      IQ = 2
      NPAR = IP + 2*IQ + 1

      PARAM(1) = 0.1e0
      PARAM(2) = -0.3e0
      PARAM(3) = -0.1e0
      PARAM(4) = 0.1e0
      PARAM(5) = 0.3e0
      PARAM(6) = 0.7e0

      DIST = 'T'

      FCALL = .TRUE.
      ISEED(1) = 111
      CALL G05KBF(IGEN,ISEED)
      CALL G05HNF(DIST,NUM,IP,IQ,PARAM,DF,HT,YT,FCALL,RVEC,IGEN,ISEED,
     +            RWSAV,IFLAG)
      FCALL = .FALSE.
      CALL G05HNF(DIST,NUM,IP,IQ,PARAM,DF,HT,YT,FCALL,RVEC,IGEN,ISEED,
     +            RWSAV,IFLAG)

      IFLAG = -1
      DO 200 I = 1, NUM
         XTERM = ZERO
         DO 180 K = 1, NREG
            XTERM = XTERM + X(I,K)*BX(K)
  180    CONTINUE
         IF (MN.EQ.1) THEN
            YT(I) = MEAN + XTERM + YT(I)
         ELSE
            YT(I) = XTERM + YT(I)
         END IF
  200 CONTINUE

      COPT = .TRUE.
      MAXIT = 50
      TOL = 1.0e-5

      DO 220 I = 1, NPAR
         THETA(I) = PARAM(I)*0.5e0
```

```
  220 CONTINUE

      THETA(NPAR+1) = DF*0.65e0

      IF (MN.EQ.1) THEN
         THETA(NPAR+1+MN) = MEAN*0.5e0
      END IF

      DO 240 I = 1, NREG
         THETA(NPAR+1+MN+I) = BX(I)*0.5e0
  240 CONTINUE

      LWK = NREG*NUM + 3*NUM + 3
      NPAR2 = 2 + IQ*2 + IP + MN + NREG

      CALL G13FGF(DIST,YT,X,LDX,NUM,IP,IQ,NREG,MN,NPAR2,THETA,SE,SC,
     +            COVAR,NPARMX,HP,ETM,HTM,LGF,COPT,MAXIT,TOL,WK,LWK,
     +            IFLAG)

      WRITE (NOUT,*)
      WRITE (NOUT,*) 'Student t-distribution'
      WRITE (NOUT,*)
      WRITE (NOUT,*) '          Parameter       Standard        Correct'
      WRITE (NOUT,*) '          estimates       errors          values'

      DO 260 I = 1, NPAR
         WRITE (NOUT,99999) THETA(I), SE(I), PARAM(I)
  260 CONTINUE

      WRITE (NOUT,99999) THETA(NPAR+1), SE(NPAR+1), DF

      IF (MN.EQ.1) THEN
         WRITE (NOUT,99999) THETA(NPAR+1+MN), SE(NPAR+1+MN), MEAN
      END IF

      DO 280 I = 1, NREG
         WRITE (NOUT,99999) THETA(NPAR+1+MN+I), SE(NPAR+1+MN+I), BX(I)
  280 CONTINUE

      NT = 4
      CALL G13FHF(NUM,NT,IP,IQ,THETA,CVAR,HTM,ETM,IFLAG)

      WRITE (NOUT,*)
      WRITE (NOUT,99998) 'Volatility forecast = ', CVAR(NT)
      WRITE (NOUT,*)
      STOP
*
99999 FORMAT (1X,3F16.4)
99998 FORMAT (1X,A,F12.4)
      END
```

9.2 Program Data

None.

9.3 Program Results

```
 G13FGF Example Program Results

 Normal distribution

           Parameter       Standard        Correct
           estimates       errors          values
           0.1153          0.0202          0.1000
          -0.3097          0.0268         -0.3000
           0.1209          0.0444          0.1000
           0.8936          0.0143          0.9000
           2.8629          0.0948          3.0000
           1.4524          0.0588          1.5000
           2.5812          0.0622          2.5000
```

```
Volatility forecast =       2.9479


Student t-distribution

          Parameter        Standard        Correct
          estimates        errors          values
          0.0575           0.0299          0.1000
         -0.3275           0.0476         -0.3000
         -0.0838           0.0666         -0.1000
          0.0400           0.0693          0.1000
          0.2601           0.0751          0.3000
          0.6797           0.0673          0.7000
          5.0868           0.6308          5.0000
          3.0098           0.0627          3.0000
          1.4348           0.0415          1.5000
          2.4781           0.0471          2.5000

Volatility forecast =       1.1276
```