G02DNF – NAG Fortran Library Routine Document

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

G02DNF gives the estimate of an estimable function along with its standard error.

2 Specification

```
 SUBROUTINE G02DNF(IP, IRANK, B, COV, P, F, EST, STAT, SESTAT, T,
1                  TOL, WK, IFAIL)
 INTEGER           IP, IRANK, IFAIL
 real              B(IP), COV(IP*(IP+1)/2), P(IP*IP+2*IP), F(IP),
1                  STAT, SESTAT, T, TOL, WK(IP)
 LOGICAL           EST
```

3 Description

This routine computes the estimates of an estimable function for a general linear regression model which is not of full rank. It is intended for use after a call to G02DAF or G02DDF. An estimable function is a linear combination of the parameters such that it has a unique estimate. For a full rank model all linear combinations of parameters are estimable.

In the case of a model not of full rank the routines use a singular value decomposition (SVD) to find the the parameter estimates, $\hat{\beta}$, and their variance-covariance matrix. Given the upper triangular matrix R obtained from the QR decomposition of the independent variables the SVD gives:

$$R = Q_* \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} P^T,$$

where D is a k by k diagonal matrix with non-zero diagonal elements, k being the rank of R, and Q_* and P are p by p orthogonal matrices. This leads to a solution:

$$\hat{\beta} = P_1 D^{-1} Q_{*_1}^T c_1$$

P_1 being the first k columns of P, i.e., $P = (P_1 P_0)$, Q_{*_1} being the first k columns of Q_* and c_1 being the first p elements of c.

Details of the SVD, are made available, in the form of the matrix P^*:

$$P^* = \begin{pmatrix} D^{-1} P_1^T \\ P_0^T \end{pmatrix}$$

as given by G02DAF and G02DDF.

A linear function of the parameters, $F = f^T\beta$, can be tested to see if it is estimable by computing $\zeta = P_0^T f$. If ζ is zero, then the function is estimable, if not, the function is not estimable. In practice $|\zeta|$ is tested against some small quantity η.

Given that F is estimable it can be estimated by $f^T\hat{\beta}$ and its standard error calculated from the variance-covariance matrix of $\hat{\beta}$, C_β, as

$$\text{se}(F) = \sqrt{f^T C_\beta f}$$

Also a t-statistic:

$$t = \frac{f^T\hat{\beta}}{\text{se}(F)},$$

can be computed. The t-statistic will have a Student's t-distribution with degrees of freedom as given by the degrees of freedom for the residual sum of squares for the model.

4 References

[1] Golub G H and van Loan C F (1996) *Matrix Computations* Johns Hopkins University Press (3rd Edition), Baltimore

[2] Hammarling S (1985) The singular value decomposition in multivariate statistics *SIGNUM Newsl.* **20 (3)** 2–25

[3] Searle S R (1971) *Linear Models* Wiley

5 Parameters

1: IP — INTEGER *Input*

On entry: the number of terms in the linear model, p.

Constraint: IP ≥ 1.

2: IRANK — INTEGER *Input*

On entry: the rank of the independent variables, k.

Constraint: $1 \leq$ IRANK $\leq$ IP.

3: B(IP) — ***real*** array *Input*

On entry: the IP values of the estimates of the parameters of the model, $\hat{\beta}$.

4: COV(IP∗(IP+1)/2) — ***real*** array *Input*

On entry: the upper triangular part of the variance-covariance matrix of the IP parameter estimates given in B. They are stored packed by column, i.e., the covariance between the parameter estimate given in B(i) and the parameter estimate given in B(j), $j \geq i$, is stored in COV($j \times (j-1)/2 + i$).

5: P(IP∗IP+2∗IP) — ***real*** array *Input*

On entry: P as returned by G02DAF or G02DDF.

6: F(IP) — ***real*** array *Input*

On entry: the linear function to be estimated, f.

7: EST — LOGICAL *Output*

On exit: EST indicates if the function was estimable.

If EST = .TRUE., then the function is estimable.
If EST = .FALSE., the function is not estimable and STAT, SESTAT and T are not set.

8: STAT — ***real*** *Output*

On exit: if EST = .TRUE., STAT contains the estimate of the function, $f^{\mathrm{T}}\hat{\beta}$.

9: SESTAT — ***real*** *Output*

On exit: if EST = .TRUE., SESTAT contains the standard error of the estimate of the function, se(F).

10: T — ***real*** *Output*

On exit: if EST = .TRUE., T contains the t-statistic for the test of the function being equal to zero.

11: TOL — ***real*** *Input*

On entry: TOL is the tolerance value used in the check for estimability, η.

If TOL ≤ 0.0, then $\sqrt{\epsilon}$, where ϵ is the ***machine precision***, is used instead.

12: WK(IP) — ***real*** array *Workspace*

13: 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 or gives a warning (see Section 6).

For this routine, because the values of output parameters may be useful even if IFAIL ≠ 0 on exit, users are recommended to set IFAIL to −1 before entry. **It is then 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 specified by the routine:

IFAIL = 1

On entry, IP < 1,
or IRANK < 1,
or IRANK > IP,

IFAIL = 2

On entry, IRANK = IP. In this case EST is returned as true and all statistics are calculated.

IFAIL = 3

Standard error of statistic = 0.0, this may be due to rounding errors if the standard error is very small or due to miss-specified inputs COV and F.

7 Accuracy

The computations are believed to be stable.

8 Further Comments

The value of estimable functions is independent of the solution chosen from the many possible solutions. While G02DNF may be used to estimate functions of the parameters of the model as computed by G02DKF, β_c, these must be expressed in terms of the original parameters, β. The relation between the two sets of parameters may not be straightforward.

9 Example

Data from an experiment with four treatments and three observations per treatment are read in. A model, with a mean term, is fitted by G02DAF. The number of functions to be tested is read in, then the linear functions themselves are read in and tested with G02DNF. The results of G02DNF are printed.

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.

```
*     G02DNF Example Program Text
*     Mark 14 Release.  NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          MMAX, NMAX
```

```
      PARAMETER        (MMAX=5,NMAX=12)
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
*     .. Local Scalars ..
      real             RSS, SESTAT, STAT, T, TOL
      INTEGER          I, IDF, IFAIL, IP, IRANK, J, M, N, NESTFN
      LOGICAL          EST, SVD
      CHARACTER        MEAN, WEIGHT
*     .. Local Arrays ..
      real             B(MMAX), COV((MMAX*MMAX+MMAX)/2), F(MMAX),
     +                 H(NMAX), P(MMAX*(MMAX+2)), Q(NMAX,MMAX+1),
     +                 RES(NMAX), SE(MMAX), WK(MMAX*MMAX+5*(MMAX-1)),
     +                 WT(NMAX), X(NMAX,MMAX), Y(NMAX)
      INTEGER          ISX(MMAX)
*     .. External Subroutines ..
      EXTERNAL         G02DAF, G02DNF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G02DNF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) N, M, WEIGHT, MEAN
      IF (N.LE.NMAX .AND. M.LT.MMAX) THEN
         IF (WEIGHT.EQ.'W' .OR. WEIGHT.EQ.'w') THEN
            DO 20 I = 1, N
               READ (NIN,*) (X(I,J),J=1,M), Y(I), WT(I)
   20       CONTINUE
         ELSE
            DO 40 I = 1, N
               READ (NIN,*) (X(I,J),J=1,M), Y(I)
   40       CONTINUE
         END IF
         READ (NIN,*) (ISX(J),J=1,M), IP
*        Set tolerance
         TOL = 0.00001e0
         IFAIL = 0
*
*        Find initial estimates using G02DAF
         CALL G02DAF(MEAN,WEIGHT,N,X,NMAX,M,ISX,IP,Y,WT,RSS,IDF,B,SE,
     +              COV,RES,H,Q,NMAX,SVD,IRANK,P,TOL,WK,IFAIL)
*
         WRITE (NOUT,*)
         WRITE (NOUT,*) 'Estimates from G02DAF'
         WRITE (NOUT,*)
         WRITE (NOUT,99999) 'Residual sum of squares = ', RSS
         WRITE (NOUT,99998) 'Degrees of freedom = ', IDF
         WRITE (NOUT,*)
         WRITE (NOUT,*) 'Variable   Parameter estimate   Standard error'
         WRITE (NOUT,*)
         DO 60 J = 1, IP
            WRITE (NOUT,99997) J, B(J), SE(J)
   60    CONTINUE
         READ (NIN,*) NESTFN
         DO 80 I = 1, NESTFN
            READ (NIN,*) (F(J),J=1,IP)
            IFAIL = -1
*
            CALL G02DNF(IP,IRANK,B,COV,P,F,EST,STAT,SESTAT,T,TOL,WK,
     +                  IFAIL)
```

```
*
            IF (IFAIL.EQ.0 .OR. IFAIL.EQ.2) THEN
               WRITE (NOUT,*)
               WRITE (NOUT,99996) 'Function ', I
               WRITE (NOUT,*)
               WRITE (NOUT,99995) (F(J),J=1,IP)
               WRITE (NOUT,*)
               IF (EST) THEN
                  WRITE (NOUT,99994) 'STAT = ', STAT, ' SE = ', SESTAT,
     +              ' T = ', T
               ELSE
                  WRITE (NOUT,*) 'Function not estimable'
               END IF
            END IF
   80    CONTINUE
      END IF
      STOP
*
99999 FORMAT (1X,A,e12.4)
99998 FORMAT (1X,A,I4)
99997 FORMAT (1X,I6,2e20.4)
99996 FORMAT (1X,A,I4)
99995 FORMAT (1X,5F8.2)
99994 FORMAT (1X,A,F10.4,A,F10.4,A,F10.4)
      END
```

9.2 Program Data

```
G02DNF Example Program Data
 12 4 'U' 'M'
1.0 0.0 0.0 0.0 33.63
0.0 0.0 0.0 1.0 39.62
0.0 1.0 0.0 0.0 38.18
0.0 0.0 1.0 0.0 41.46
0.0 0.0 0.0 1.0 38.02
0.0 1.0 0.0 0.0 35.83
0.0 0.0 0.0 1.0 35.99
1.0 0.0 0.0 0.0 36.58
0.0 0.0 1.0 0.0 42.92
1.0 0.0 0.0 0.0 37.80
0.0 0.0 1.0 0.0 40.43
0.0 1.0 0.0 0.0 37.89
 1   1   1   1   5
 3
 1.0 1.0  0.0 0.0 0.0
 0.0 1.0 -1.0 0.0 0.0
 0.0 1.0  0.0 0.0 0.0
```

9.3 Program Results

```
G02DNF Example Program Results

Estimates from G02DAF

Residual sum of squares =   0.2223E+02
Degrees of freedom =    8

Variable   Parameter estimate   Standard error

     1          0.3056E+02          0.3849E+00
     2          0.5447E+01          0.8390E+00
     3          0.6743E+01          0.8390E+00
     4          0.1105E+02          0.8390E+00
     5          0.7320E+01          0.8390E+00

Function    1

    1.00    1.00    0.00    0.00    0.00

STAT =    36.0033 SE =     0.9623 T =    37.4119

Function    2

    0.00    1.00   -1.00    0.00    0.00

STAT =    -1.2967 SE =     1.3610 T =    -0.9528

Function    3

    0.00    1.00    0.00    0.00    0.00

Function not estimable
```