NAG Library Routine Document

G02GKF

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

G02GKF calculates the estimates of the parameters of a generalized linear model for given constraints from the singular value decomposition results.

2 Specification

```
      SUBROUTINE G02GKF(IP, ICONST, V, LDV, C, LDC, B, S, SE, COV, WK, IFAIL)
      INTEGER            IP, ICONST, LDV, LDC, IFAIL
      double precision   V(LDV,IP+7), C(LDC,ICONST), B(IP), S, SE(IP),
     1                   COV(IP*(IP+1)/2),
     2                   WK(2*IP*IP+IP*ICONST+2*ICONST*ICONST+4*ICONST)
```

3 Description

G02GKF computes the estimates given a set of linear constraints for a generalized linear model which is not of full rank. It is intended for use after a call to G02GAF, G02GBF, G02GCF or G02GDF.

In the case of a model not of full rank the routines use a singular value decomposition to find the parameter estimates, $\hat{\beta}_{\mathrm{svd}}$, and their variance-covariance matrix. Details of the SVD are made available in the form of the matrix P^*:

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

as described by G02GAF, G02GBF, G02GCF and G02GDF. Alternative solutions can be formed by imposing constraints on the parameters. If there are p parameters and the rank of the model is k then $n_{\mathrm{c}} = p - k$ constraints will have to be imposed to obtain a unique solution.

Let C be a p by n_{c} matrix of constraints, such that

$$C^{\mathrm{T}}\beta = 0,$$

then the new parameter estimates $\hat{\beta}_{\mathrm{c}}$ are given by:

$$\begin{aligned}\hat{\beta}_{\mathrm{c}} &= A\hat{\beta}_{\mathrm{svd}} \\ &= \left(I - P_0\left(C^{\mathrm{T}}P_0\right)^{-1}\right)\hat{\beta}_{\mathrm{svd}}, \qquad \text{where } I \text{ is the identity matrix,}\end{aligned}$$

and the variance-covariance matrix is given by

$$AP_1D^{-2}P_1^{\mathrm{T}}A^{\mathrm{T}}$$

provided $\left(C^{\mathrm{T}}P_0\right)^{-1}$ exists.

4 References

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

McCullagh P and Nelder J A (1983) *Generalized Linear Models* Chapman and Hall

Searle S R (1971) *Linear Models* Wiley

5 Parameters

1: IP – INTEGER *Input*

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

Constraint: $\text{IP} \geq 1$.

2: ICONST – INTEGER *Input*

On entry: the number of constraints to be imposed on the parameters, n_c.

Constraint: $0 < \text{ICONST} < \text{IP}$.

3: $\text{V}(\text{LDV},\text{IP}+7)$ – ***double precision*** array *Input*

On entry: the array V as returned by G02GAF, G02GBF, G02GCF or G02GDF.

4: LDV – INTEGER *Input*

On entry: the first dimension of the array V as declared in the (sub)program from which G02GKF is called.

Constraint: $\text{LDV} \geq \text{IP}$. LDV should be as supplied to G02GAF, G02GBF, G02GCF or G02GDF

5: $\text{C}(\text{LDC},\text{ICONST})$ – ***double precision*** array *Input*

On entry: contains the ICONST constraints stored by column, i.e., the ith constraint is stored in the ith column of C.

6: LDC – INTEGER *Input*

On entry: the first dimension of the array C as declared in the (sub)program from which G02GKF is called.

Constraint: $\text{LDC} \geq \text{IP}$.

7: $\text{B}(\text{IP})$ – ***double precision*** array *Input/Output*

On entry: the parameter estimates computed by using the singular value decomposition, $\hat{\beta}_{\text{svd}}$.

On exit: the parameter estimates of the parameters with the constraints imposed, $\hat{\beta}_c$.

8: S – ***double precision*** *Input*

On entry: the estimate of the scale parameter.

For results from G02GAF and G02GDF then S is the scale parameter for the model, σ^2.

For results from G02GBF and G02GCF then S should be set to 1.0.

Constraint: $\text{S} > 0.0$.

9: $\text{SE}(\text{IP})$ – ***double precision*** array *Output*

On exit: the standard error of the parameter estimates in B.

10: $\text{COV}(\text{IP} \times (\text{IP}+1)/2)$ – ***double precision*** array *Output*

On exit: 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 $\text{B}(i)$ and the parameter estimate given in $\text{B}(j)$, $j \geq i$, is stored in $\text{COV}(j \times (j-1)/2+i)$.

11: $\text{WK}(2 \times \text{IP} \times \text{IP} + \text{IP} \times \text{ICONST} + 2 \times \text{ICONST} \times \text{ICONST} + 4 \times \text{ICONST})$ – ***double precision*** array *Workspace*

Note: a simple upper bound for the size of the workspace is $5 \times \text{IP} \times \text{IP} + 4 \times \text{IP}$.

12: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction 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, if you are 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, IP < 1.
or ICONST $\geq$ IP,
or ICONST ≤ 0,
or LDV $<$ IP,
or LDC $<$ IP,
or S ≤ 0.0.

IFAIL $= 2$

C does not give a model of full rank.

7 Accuracy

It should be noted that due to rounding errors a parameter that should be zero when the constraints have been imposed may be returned as a value of order ***machine precision***.

8 Further Comments

G02GKF is intended for use in situations in which dummy $(0-1)$ variables have been used such as in the analysis of designed experiments when you do not wish to change the parameters of the model to give a full rank model. The routine is not intended for situations in which the relationships between the independent variables are only approximate.

9 Example

A loglinear model is fitted to a 3 by 5 contingency table by G02GCF. The model consists of terms for rows and columns. The table is

$$\begin{array}{rrrrr} 141 & 67 & 114 & 79 & 39 \\ 131 & 66 & 143 & 72 & 35\,. \\ 36 & 14 & 38 & 28 & 16 \end{array}$$

The constraints that the sum of row effects and the sum of column effects are zero are then read in and the parameter estimates with these constraints imposed are computed by G02GKF and printed.

9.1 Program Text

```
*     G02GKF Example Program Text
*     Mark 14 Release. NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          NMAX, MMAX, LDV, LDC
      PARAMETER        (NMAX=15,MMAX=9,LDV=NMAX,LDC=MMAX)
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
*     .. Local Scalars ..
      DOUBLE PRECISION A, DEV, EPS, TOL
      INTEGER          I, ICONST, IDF, IFAIL, IP, IPRINT, IRANK, J, M,
     +                 MAXIT, N
*     .. Local Arrays ..
      DOUBLE PRECISION B(MMAX), C(LDC,MMAX), COV((MMAX*MMAX+MMAX)/2),
     +                 SE(MMAX), V(LDV,7+MMAX), WK(5*MMAX*MMAX),
     +                 WT(NMAX), X(NMAX,MMAX), Y(NMAX)
      INTEGER          ISX(MMAX)
*     .. External Subroutines ..
      EXTERNAL         G02GCF, G02GKF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G02GKF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) N, M, IPRINT
      IF (N.LE.NMAX .AND. M.LT.MMAX) THEN
         DO 20 I = 1, N
            READ (NIN,*) (X(I,J),J=1,M), Y(I)
   20    CONTINUE
         READ (NIN,*) (ISX(J),J=1,M), IP
*        Set control parameters
         EPS = 0.000001D0
         TOL = 0.00005D0
         MAXIT = 10
         IFAIL = 1
*
*        Fit Log-linear model using G02GCF
         CALL G02GCF('L','M','N','U',N,X,NMAX,M,ISX,IP,Y,WT,A,DEV,IDF,B,
     +               IRANK,SE,COV,V,LDV,TOL,MAXIT,IPRINT,EPS,WK,IFAIL)
*
         IF (IFAIL.EQ.0 .OR. IFAIL.GE.7) THEN
            WRITE (NOUT,*)
            WRITE (NOUT,99999) 'Deviance = ', DEV
            WRITE (NOUT,99998) 'Degrees of freedom = ', IDF
            WRITE (NOUT,*)
*           Input constraints
            ICONST = IP - IRANK
            DO 40 I = 1, IP
               READ (NIN,*) (C(I,J),J=1,ICONST)
   40       CONTINUE
*
            IFAIL = 1
            CALL G02GKF(IP,ICONST,V,LDV,C,LDC,B,1.0D0,SE,COV,WK,IFAIL)
*
            IF (IFAIL.EQ.0) THEN
               WRITE (NOUT,*) '      Estimate     Standard error'
               WRITE (NOUT,*)
               DO 60 I = 1, IP
                  WRITE (NOUT,99997) B(I), SE(I)
   60          CONTINUE
            ELSE
               WRITE (NOUT,*)
               WRITE (NOUT,99996) ' ** G02GKF returned with IFAIL = ',
     +           IFAIL
            END IF
         ELSE
            WRITE (NOUT,*)
            WRITE (NOUT,99996) ' ** G02GCF returned with IFAIL = ',
     +        IFAIL
         END IF
      END IF
```

```
*
99999 FORMAT (1X,A,E12.4)
99998 FORMAT (1X,A,I2)
99997 FORMAT (1X,2F14.4)
99996 FORMAT (1X,A,I5)
      END
```

9.2 Program Data

```
G02GKF Example Program Data
 15 8 0
1.0 0.0 0.0 1.0 0.0 0.0 0.0 0.0 141.0
1.0 0.0 0.0 0.0 1.0 0.0 0.0 0.0  67.0
1.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 114.0
1.0 0.0 0.0 0.0 0.0 0.0 1.0 0.0  79.0
1.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0  39.0
0.0 1.0 0.0 1.0 0.0 0.0 0.0 0.0 131.0
0.0 1.0 0.0 0.0 1.0 0.0 0.0 0.0  66.0
0.0 1.0 0.0 0.0 0.0 1.0 0.0 0.0 143.0
0.0 1.0 0.0 0.0 0.0 0.0 1.0 0.0  72.0
0.0 1.0 0.0 0.0 0.0 0.0 0.0 1.0  35.0
0.0 0.0 1.0 1.0 0.0 0.0 0.0 0.0  36.0
0.0 0.0 1.0 0.0 1.0 0.0 0.0 0.0  14.0
0.0 0.0 1.0 0.0 0.0 1.0 0.0 0.0  38.0
0.0 0.0 1.0 0.0 0.0 0.0 1.0 0.0  28.0
0.0 0.0 1.0 0.0 0.0 0.0 0.0 1.0  16.0
 1   1   1   1   1   1   1   1   9
0.0 0.0
1.0 0.0
1.0 0.0
1.0 0.0
0.0 1.0
0.0 1.0
0.0 1.0
0.0 1.0
0.0 1.0
```

9.3 Program Results

```
 G02GKF Example Program Results

 Deviance =   0.9038E+01
 Degrees of freedom =  8

       Estimate     Standard error

         3.9831       0.0396
         0.3961       0.0458
         0.4118       0.0457
        -0.8079       0.0622
         0.5112       0.0562
        -0.2285       0.0727
         0.4680       0.0569
        -0.0316       0.0675
        -0.7191       0.0887
```