

NAG Library Routine Document

G02LDF

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

G02LDF calculates predictions given the output from an orthogonal scores PLS regression (G02LAF or G02LBF) and G02LCF.

2 Specification

```

SUBROUTINE G02LDF(IP, MY, ORIG, XBAR, YBAR, ISCALE, XSTD, YSTD, B, LDB,
1          N, MZ, ISZ, Z, LDZ, YHAT, LDYHAT, IFAIL)

  INTEGER          IP, MY, ORIG, ISCALE, LDB, N, MZ, ISZ(MZ), LDZ,
1          LDYHAT, IFAIL
  double precision XBAR(IP), YBAR(MY), XSTD(IP), YSTD(MY), B(LDB,MY),
1          Z(LDZ,MZ), YHAT(LDYHAT,MY)

```

3 Description

G02LDF calculates the predictions $\hat{Y}$ of a PLS model given a set Z of test data and a set B of parameter estimates as returned by G02LCF.

If G02LCF returns parameter estimates for the original data scale, no further information is required.

If G02LCF returns parameter estimates for the centred, and possibly scaled, data, further information is required. The means of variables in the fitted model must be supplied. In the case of a PLS model fitted by using scaled data, the standard deviations of variables in the fitted model must also be supplied. These means and standard deviations are those returned by either G02LAF and G02LBF.

4 References

None.

5 Parameters

- 1: IP – INTEGER *Input*
On entry: the number of predictor variables in the fitted model. IP must take the same value as that supplied to G02LAF or G02LBF to fit the model.
Constraint: IP > 1.
- 2: MY – INTEGER *Input*
On entry: the number of response variables in the fitted model. MY must take the same value as that supplied to G02LAF or G02LBF to fit the model.
Constraint: MY ≥ 1.
- 3: ORIG – INTEGER *Input*
On entry: indicates how parameter estimates are supplied.
ORIG = 1
Parameter estimates are for the original data.

ORIG = -1

Parameter estimates are for the centred, and possibly scaled, data.

Constraint: ORIG = -1 or 1.

- 4: XBAR(IP) – *double precision* array *Input*
On entry: if ORIG = -1, XBAR must contain mean values of predictor variables in the model; otherwise XBAR is not referenced.

- 5: YBAR(MY) – *double precision* array *Input*
On entry: if ORIG = -1, YBAR must contain the mean value of each response variable in the model; otherwise YBAR is not referenced.

- 6: ISCALE – INTEGER *Input*
On entry: if ORIG = -1, ISCALE must take the value supplied to either G02LAF or G02LBF; otherwise ISCALE is not referenced.
Constraint: if ORIG = -1, ISCALE = -1, 1 or 2.

- 7: XSTD(IP) – *double precision* array *Input*
On entry: if ORIG = -1 and ISCALE $\neq$ -1, XSTD must contain the scalings of predictor variables in the model as returned from either G02LAF or G02LBF; otherwise XSTD is not referenced.

- 8: YSTD(MY) – *double precision* array *Input*
On entry: if ORIG = -1 and ISCALE $\neq$ -1, YSTD must contain the scalings of response variables as returned from either G02LAF or G02LBF; otherwise YSTD is not referenced.

- 9: B(LDB,MY) – *double precision* array *Input*
On entry: if ORIG = -1, B must contain the parameter estimate for the centred, and possibly scaled, data as returned by G02LCF; otherwise B must contain the parameter estimates for the original data as returned by G02LCF.

- 10: LDB – INTEGER *Input*
On entry: the first dimension of the array B as declared in the (sub)program from which G02LDF is called. If ORIG = -1, LDB must be at least IP; otherwise B also contains the estimate for the intercept parameter and consequently LDB must be at least 1 + IP.
Constraints:
 if ORIG = -1, LDB $\geq$ IP;
 if ORIG = 1, LDB $\geq$ 1 + IP.

- 11: N – INTEGER *Input*
On entry: n , the number of observations in the test data Z .
Constraint: $N \geq 1$.

- 12: MZ – INTEGER *Input*
On entry: the number of available predictor variables in the test data.
Constraint: $MZ \geq IP$.

- 13: ISZ(MZ) – INTEGER array *Input*
On entry: indicates which predictor variables are to be included in the model. Predictor variables included from Z must be in the same order as those included in the fitted model.

If $ISZ(j) = 1$, the j th predictor variable is included in the model, for $j = 1, 2, \dots, MZ$, otherwise $ISZ(j) = 0$.

Constraints:

$$ISZ(j) = 0 \text{ or } 1, \text{ for } j = 1, 2, \dots, MZ;$$

$$\sum_j ISZ(j) = IP.$$

14: $Z(LDZ, MZ)$ – *double precision* array *Input*

On entry: $Z(i, j)$ contains the i th observation on the j th available predictor variable, for $i = 1, 2, \dots, N$ and $j = 1, 2, \dots, MZ$.

15: LDZ – INTEGER *Input*

On entry: the first dimension of the array Z as declared in the (sub)program from which G02LDF is called.

Constraint: $LDZ \geq N$.

16: $YHAT(LDYHAT, MY)$ – *double precision* array *Output*

On exit: $YHAT(i, j)$ contains the i th predicted value of the j th y -variable in the model.

17: $LDYHAT$ – INTEGER *Input*

On entry: the first dimension of the array $YHAT$ as declared in the (sub)program from which G02LDF is called.

Constraint: $LDYHAT \geq N$.

18: $IFAIL$ – 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, $IP \leq 1$,
or $MY < 1$,
or $ORIG \neq -1$ or 1 ,
or $ORIG = -1$ and $ISCALE \neq -1, 1$ or 2 ,
or $N < 1$,
or an element of $ISZ \neq 0$ or 1 .

$IFAIL = 2$

On entry, $ORIG = -1$ and $LDB < IP$,
or $ORIG = 1$ and $LDB < 1 + IP$,
or $MZ < IP$,

or LDZ < N,
or LDYHAT < N.

IFAIL = 3

The sum of elements in ISZ does not equal IP.

7 Accuracy

Not applicable.

8 Further Comments

G02LDF allocates internally $3 \times IP + MY$ elements of *double precision* storage.

9 Example

This example reads in parameter estimates for a fitted PLS model and prediction data, and the PLS model predictions are calculated.

9.1 Program Text

```
*      G02LDF Example Program Text
*      Mark 22 Release. NAG Copyright 2008.
*      .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          NMAX, MZMAX, MYMAX
      PARAMETER        (NMAX=25,MZMAX=28,MYMAX=10)
*      .. Local Scalars ..
      INTEGER          I, IFAIL, IP, ISCALE, J, L, LDB, LDYHAT, LDZ, MY,
+                     MZ, N, ORIG
*      .. Local Arrays ..
      DOUBLE PRECISION B(MZMAX+1,MYMAX), XBAR(MZMAX), XSTD(MZMAX),
+                     YBAR(MYMAX), YHAT(NMAX,MYMAX), YSTD(MYMAX),
+                     Z(NMAX,MZMAX)
      INTEGER          ISZ(MZMAX)
*      .. External Subroutines ..
      EXTERNAL         G02LDF, X04CAF
*      .. Executable Statements ..
      WRITE (NOUT,*) 'G02LDF example program results'
*      Skip header in data file.
      READ (NIN,*)
*      Read data values.
      READ (NIN,*) IP, MY, ORIG, ISCALE, N, MZ
      LDB = MZMAX + 1
      LDYHAT = NMAX
      LDZ = NMAX
*      Read prediction x-data
      DO I = 1, N
         READ (NIN,*) (Z(I,J),J=1,MZ)
      END DO
*      Read elements of ISZ
      READ (NIN,*) (ISZ(J),J=1,MZ)
*      Read parameter estimates
      L = IP
      IF (ORIG.NE.-1) L = L + 1
      READ (NIN,*) ((B(J,I),I=1,MY),J=1,L)
*      Read means
      READ (NIN,*) (XBAR(J),J=1,IP)
      READ (NIN,*) (YBAR(L),L=1,MY)
      IF ((ORIG.EQ.-1) .AND. (ISCALE.NE.-1)) THEN
         READ (NIN,*) (XSTD(J),J=1,IP)
         READ (NIN,*) (YSTD(L),L=1,MY)
      END IF
*      Calculate predictions
```

```

      IFAIL = 1
      CALL G02LDF(IP,MY,ORIG,XBAR,YBAR,ISCALE,XSTD,YSTD,B,LDB,N,MZ,ISZ,
+             Z,LDZ,YHAT,LDYHAT,IFAIL)
      IF (IFAIL.NE.0) THEN
          WRITE (NOUT,99999) IFAIL
      ELSE
          CALL X04CAF('G','X',N,MY,YHAT,LDYHAT,'YHAT',IFAIL)
      END IF
*
99999 FORMAT (1X,/1X,' ** G02LDF returned with IFAIL = ',I5)
      END

```

9.2 Program Data

G02LDF Example Program Data

[illegible]

```
0.4520                                     : YBAR
1.4956  1.3233  0.5829  0.7735  0.6247  0.7966
2.4113  2.0421  0.4678  0.8197  0.9420  0.1735
1.0475  0.1359  1.3853                                     : XSTD
0.9062                                     : YSTD
```

9.3 Program Results

G02LDF example program results

YHAT

```
1
1      0.2132
2      0.5152
3      0.1437
4      0.4459
5      0.1716
6      2.4809
7      0.0964
8      1.4475
9     -0.1546
10     -0.5492
11     0.5393
12     0.2686
13    -1.1332
14     1.7975
15     0.4973
```
