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, N,     &
                   MZ, ISZ, Z, LDZ, YHAT, LDYHAT, IFAIL)

INTEGER            IP, MY, ORIG, ISCALE, LDB, N, MZ, ISZ(MZ), LDZ, LDYHAT,      &
                   IFAIL
REAL (KIND=nag_wp) XBAR(IP), YBAR(MY), XSTD(IP), YSTD(MY), B(LDB,MY),           &
                   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 means and 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) – REAL (KIND=nag_wp) 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) – REAL (KIND=nag_wp) 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) – REAL (KIND=nag_wp) 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) – REAL (KIND=nag_wp) 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) – REAL (KIND=nag_wp) 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 $1 + \text{IP}$.

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 ≥ 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 jth predictor variable is included in the model, for $j = 1, 2, \ldots, \text{MZ}$, otherwise ISZ($j$) = 0.

Constraints:

ISZ(j) = 0 or 1, for $j = 1, 2, \ldots, \text{MZ}$;
$\sum_j \text{ISZ}(j) = \text{IP}$.

14: Z(LDZ,MZ) – REAL (KIND=nag_wp) array *Input*

On entry: Z(i, j) contains the ith observation on the jth available predictor variable, for $i = 1, 2, \ldots, \text{N}$ and $j = 1, 2, \ldots, \text{MZ}$.

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) – REAL (KIND=nag_wp) array *Output*

On exit: YHAT(i, j) contains the ith predicted value of the jth 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.

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.**

On exit: IFAIL = 0 unless the routine detects an error or a warning has been flagged (see Section 6).

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 $3 \times \text{IP} + \text{MY}$ elements of real 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

```
    Program g02ldfe

!     G02LDF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: g02ldf, nag_wp, x04caf
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: nin = 5, nout = 6
!     .. Local Scalars ..
      Integer                          :: i, ifail, ip, iscale, ldb, ldyhat,   &
                                          ldz, my, mz, n, orig
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: b(:,:), xbar(:), xstd(:), ybar(:),   &
                                          yhat(:,:), ystd(:), z(:,:)
      Integer, Allocatable             :: isz(:)
!     .. Intrinsic Procedures ..
      Intrinsic                        :: sum
!     .. Executable Statements ..
      Write (nout,*) 'G02LDF Example Program Results'
      Write (nout,*)
      Flush (nout)

!     Skip heading in data file
      Read (nin,*)

!     Read in problem size
      Read (nin,*) my, orig, iscale, n, mz

      ldyhat = n
      ldz = n
      Allocate (ybar(my),ystd(my),isz(mz),z(ldz,mz),yhat(ldyhat,my))

!     Read prediction x-data
      Read (nin,*)(z(i,1:mz),i=1,n)

!     Read in elements of ISZ
      Read (nin,*) isz(1:mz)

!     Calculate IP
```

```
      ip = sum(isz(1:mz))

      ldb = ip
      If (orig==1) Then
        ldb = ldb + 1
      End If
      Allocate (xbar(ip),xstd(ip),b(ldb,my))

!     Read parameter estimates
      Read (nin,*)(b(i,1:my),i=1,ldb)

!     Read means
      If (orig==-1) Then
        Read (nin,*) xbar(1:ip)
        Read (nin,*) ybar(1:my)

        If (iscale/=-1) Then
!         Read scalings
          Read (nin,*) xstd(1:ip)
          Read (nin,*) ystd(1:my)
        End If
      End If

!     Calculate predictions
      ifail = 0
      Call g02ldf(ip,my,orig,xbar,ybar,iscale,xstd,ystd,b,ldb,n,mz,isz,z,ldz, &
        yhat,ldyhat,ifail)

!     Display results
      ifail = 0
      Call x04caf('General',' ',n,my,yhat,ldyhat,'YHAT',ifail)

    End Program g02ldfe
```

9.2 Program Data

```
G02LDF Example Program Data
 1   -1   1   15   15                              : MY, ORIG, SCALE, N, MZ
 -2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
  1.9607 -1.6324  0.5746  1.9607 -1.6324  0.5740
  2.8369  1.4092 -3.1398
 -2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
  1.9607 -1.6324  0.5746  0.0744 -1.7333  0.0902
  2.8369  1.4092 -3.1398
 -2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
  0.0744 -1.7333  0.0902  1.9607 -1.6324  0.5746
  2.8369  1.4092 -3.1398
 -2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
  0.0744 -1.7333  0.0902  0.0744 -1.7333  0.0902
  2.8369  1.4092 -3.1398
 -2.6931 -2.5271 -1.2871  2.8369  1.4092 -3.1398
  0.0744 -1.7333  0.0902  0.0744 -1.7333  0.0902
  2.8369  1.4092 -3.1398
 -2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
 -4.7548  3.6521  0.8524  0.0744 -1.7333  0.0902
  2.8369  1.4092 -3.1398
 -2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
  0.0744 -1.7333  0.0902  0.0744 -1.7333  0.0902
 -1.2201  0.8829  2.2253
 -2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
  2.4064  1.7438  1.1057  0.0744 -1.7333  0.0902
  2.8369  1.4092 -3.1398
 -2.6931 -2.5271 -1.2871  0.0744 -1.7333  0.0902
  0.0744 -1.7333  0.0902  0.0744 -1.7333  0.0902
  2.8369  1.4092 -3.1398
  2.2261 -5.3648  0.3049  3.0777  0.3891 -0.0701
  0.0744 -1.7333  0.0902  0.0744 -1.7333  0.0902
  2.8369  1.4092 -3.1398
 -4.1921 -1.0285 -0.9801  3.0777  0.3891 -0.0701
  0.0744 -1.7333  0.0902  0.0744 -1.7333  0.0902
```

```
 2.8369  1.4092 -3.1398
-4.9217  1.2977  0.4473  3.0777  0.3891 -0.0701
 0.0744 -1.7333  0.0902  0.0744 -1.7333  0.0902
 2.8369  1.4092 -3.1398
-2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
 2.2261 -5.3648  0.3049  2.2261 -5.3648  0.3049
 2.8369  1.4092 -3.1398
-2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
-4.9217  1.2977  0.4473  0.0744 -1.7333  0.0902
 2.8369  1.4092 -3.1398
-2.6931 -2.5271 -1.2871  3.0777  0.3891 -0.0701
-4.1921 -1.0285 -0.9801  0.0744 -1.7333  0.0902
 2.8369  1.4092 -3.1398                          : End of Z
 1  1  1  1  1  1  1  1  1  1  1  1  1  1  1     : Elements of ISZ
-0.1383  0.0572 -0.1906  0.1238  0.0591  0.0936
-0.2842  0.4713  0.2661 -0.0914  0.1226 -0.0488
 0.0332  0.0332 -0.0332                          : End of B
-2.6137 -2.3614 -1.0449  2.8614  0.3156 -0.2641
-0.3146 -1.1221  0.2401  0.4694 -1.9619  0.1691
 2.5664  1.3741 -2.7821                          : End of XBAR
 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                          : End of 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
```