NAG Library Routine Document

G02AEF

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

G02AEF computes the factor loading matrix associated with the nearest correlation matrix with k-factor structure, in the Frobenius norm, to a given square, input matrix.

2 Specification

```
SUBROUTINE G02AEF (G, LDG, N, K, ERRTOL, MAXIT, X, LDX, ITER, FEVAL,          &
                   NRMPGD, IFAIL)

INTEGER            LDG, N, K, MAXIT, LDX, ITER, FEVAL, IFAIL
REAL (KIND=nag_wp) G(LDG,N), ERRTOL, X(LDX,K), NRMPGD
```

3 Description

A correlation matrix C with k-factor structure may be characterised as a real square matrix that is symmetric, has a unit diagonal, is positive semidefinite and can be written as $C = XX^{\mathrm{T}} + \mathrm{diag}\left(I - XX^{\mathrm{T}}\right)$, where I is the identity matrix and X has n rows and k columns. X is often referred to as the factor loading matrix.

G02AEF applies a spectral projected gradient method to the modified problem $\min \left\|G - XX^{\mathrm{T}} + \mathrm{diag}\left(XX^{\mathrm{T}} - I\right)\right\|_F$ such that $\left\|x_i^{\mathrm{T}}\right\|_2 \leq 1$, for $i = 1, 2, \ldots, n$, where x_i is the ith row of the factor loading matrix, X, which gives us the solution.

4 References

Birgin E G, Martínez J M and Raydan M (2001) Algorithm 813: SPG–software for convex-constrained optimization *ACM Trans. Math. Software* **27** 340–349

Borsdorf R, Higham N J and Raydan M (2010) Computing a nearest correlation matrix with factor structure. *SIAM J. Matrix Anal. Appl.* **31(5)** 2603–2622

5 Parameters

1: G(LDG,N) – REAL (KIND=nag_wp) array *Input/Output*

On entry: G, the initial matrix.

On exit: a symmetric matrix $\frac{1}{2}\left(G + G^{\mathrm{T}}\right)$ with the diagonal elements set to unity.

2: LDG – INTEGER *Input*

On entry: the first dimension of the array G as declared in the (sub)program from which G02AEF is called.

Constraint: LDG $\geq$ N.

3: N – INTEGER *Input*

On entry: n, the order of the matrix G.

Constraint: N > 0.

4: K – INTEGER *Input*

On entry: k, the number of factors and columns of X.

Constraint: $0 < \mathrm{K} \leq \mathrm{N}$.

5: ERRTOL – REAL (KIND=nag_wp) *Input*

On entry: the termination tolerance for the projected gradient norm. See references for further details. If ERRTOL ≤ 0.0 then 0.01 is used. This is often a suitable default value.

6: MAXIT – INTEGER *Input*

On entry: specifies the maximum number of iterations in the spectral projected gradient method.

If MAXIT ≤ 0, 40000 is used.

7: X(LDX,K) – REAL (KIND=nag_wp) array *Output*

On exit: contains the matrix X.

8: LDX – INTEGER *Input*

On entry: the first dimension of the array X as declared in the (sub)program from which G02AEF is called.

Constraint: LDX $\geq$ N.

9: ITER – INTEGER *Output*

On exit: the number of steps taken in the spectral projected gradient method.

10: FEVAL – INTEGER *Output*

On exit: the number of evaluations $\min \left\|G - XX^{\mathrm{T}} + \mathrm{diag}\left(XX^{\mathrm{T}} - I\right)\right\|_F$.

11: NRMPGD – REAL (KIND=nag_wp) *Output*

On exit: the norm of the projected gradient at the final iteration.

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.

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, N ≤ 0,
or LDG $<$ N,
or LDX $<$ N.

or $K \leq 0$,
or $K > N$,

IFAIL $= 2$

The routine fails to converge in MAXIT iterations. Increase MAXIT or check the call to the routine.

IFAIL $= -999$

Internal memory allocation failed.

7 Accuracy

The returned accuracy is controlled by ERRTOL and limited by ***machine precision***.

8 Further Comments

Arrays are internally allocated by G02AEF. The total size of these arrays is $N \times N + 4 \times N \times K + (nb + 3) \times N + N + 50$ real elements and $6 \times N$ integer elements. Here *nb* is the block size required for optimal performance by F08FEF (DSYTRD) and F08FGF (DORMTR) which are called internally. All allocated memory is freed before return of G02AEF.

See G03CAF for constructing the factor loading matrix from a known correlation matrix.

9 Example

This example finds the nearest correlation matrix with $k = 2$ factor structure to:

$$G = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

9.1 Program Text

```
    Program g02aefe

!     G02AEF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: dgemm, g02aef, nag_wp, x04caf
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Real (Kind=nag_wp), Parameter    :: one = 1.0_nag_wp
      Real (Kind=nag_wp), Parameter    :: zero = 0.0_nag_wp
      Integer, Parameter               :: nin = 5, nout = 6
!     .. Local Scalars ..
      Real (Kind=nag_wp)               :: errtol, nrmpgd
      Integer                          :: feval, i, ifail, iter, k, lda, ldg,  &
                                          ldx, maxit, n
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: a(:,:), g(:,:), x(:,:)
!     .. Executable Statements ..
      Write (nout,*) 'G02AEF Example Program Results'
      Write (nout,*)
      Flush (nout)

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

!     Read in the problem size
```

```
      Read (nin,*) n

      lda = n
      ldg = n
      ldx = n
      Allocate (a(lda,n),g(ldg,n),x(ldx,n))

!     Read in the matrix G
      Read (nin,*)(g(i,1:n),i=1,n)

!     Use the defaults for ERRTOL and MAXIT
      errtol = zero
      maxit = 0

!     Set k value
      k = 2

!     Calculate the nearest factor loading matrix
      ifail = 0

      Call g02aef(g,ldg,n,k,errtol,maxit,x,ldx,iter,feval,nrmpgd,ifail)

!     Display results
      ifail = 0
      Call x04caf('General',' ',n,k,x,ldx,'Factor Loading Matrix X',ifail)
      Write (nout,*)
      Write (nout,99999) 'Number of steps taken:', iter
      Write (nout,99998) 'Number of function evaluations:', feval

!     Generate Nearest k factor correlation matrix
!     The NAG name equivalent of dgemm is f06yaf
      Call dgemm('N','T',n,n,k,one,x,n,x,n,zero,a,n)
      Do i = 1, n
        a(i,i) = one
      End Do
      Write (nout,*)
      Flush (nout)
      ifail = 0
      Call x04caf('General',' ',n,n,a,lda,'Nearest Correlation Matrix',ifail)

99999 Format (1X,A,I11)
99998 Format (1X,A,I9)

    End Program g02aefe
```

9.2 Program Data

```
G02AEF Example Program Data
 4                          :: N
 2.0   -1.0    0.0    0.0
-1.0    2.0   -1.0    0.0
 0.0   -1.0    2.0   -1.0
 0.0    0.0   -1.0    2.0  :: End of G
```

9.3 Program Results

```
 G02AEF Example Program Results

 Factor Loading Matrix X
          1       2
 1   0.7665 -0.6271
 2  -0.4250  0.9052
 3  -0.4250 -0.9052
 4   0.7665  0.6271

 Number of steps taken:          5
 Number of function evaluations:        6

 Nearest Correlation Matrix
```

```
          1       2       3       4
 1   1.0000 -0.8935  0.2419  0.1943
 2  -0.8935  1.0000 -0.6388  0.2419
 3   0.2419 -0.6388  1.0000 -0.8935
 4   0.1943  0.2419 -0.8935  1.0000
```