NAG Library Routine Document

F04JAF

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

F04JAF finds the minimal solution of a linear least-squares problem, $Ax = b$, where A is a real m by n $(m \geq n)$ matrix and b is an m element vector.

2 Specification

```
 SUBROUTINE F04JAF(M, N, A, LDA, B, TOL, SIGMA, IRANK, WORK, LWORK,
1                  IFAIL)
 INTEGER           M, N, LDA, IRANK, LWORK, IFAIL
 double precision  A(LDA,N), B(M), TOL, SIGMA, WORK(LWORK)
```

3 Description

The minimal least-squares solution of the problem $Ax = b$ is the vector x of minimum (Euclidean) length which minimizes the length of the residual vector $r = b - Ax$.

The real m by $n(m \geq n)$ matrix A may be factorized by the singular value decomposition (SVD) as

$$A = Q\begin{pmatrix} D \\ 0 \end{pmatrix}P^{\mathrm{T}},$$

where Q is an m by m orthogonal matrix, P is an n by n orthogonal matrix and D is the n by n diagonal matrix

$$D = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$, these being the singular values of A. The columns of P are the right-hand singular vectors of A.

If the singular values $\sigma_{k+1}, \ldots, \sigma_n$ are negligible, but σ_k is not negligible, relative to the data errors in A, then the rank of A is taken to be k and the minimal least-squares solution is given by

$$x = P\begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix}Q^{\mathrm{T}}b,$$

where $S = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k)$. The routine also returns the value of the standard error

$$\begin{aligned} \sigma &= \sqrt{\tfrac{r^{\mathrm{T}}r}{m-k}}, \quad \text{if } m > k \\ &= 0, \qquad\quad \text{if } m = k, \end{aligned}$$

$r^{\mathrm{T}}r$ being the residual sum of squares.

4 References

Lawson C L and Hanson R J (1974) *Solving Least-squares Problems* Prentice–Hall

5 Parameters

1: M – INTEGER *Input*

On entry: m, the number of rows of the matrix A.

Constraint: $\text{M} \geq \text{N}$.

2: N – INTEGER *Input*

On entry: n, the number of columns of the matrix A.

Constraint: $1 \leq \text{N} \leq \text{M}$.

3: A(LDA,N) – ***double precision*** array *Input/Output*

On entry: the m by n matrix A.

On exit: the first n rows of A are overwritten by the n by n matrix P^{T}, i.e., the right-hand singular vectors, stored by **rows**. The rest of the array is used as workspace.

4: LDA – INTEGER *Input*

On entry: the first dimension of the array A as declared in the (sub)program from which F04JAF is called.

Constraint: $\text{LDA} \geq \text{M}$.

5: B(M) – ***double precision*** array *Input/Output*

On entry: the right-hand side vector b.

On exit: the first n elements of B contain the minimal least-squares solution vector x. The remaining $m - n$ elements are used for workspace.

6: TOL – ***double precision*** *Input*

On entry: a relative tolerance to be used to determine the rank of A. TOL should be chosen as approximately the largest relative error in the elements of A. For example, if the elements of A are correct to about 4 significant figures, then TOL should be set to about 5×10^{-4}. See Section 8 for a description of how TOL is used to determine rank. If TOL is outside the range $(\epsilon, 1.0)$, where ϵ is the ***machine precision***, then the value ϵ is used instead of TOL. For most problems this is unreasonably small.

7: SIGMA – ***double precision*** *Output*

On exit: the standard error, i.e., the value $\sqrt{r^{\mathrm{T}}r/(m-k)}$ when $m > k$, and the value zero when $m = k$. Here r is the residual vector $b - Ax$ and k is the rank of A.

8: IRANK – INTEGER *Output*

On exit: k, the rank of the matrix A.

9: WORK(LWORK) – ***double precision*** array *Output*

On exit: the first n elements of WORK contain the singular values of A arranged in descending order. WORK$(n+1)$ contains the total number of iterations taken by the QR algorithm. The rest of WORK is used as workspace.

10: LWORK – INTEGER *Input*

On entry: the dimension of the array WORK as declared in the (sub)program from which F04JAF is called.

Constraint: LWORK $\geq 4 \times$ N.

11: 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, N < 1,
or M $<$ N,
or LDA $<$ M,
or LWORK $< 4 \times$ N.

IFAIL $= 2$

The QR algorithm has failed to converge to the singular values in $50 \times$ N iterations. This failure is not likely to occur.

7 Accuracy

The computed factors Q, D and P^{T} satisfy the relation

$$Q\begin{pmatrix} D \\ 0 \end{pmatrix}P^{\mathrm{T}} = A + E,$$

where $\|E\|_2 \leq c\epsilon\|A\|_2$, ϵ being the ***machine precision*** and c being a modest function of m and n. Note that $\|A\|_2 = \sigma_1$.

For a fuller discussion covering the accuracy of the solution x, see Lawson and Hanson (1974), especially pages 50 and 95.

8 Further Comments

The time taken by F04JAF is approximately proportional to $n^2(m + 6n)$.

This routine is column-biased and so is suitable for use in paged environments.

F04JDF gives the minimal least-squares solution for the case $m < n$.

The rank of A, say k, is returned as the largest integer such that

$$\sigma_k > \mathrm{TOL} \times \sigma_1,$$

unless $\sigma_1 = 0$ in which case k is returned as zero. That is, singular values which satisfy $\sigma_i \leq \text{TOL} \times \sigma_1$ are regarded as negligible because relative perturbations of order TOL can make such singular values zero.

9 Example

This example obtains the minimal least-squares solution for $r = b - Ax$ where

$$A = \begin{pmatrix} 0.05 & 0.05 & 0.25 & -0.25 \\ 0.25 & 0.25 & 0.05 & -0.05 \\ 0.35 & 0.35 & 1.75 & -1.75 \\ 1.75 & 1.75 & 0.35 & -0.35 \\ 0.30 & -0.30 & 0.30 & 0.30 \\ 0.40 & -0.40 & 0.40 & 0.40 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}$$

and the value TOL is to be taken as 5×10^{-4}.

9.1 Program Text

```
*     F04JAF Example Program Text
*     Mark 14 Revised. NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          MMAX, NMAX, LDA, LWORK
      PARAMETER        (MMAX=8,NMAX=MMAX,LDA=MMAX,LWORK=4*NMAX)
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
*     .. Local Scalars ..
      DOUBLE PRECISION SIGMA, TOL
      INTEGER          I, IFAIL, IRANK, J, M, N
*     .. Local Arrays ..
      DOUBLE PRECISION A(LDA,NMAX), B(MMAX), WORK(LWORK)
*     .. External Subroutines ..
      EXTERNAL         F04JAF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'F04JAF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) M, N
      TOL = 5.0D-4
      WRITE (NOUT,*)
      IF (M.GT.0 .AND. M.LE.MMAX .AND. N.GT.0 .AND. N.LE.NMAX) THEN
         READ (NIN,*) ((A(I,J),J=1,N),I=1,M)
         READ (NIN,*) (B(I),I=1,M)
         IFAIL = 1
*
         CALL F04JAF(M,N,A,LDA,B,TOL,SIGMA,IRANK,WORK,LWORK,IFAIL)
*
         IF (IFAIL.EQ.0) THEN
            WRITE (NOUT,*) 'Solution vector'
            WRITE (NOUT,99997) (B(I),I=1,N)
            WRITE (NOUT,*)
            WRITE (NOUT,99998) 'Standard error = ', SIGMA,
     +        '    Rank = ', IRANK
         ELSE
            WRITE (NOUT,99996) IFAIL
         END IF
      ELSE
         WRITE (NOUT,99999) 'M or N out of range: M = ', M, '  N = ', N
      END IF
*
99999 FORMAT (1X,A,I5,A,I5)
99998 FORMAT (1X,A,F6.3,A,I2)
99997 FORMAT (1X,8F9.3)
99996 FORMAT (1X,' ** F04JAF returned with IFAIL = ',I5)
      END
```

9.2 Program Data

```
F04JAF Example Program Data
  6  4
  0.05  0.05  0.25 -0.25
  0.25  0.25  0.05 -0.05
  0.35  0.35  1.75 -1.75
  1.75  1.75  0.35 -0.35
  0.30 -0.30  0.30  0.30
  0.40 -0.40  0.40  0.40
  1.0   2.0   3.0   4.0   5.0   6.0
```

9.3 Program Results

```
 F04JAF Example Program Results

 Solution vector
     4.967   -2.833    4.567    3.233

 Standard error =  0.909    Rank =  3
```