NAG Toolbox for MATLAB

f07pb

1 Purpose

f07pb uses the diagonal pivoting factorization

$$A = UDU^{\mathrm{T}} \qquad \text{or} \qquad A = LDL^{\mathrm{T}}$$

to compute the solution to a real system of linear equations

$$AX = B,$$

where A is an n by n symmetric matrix stored in packed format and X and B are n by r matrices. Error bounds on the solution and a condition estimate are also provided.

2 Syntax

```
[afp, ipiv, x, rcond, ferr, berr, info] = f07pb(fact, uplo, ap, afp,
ipiv, b, 'n', n, 'nrhs_p', nrhs_p)
```

3 Description

f07pb performs the following steps:

1. If **fact** = 'N', the diagonal pivoting method is used to factor A as $A = UDU^{\mathrm{T}}$ if **uplo** = 'U' or $A = LDL^{\mathrm{T}}$ if **uplo** = 'L', where U (or L) is a product of permutation and unit upper (lower) triangular matrices and D is symmetric and block diagonal with 1 by 1 and 2 by 2 diagonal blocks.
2. If some $d_{ii} = 0$, so that D is exactly singular, then the function returns with **info** $= i$. Otherwise, the factored form of A is used to estimate the condition number of the matrix A. If the reciprocal of the condition number is less than ***machine precision***, **info** $\geq N + 1$ is returned as a warning, but the function still goes on to solve for X and compute error bounds as described below.
3. The system of equations is solved for X using the factored form of A.
4. Iterative refinement is applied to improve the computed solution matrix and to calculate error bounds and backward error estimates for it.

4 References

Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J J, Du Croz J J, Greenbaum A, Hammarling S, McKenney A and Sorensen D 1999 *LAPACK Users' Guide* (3rd Edition) SIAM, Philadelphia URL: http://www.netlib.org/lapack/lug

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

Higham N J 2002 *Accuracy and Stability of Numerical Algorithms* (2nd Edition) SIAM, Philadelphia

5 Parameters

5.1 Compulsory Input Parameters

1: **fact – string**

Specifies whether or not the factorized form of the matrix A has been supplied.

fact = 'F'

afp and **ipiv** contain the factorized form of the matrix A. **ap**, **afp** and **ipiv** will not be modified.

fact = 'N'

The matrix A will be copied to **afp** and factorized.

Constraint: **fact** = 'F' or 'N'.

2: **uplo – string**

If **uplo** = 'U', the upper triangle of A is stored.

If **uplo** = 'L', the lower triangle of A is stored.

Constraint: **uplo** = 'U' or 'L'.

3: **ap**$(*)$ **– double array**

Note: the dimension of the array **ap** must be at least $\max(1, \mathbf{n} \times (\mathbf{n} + 1)/2)$.

The n by n symmetric matrix A, packed by columns.

More precisely,

if **uplo** = 'U', the upper triangle of A must be stored with element A_{ij} in $\mathbf{ap}(i + j(j-1)/2)$ for $i \leq j$;
if **uplo** = 'L', the lower triangle of A must be stored with element A_{ij} in $\mathbf{ap}(i + (2n - j)(j - 1)/2)$ for $i \geq j$.

4: **afp**$(*)$ **– double array**

Note: the dimension of the array **afp** must be at least $\max(1, \mathbf{n} \times (\mathbf{n} + 1)/2)$.

If **fact** = 'F', **afp** contains the block diagonal matrix D and the multipliers used to obtain the factor U or L from the factorization $A = UDU^{\mathrm{T}}$ or $A = LDL^{\mathrm{T}}$ as computed by f07pd, stored as a packed triangular matrix in the same storage format as A.

5: **ipiv**$(*)$ **– int32 array**

Note: the dimension of the array **ipiv** must be at least $\max(1, \mathbf{n})$.

If **fact** = 'F', **ipiv** contains details of the interchanges and the block structure of D, as determined by f07pd.

$\mathbf{ipiv}(k) > 0$

Rows and columns k and $\mathbf{ipiv}(k)$ were interchanged and $D(k,k)$ is a 1 by 1 diagonal block.

uplo = 'U' and $\mathbf{ipiv}(k) = \mathbf{ipiv}(k-1) < 0$

Rows and columns $k-1$ and $-\mathbf{ipiv}(k)$ were interchanged and $D(k-1:k, k-1:k)$ is a 2 by 2 diagonal block.

uplo = 'L' and $\mathbf{ipiv}(k) = \mathbf{ipiv}(k+1) < 0$

Rows and columns $k+1$ and $-\mathbf{ipiv}(k)$ were interchanged and $D(k:k+1, k:k+1)$ is a 2 by 2 diagonal block.

6: **b(ldb,**$*$**) – double array**

The first dimension of the array **b** must be at least $\max(1, \mathbf{n})$

The second dimension of the array must be at least $\max(1, \mathbf{nrhs\_p})$

The n by r right-hand side matrix B.

5.2 Optional Input Parameters

1: **n – int32 scalar**

Default: The dimension of the array **ipiv**.

n, the number of linear equations, i.e., the order of the matrix A.

Constraint: $\mathbf{n} \geq 0$.

2: **nrhs_p – int32 scalar**

Default: The second dimension of the array **b**.

r, the number of right-hand sides, i.e., the number of columns of the matrix B.

Constraint: $\mathbf{nrhs\_p} \geq 0$.

5.3 Input Parameters Omitted from the MATLAB Interface

ldb, ldx, work, iwork

5.4 Output Parameters

1: **afp**$(*)$ **– double array**

Note: the dimension of the array **afp** must be at least $\max(1, \mathbf{n} \times (\mathbf{n} + 1)/2)$.

If **fact** = 'N', **afp** contains the block diagonal matrix D and the multipliers used to obtain the factor U or L from the factorization $A = UDU^{\mathrm{T}}$ or $A = LDL^{\mathrm{T}}$ as computed by f07pd, stored as a packed triangular matrix in the same storage format as A.

2: **ipiv**$(*)$ **– int32 array**

Note: the dimension of the array **ipiv** must be at least $\max(1, \mathbf{n})$.

If **fact** = 'N', **ipiv** contains details of the interchanges and the block structure of D, as determined by f07pd.

3: **x**(**ldx**,$*$) **– double array**

The first dimension of the array **x** must be at least $\max(1, \mathbf{n})$

The second dimension of the array must be at least $\max(1, \mathbf{nrhs\_p})$

If **info** $= 0$ or **info** $\geq N + 1$, the n by r solution matrix X.

4: **rcond – double scalar**

The estimate of the reciprocal condition number of the matrix A. If **rcond** $= 0$, the matrix may be exactly singular. This condition is indicated by a return code of **info** > 0leqN. Otherwise, if **rcond** is less than the ***machine precision***, the matrix is singular to working precision. This condition is indicated by a return code of **info** $\geq N + 1$.

5: **ferr**$(*)$ **– double array**

Note: the dimension of the array **ferr** must be at least $\max(1, \mathbf{nrhs\_p})$.

If **info** $= 0$ or **info** $\geq N + 1$, an estimate of the forward error bound for each computed solution vector, such that $\|\hat{x}_j - x_j\|_\infty / \|x_j\|_\infty \leq \mathbf{ferr}(j)$ where $\hat{x}_j$ is the jth column of the computed solution returned in the array **x** and x_j is the corresponding column of the exact solution X. The estimate is as reliable as the estimate for **rcond**, and is almost always a slight overestimate of the true error.

6: **berr**$(*)$ **– double array**

Note: the dimension of the array **berr** must be at least $\max(1, \textbf{nrhs\_p})$.

If **info** $= 0$ or **info** $\geq N + 1$, an estimate of the component-wise relative backward error of each computed solution vector $\hat{x}_j$ (i.e., the smallest relative change in any element of A or B that makes $\hat{x}_j$ an exact solution).

7: **info – int32 scalar**

info $= 0$ unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

info $= -i$

If **info** $= -i$, parameter i had an illegal value on entry. The parameters are numbered as follows:

1: **fact**, 2: **uplo**, 3: **n**, 4: **nrhs_p**, 5: **ap**, 6: **afp**, 7: **ipiv**, 8: **b**, 9: **ldb**, 10: **x**, 11: **ldx**, 12: **rcond**, 13: **ferr**, 14: **berr**, 15: **work**, 16: **iwork**, 17: **info**.

It is possible that **info** refers to a parameter that is omitted from the MATLAB interface. This usually indicates that an error in one of the other input parameters has caused an incorrect value to be inferred.

info > 0 and **info** $\leq N$

If **info** $\leq$ **n**, $d(i,i)$ is exactly zero. The factorization has been completed, but the factor D is exactly singular, so the solution and error bounds could not be computed. **rcond** $= 0$ is returned.

info $= N + 1$

D is nonsingular, but **rcond** is less than ***machine precision***, meaning that the matrix is singular to working precision. Nevertheless, the solution and error bounds are computed because there are a number of situations where the computed solution can be more accurate than the value of **rcond** would suggest.

7 Accuracy

For each right-hand side vector b, the computed solution $\hat{x}$ is the exact solution of a perturbed system of equations $(A + E)\hat{x} = b$, where

$$\|E\|_1 = O(\epsilon)\|A\|_1,$$

where ϵ is the ***machine precision***. See Chapter 11 of Higham 2002 for further details.

If $\hat{x}$ is the true solution, then the computed solution x satisfies a forward error bound of the form

$$\frac{\|x - \hat{x}\|_\infty}{\|\hat{x}\|_\infty} \leq w_c \operatorname{cond}(A, \hat{x}, b)$$

where $\operatorname{cond}(A, \hat{x}, b) = \left\| |A^{-1}|(|A||\hat{x}| + |b|) \right\|_\infty / \|\hat{x}\|_\infty \leq \operatorname{cond}(A) = \left\| |A^{-1}||A| \right\|_\infty \leq \kappa_\infty(A)$. If $\hat{x}$ is the jth column of X, then w_c is returned in **berr**(j) and a bound on $\|x - \hat{x}\|_\infty / \|\hat{x}\|_\infty$ is returned in **ferr**(j). See Section 4.4 of Anderson *et al.* 1999 for further details.

8 Further Comments

The factorization of A requires approximately $\frac{1}{3}n^3$ floating-point operations.

For each right-hand side, computation of the backward error involves a minimum of $4n^2$ floating-point operations. Each step of iterative refinement involves an additional $6n^2$ operations. At most five steps of iterative refinement are performed, but usually only one or two steps are required. Estimating the forward error involves solving a number of systems of equations of the form $Ax = b$; the number is usually 4 or 5 and never more than 11. Each solution involves approximately $2n^2$ operations.

The complex analogues of this function are f07pp for Hermitian matrices, and f07qp for symmetric matrices.

9 Example

```
fact = 'Not factored';
uplo = 'U';
ap = [-1.81;
     2.06;
     1.15;
     0.63;
     1.87;
     -0.21;
     -1.15;
     4.2;
     3.87;
     2.07];
afp = zeros(10, 1);
ipiv = zeros(4, 1, 'int32');
b = [0.96, 3.93;
     6.07, 19.25;
     8.38, 9.9;
     9.5, 27.85];
[afpOut, ipivOut, x, rcond, ferr, berr, info] = f07pb(fact, uplo, ap,
afp, ipiv, b)
```

```
afpOut =
    0.4074
    0.3031
   -2.5907
   -0.5960
    0.8115
    1.1500
    0.6537
    0.2230
    4.2000
    2.0700
ipivOut =
           1
           2
          -2
          -2
x =
   -5.0000    2.0000
   -2.0000    3.0000
    1.0000    4.0000
    4.0000    1.0000
rcond =
    0.0132
ferr =
   1.0e-13 *
    0.2513
    0.3156
berr =
   1.0e-16 *
    0.9924
    0.8270
info =
           0
```