

NAG Library Routine Document

E04WDF

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.

Note: *this routine uses **optional parameters** to define choices in the problem specification and in the details of the algorithm. If you wish to use default settings for all of the optional parameters, you need only read Sections 1 to 10 of this document. If, however, you wish to reset some or all of the settings please refer to Section 11 for a detailed description of the algorithm, to Section 12 for a detailed description of the specification of the optional parameters and to Section 13 for a detailed description of the monitoring information produced by the routine.*

1 Purpose

E04WDF is designed to minimize an arbitrary smooth function subject to constraints (which may include simple bounds on the variables, linear constraints and smooth nonlinear constraints) using a sequential quadratic programming (SQP) method. As many first derivatives as possible should be supplied by you; any unspecified derivatives are approximated by finite differences. It is not intended for large sparse problems.

E04WDF may also be used for unconstrained, bound-constrained and linearly constrained optimization.

E04WDF uses **forward communication** for evaluating the objective function, the nonlinear constraint functions, and any of their derivatives.

The initialization routine E04WCF **must** have been called before to calling E04WDF.

2 Specification

```
SUBROUTINE E04WDF (N, NCLIN, NCNLN, LDA, LDCJ, LDH, A, BL, BU, CONFUN,      &
                   OBJFUN, MAJITS, ISTATE, CCON, CJAC, CLAMDA, OBJF,      &
                   GRAD, H, X, IW, LENIW, RW, LENRW, IUSER, RUSER,      &
                   IFAIL)
INTEGER              N, NCLIN, NCNLN, LDA, LDCJ, LDH, MAJITS,              &
                   ISTATE(N+NCLIN+NCNLN), IW(LENIW), LENIW, LENRW,      &
                   IUSER(*), IFAIL
REAL (KIND=nag_wp)  A(LDA,*), BL(N+NCLIN+NCNLN), BU(N+NCLIN+NCNLN),      &
                   CCON(max(1,NCNLN)), CJAC(LDCJ,*),                    &
                   CLAMDA(N+NCLIN+NCNLN), OBJF, GRAD(N), H(LDH,*),      &
                   X(N), RW(LENRW), RUSER(*)
EXTERNAL             CONFUN, OBJFUN
```

Before calling E04WDF, or any of the option setting routines E04WEF, E04WFF, E04WGF or E04WHF, E04WCF **must** be called. The specification for E04WCF is:

```
SUBROUTINE E04WCF (IW, LENIW, RW, LENRW, IFAIL)
INTEGER              IW(LENIW), LENIW, LENRW, IFAIL
REAL (KIND=nag_wp)  RW(LENRW)
```

E04WCF **must** be called with LENIW and LENRW, the declared lengths of IW and RW respectively, satisfying:

$$\text{LENIW} \geq 600;$$

$$\text{LENRW} \geq 600.$$

The contents of the arrays IW and RW **must not** be altered between calls of the routines E04WCF, E04WDF, E04WEF, E04WGF or E04WHF.

3 Description

E04WDF is designed to solve nonlinear programming problems – the minimization of a smooth nonlinear function subject to a set of constraints on the variables. E04WDF is suitable for small dense problems. The problem is assumed to be stated in the following form:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} F(x) \quad \text{subject to} \quad l \leq \begin{pmatrix} x \\ A_L x \\ c(x) \end{pmatrix} \leq u, \quad (1)$$

where $F(x)$ (the *objective function*) is a nonlinear scalar function, A_L is an n_L by n constant matrix, and $c(x)$ is an n_N -vector of nonlinear constraint functions. (The matrix A_L and the vector $c(x)$ may be empty.) The objective function and the constraint functions are assumed to be smooth, here meaning at least twice-continuously differentiable. (The method of E04WDF will usually solve (1) if there are only isolated discontinuities away from the solution.) We also write $r(x)$ for the vector of combined functions:

$$r(x) = \begin{pmatrix} x & A_L x & c(x) \end{pmatrix}^T.$$

Note that although the bounds on the variables could be included in the definition of the linear constraints, we prefer to distinguish between them for reasons of computational efficiency. For the same reason, the linear constraints should **not** be included in the definition of the nonlinear constraints. Upper and lower bounds are specified for all the variables and for all the constraints. An *equality* constraint on r_i can be specified by setting $l_i = u_i$. If certain bounds are not present, the associated elements of l or u can be set to special values that will be treated as $-\infty$ or $+\infty$. (See the description of the optional parameter **Infinite Bound Size**.)

A typical invocation of E04WDF is:

```
CALL E04WCF (IW, LENIW, ...)
CALL E04WEF (ISPECS, IW, ...)
CALL E04WDF (N, NCLIN, NCNLN, ...)
```

where E04WEF reads a set of optional parameter definitions from the file with unit number ISPECS.

Figure 1 illustrates the feasible region for the j th pair of constraints $l_j \leq r_j(x) \leq u_j$. The quantity of δ is the **Feasibility Tolerance**, which can be set by you (see Section 12). The constraints $l_j \leq r_j \leq u_j$ are considered ‘satisfied’ if r_j lies in Regions 2, 3 or 4, and ‘inactive’ if r_j lies in Region 3. The constraint $r_j \geq l_j$ is considered ‘active’ in Region 2, and ‘violated’ in Region 1. Similarly, $r_j \leq u_j$ is active in Region 4, and violated in Region 5. For equality constraints ($l_j = u_j$), Regions 2 and 4 are the same and Region 3 is empty.

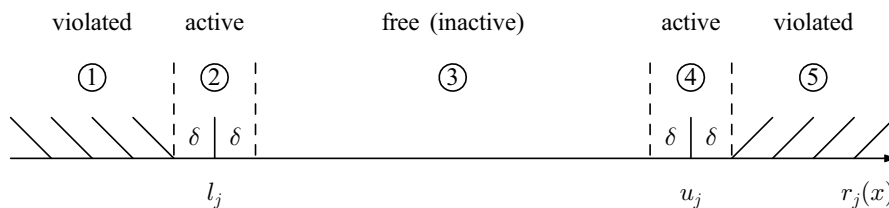


Figure 1

Illustration of the constraints $l_j \leq r_j(x) \leq u_j$

If there are no nonlinear constraints in (1) and F is linear or quadratic, then it will generally be more efficient to use one of E04MFF/E04MFA, E04NCF/E04NCA or E04NFF/E04NFA. If the problem is large and sparse and does have nonlinear constraints, then E04VHF should be used, since E04WDF treats all matrices as dense.

You must supply an initial estimate of the solution to (1), together with subroutines that define $F(x)$ and $c(x)$ with as many first partial derivatives as possible; unspecified derivatives are approximated by finite differences; see Section 12.1 for a discussion of the optional parameter **Derivative Level**.

The objective function is defined by OBJFUN, and the nonlinear constraints are defined by CONFUN. Note that if there *are* any nonlinear constraints then the *first* call to CONFUN will precede the *first* call to OBJFUN.

For maximum reliability, it is preferable for you to provide all partial derivatives (see Chapter 8 of Gill *et al.* (1981), for a detailed discussion). If all gradients cannot be provided, it is similarly advisable to provide as many as possible. While developing OBJFUN and CONFUN, the optional parameter **Verify Level** should be used to check the calculation of any known gradients.

The method used by E04WDF is based on NPOPT, which is part of the SNOPT package described in Gill *et al.* (2005b), and the algorithm it uses is described in detail in Section 11.

4 References

- Eldersveld S K (1991) Large-scale sequential quadratic programming algorithms *PhD Thesis* Department of Operations Research, Stanford University, Stanford
- Fourer R (1982) Solving staircase linear programs by the simplex method *Math. Programming* **23** 274–313
- Gill P E, Murray W and Saunders M A (2002) *SNOPT: An SQP Algorithm for Large-scale Constrained Optimization* **12** 979–1006 SIAM J. Optim.
- Gill P E, Murray W and Saunders M A (2005a) Users' guide for SQOPT 7: a Fortran package for large-scale linear and quadratic programming *Report NA 05-1* Department of Mathematics, University of California, San Diego <http://www.ccom.ucsd.edu/~peg/papers/sqdoc7.pdf>
- Gill P E, Murray W and Saunders M A (2005b) Users' guide for SNOPT 7.1: a Fortran package for large-scale linear nonlinear programming *Report NA 05-2* Department of Mathematics, University of California, San Diego <http://www.ccom.ucsd.edu/~peg/papers/sndoc7.pdf>
- Gill P E, Murray W, Saunders M A and Wright M H (1986) Users' guide for NPSOL (Version 4.0): a Fortran package for nonlinear programming *Report SOL 86-2* Department of Operations Research, Stanford University
- Gill P E, Murray W, Saunders M A and Wright M H (1992) Some theoretical properties of an augmented Lagrangian merit function *Advances in Optimization and Parallel Computing* (ed P M Pardalos) 101–128 North Holland
- Gill P E, Murray W and Wright M H (1981) *Practical Optimization* Academic Press
- Hock W and Schittkowski K (1981) *Test Examples for Nonlinear Programming Codes. Lecture Notes in Economics and Mathematical Systems* **187** Springer–Verlag

5 Arguments

- | | | |
|----|--|--------------|
| 1: | N – INTEGER | <i>Input</i> |
| | <i>On entry:</i> n , the number of variables. | |
| | <i>Constraint:</i> $N > 0$. | |
| | | |
| 2: | NCLIN – INTEGER | <i>Input</i> |
| | <i>On entry:</i> n_L , the number of general linear constraints. | |
| | <i>Constraint:</i> $NCLIN \geq 0$. | |
| | | |
| 3: | NCNLN – INTEGER | <i>Input</i> |
| | <i>On entry:</i> n_N , the number of nonlinear constraints. | |
| | <i>Constraint:</i> $NCNLN \geq 0$. | |

- 4: LDA – INTEGER *Input*
On entry: the first dimension of the array A as declared in the (sub)program from which E04WDF is called.
Constraint: $LDA \geq \max(1, NCLIN)$.
- 5: LDCJ – INTEGER *Input*
On entry: the first dimension of the array CJAC as declared in the (sub)program from which E04WDF is called.
Constraint: $LDCJ \geq \max(1, NCNLN)$.
- 6: LDH – INTEGER *Input*
On entry: the first dimension of the array H as declared in the (sub)program from which E04WDF is called.
Constraint: $LDH \geq N$ unless optional parameter **Hessian Limited Memory** is in effect.
- 7: A(LDA,*) – REAL (KIND=nag_wp) array *Input*
Note: the second dimension of the array A must be at least N if $NCLIN > 0$, and at least 1 otherwise.
On entry: the i th row of A contains the i th row of the matrix A_L of general linear constraints in (1). That is, the i th row contains the coefficients of the i th general linear constraint, for $i = 1, 2, \dots, NCLIN$.
 If $NCLIN = 0$, the array A is not referenced.
- 8: BL(N + NCLIN + NCNLN) – REAL (KIND=nag_wp) array *Input*
 9: BU(N + NCLIN + NCNLN) – REAL (KIND=nag_wp) array *Input*
On entry: BL must contain the lower bounds and BU the upper bounds for all the constraints, in the following order. The first n elements of each array must contain the bounds on the variables, the next n_L elements the bounds for the general linear constraints (if any) and the next n_N elements the bounds for the general nonlinear constraints (if any). To specify a nonexistent lower bound (i.e., $l_j = -\infty$), set $BL(j) \leq -bigbnd$, and to specify a nonexistent upper bound (i.e., $u_j = +\infty$), set $BU(j) \geq bigbnd$; where $bigbnd$ is the optional parameter **Infinite Bound Size**. To specify the j th constraint as an *equality*, set $BL(j) = BU(j) = \beta$, say, where $|\beta| < bigbnd$.
Constraints:

$$BL(j) \leq BU(j), \text{ for } j = 1, 2, \dots, N + NCLIN + NCNLN;$$

$$\text{if } BL(j) = BU(j) = \beta, |\beta| < bigbnd.$$
- 10: CONFUN – SUBROUTINE, supplied by the NAG Library or the user. *External Procedure*
 CONFUN must calculate the vector $c(x)$ of nonlinear constraint functions and (optionally) its Jacobian, $\frac{\partial c}{\partial x}$, for a specified n -vector x . If there are no nonlinear constraints (i.e., $NCNLN = 0$), E04WDF will never call CONFUN, so it may be the dummy routine E04WDP. (E04WDP is included in the NAG Library). If there are nonlinear constraints, the first call to CONFUN will occur before the first call to OBJFUN.
 If all constraint gradients (Jacobian elements) are known (i.e., **Derivative Level** = 2 or 3), any *constant* elements may be assigned to CJAC once only at the start of the optimization. An element of CJAC that is not subsequently assigned in CONFUN will retain its initial value throughout. Constant elements may be loaded in CJAC during the first call to CONFUN (signalled by the value of $NSTATE = 1$). The ability to preload constants is useful when many Jacobian elements are identically zero, in which case CJAC may be initialized to zero and nonzero elements may be reset by CONFUN.

It must be emphasized that, if **Derivative Level** < 2 , unassigned elements of CJAC are *not* treated as constant; they are estimated by finite differences, at nontrivial expense.

The specification of CONFUN is:

```
SUBROUTINE CONFUN (MODE, NCNLN, N, LDCJ, NEEDC, X, CCON, CJAC,      &
                   NSTATE, IUSER, RUSER)
INTEGER               MODE, NCNLN, N, LDCJ, NEEDC(NCNLN), NSTATE,      &
                   IUSER(*)
REAL (KIND=nag_wp) X(N), CCON(max(1,NCNLN)), CJAC(LDCJ,N),          &
                   RUSER(*)
```

1: **MODE** – INTEGER *Input/Output*

On entry: is set by E04WDF to indicate which values must be assigned during each call of CONFUN. Only the following values need be assigned, for each value of i such that $NEEDC(i) > 0$:

MODE = 0

The components of CCON corresponding to positive values in NEEDC must be set. Other components and the array CJAC are ignored.

MODE = 1

The known components of the rows of CJAC corresponding to positive values in NEEDC must be set. Other rows of CJAC and the array CCON will be ignored.

MODE = 2

Only the elements of CCON corresponding to positive values of NEEDC need to be set (and similarly for the known components of the rows of CJAC).

On exit: may be used to indicate that you are unable or unwilling to evaluate the constraint functions at the current x .

During the linesearch, the constraint functions are evaluated at points of the form $x = x_k + \alpha p_k$ after they have already been evaluated satisfactorily at x_k . At any such α , if you set **MODE** = -1, E04WDF will evaluate the functions at some point closer to x_k (where they are more likely to be defined).

If for some reason you wish to terminate the current problem, set **MODE** < -1 .

2: **NCNLN** – INTEGER *Input*

On entry: n_N , the number of nonlinear constraints.

3: **N** – INTEGER *Input*

On entry: n , the number of variables.

4: **LDCJ** – INTEGER *Input*

On entry: the first dimension of the array CJAC as declared in the (sub)program from which E04WDF is called.

5: **NEEDC(NCNLN)** – INTEGER array *Input*

On entry: the indices of the elements of CCON and/or CJAC that must be evaluated by CONFUN. If $NEEDC(i) > 0$, the i th element of CCON and/or the available elements of the i th row of CJAC (see argument **MODE**) must be evaluated at x .

6: **X(N)** – REAL (KIND=nag_wp) array *Input*

On entry: x , the vector of variables at which the constraint functions and/or the available elements of the constraint Jacobian are to be evaluated.

7:	CCON(max(1,NCNLN)) – REAL (KIND=nag_wp) array	<i>Output</i>
	<i>On exit:</i> if NEEDC(<i>i</i>) > 0 and MODE = 0 or 2, CCON(<i>i</i>) must contain the value of the <i>i</i> th constraint at <i>x</i> . The remaining elements of CCON, corresponding to the non-positive elements of NEEDC, are ignored.	
8:	CJAC(LDCJ,N) – REAL (KIND=nag_wp) array	<i>Input/Output</i>
	<i>On entry:</i> the elements of CJAC are set to special values that enable E04WDF to detect whether they are reset by CONFUN.	
	<i>On exit:</i> if NEEDC(<i>i</i>) > 0 and MODE = 1 or 2, the <i>i</i> th row of CJAC must contain the available elements of the vector ∇c_i given by	
	$\nabla c_i = \left(\frac{\partial c_i}{\partial x_1}, \frac{\partial c_i}{\partial x_2}, \dots, \frac{\partial c_i}{\partial x_n} \right)^T,$	
	where $\frac{\partial c_i}{\partial x_j}$ is the partial derivative of the <i>i</i> th constraint with respect to the <i>j</i> th variable, evaluated at the point <i>x</i> . See also the argument NSTATE. The remaining rows of CJAC, corresponding to non-positive elements of NEEDC, are ignored.	
	If all elements of the constraint Jacobian are known (i.e., Derivative Level = 2 or 3), any constant elements may be assigned to CJAC one time only at the start of the optimization. An element of CJAC that is not subsequently assigned in CONFUN will retain its initial value throughout. Constant elements may be loaded into CJAC during the first call to CONFUN (signalled by the value NSTATE = 1). The ability to preload constants is useful when many Jacobian elements are identically zero, in which case CJAC may be initialized to zero and nonzero elements may be reset by CONFUN.	
	Note that constant nonzero elements do affect the values of the constraints. Thus, if CJAC(<i>i</i> , <i>j</i>) is set to a constant value, it need not be reset in subsequent calls to CONFUN, but the value CJAC(<i>i</i> , <i>j</i>) × X(<i>j</i>) must nonetheless be added to CCON(<i>i</i>). For example, if CJAC(1,1) = 2 and CJAC(1,2) = −5 then the term 2 × X(1) − 5 × X(2) must be included in the definition of CCON(1).	
	It must be emphasized that, if Derivative Level = 0 or 1, unassigned elements of CJAC are not treated as constant; they are estimated by finite differences, at nontrivial expense. If you do not supply a value for the optional parameter Difference Interval , an interval for each element of <i>x</i> is computed automatically at the start of the optimization. The automatic procedure can usually identify constant elements of CJAC, which are then computed once only by finite differences.	
9:	NSTATE – INTEGER	<i>Input</i>
	<i>On entry:</i> if NSTATE = 1 then E04WDF is calling CONFUN for the first time. This argument setting allows you to save computation time if certain data must be read or calculated only once.	
10:	IUSER(*) – INTEGER array	<i>User Workspace</i>
11:	RUSER(*) – REAL (KIND=nag_wp) array	<i>User Workspace</i>
	CONFUN is called with the arguments IUSER and RUSER as supplied to E04WDF. You should use the arrays IUSER and RUSER to supply information to CONFUN.	

CONFUN must either be a module subprogram USED by, or declared as EXTERNAL in, the (sub)program from which E04WDF is called. Arguments denoted as *Input* must **not** be changed by this procedure.

CONFUN should be tested separately before being used in conjunction with E04WDF. See also the description of the optional parameter **Verify Level**.

11: OBJFUN – SUBROUTINE, supplied by the user.

External Procedure

OBJFUN must calculate the objective function $F(x)$ and (optionally) its gradient $g(x) = \frac{\partial F}{\partial x}$ for a specified n -vector x .

The specification of OBJFUN is:

```
SUBROUTINE OBJFUN (MODE, N, X, OBJF, GRAD, NSTATE, IUSER, RUSER)
  INTEGER                MODE, N, NSTATE, IUSER(*)
  REAL (KIND=nag_wp) X(N), OBJF, GRAD(N), RUSER(*)
```

1: MODE – INTEGER

Input/Output

On entry: is set by E04WDF to indicate which values must be assigned during each call of OBJFUN. Only the following values need be assigned:

MODE = 0
OBJF.

MODE = 1
All available elements of GRAD.

MODE = 2
OBJF and all available elements of GRAD.

On exit: may be used to indicate that you are unable or unwilling to evaluate the objective function at the current x .

During the linesearch, the function is evaluated at points of the form $x = x_k + \alpha p_k$ after they have already been evaluated satisfactorily at x_k . For any such x , if you set MODE = -1, E04WDF will reduce α and evaluate the functions again (closer to x_k , where they are more likely to be defined).

If for some reason you wish to terminate the current problem, set MODE < -1.

2: N – INTEGER

Input

On entry: n , the number of variables.

3: X(N) – REAL (KIND=nag_wp) array

Input

On entry: x , the vector of variables at which the objective function and/or all available elements of its gradient are to be evaluated.

4: OBJF – REAL (KIND=nag_wp)

Output

On exit: if MODE = 0 or 2, OBJF must be set to the value of the objective function at x .

5: GRAD(N) – REAL (KIND=nag_wp) array

Input/Output

On entry: the elements of GRAD are set to special values.

On exit: if MODE = 1 or 2, GRAD must return the available elements of the gradient evaluated at x , i.e., GRAD(i) contains the partial derivative $\frac{\partial F}{\partial x_i}$.

6: NSTATE – INTEGER

Input

On entry: if NSTATE = 1 then E04WDF is calling OBJFUN for the first time. This argument setting allows you to save computation time if certain data must be read or calculated only once.

7:	IUSER(*) – INTEGER array	<i>User Workspace</i>
8:	RUSER(*) – REAL (KIND=nag_wp) array	<i>User Workspace</i>
OBJFUN is called with the arguments IUSER and RUSER as supplied to E04WDF. You should use the arrays IUSER and RUSER to supply information to OBJFUN.		

OBJFUN must either be a module subprogram USED by, or declared as EXTERNAL in, the (sub) program from which E04WDF is called. Arguments denoted as *Input* must **not** be changed by this procedure.

OBJFUN should be tested separately before being used in conjunction with E04WDF. See also the description of the optional parameter **Verify Level**.

12: MAJITS – INTEGER *Output*

On exit: the number of major iterations performed.

13: ISTATE(N + NCLIN + NCNLN) – INTEGER array *Input/Output*

On entry: is an integer array that need not be initialized if E04WDF is called with the **Cold Start** option (the default).

If optional parameter **Warm Start** has been chosen, every element of ISTATE must be set. If E04WDF has just been called on a problem with the same dimensions, ISTATE already contains valid values. Otherwise, ISTATE(*j*) should indicate whether either of the constraints $r_j(x) \geq l_j$ or $r_j(x) \leq u_j$ is expected to be active at a solution of (1).

The ordering of ISTATE is the same as for BL, BU and $r(x)$, i.e., the first N components of ISTATE refer to the upper and lower bounds on the variables, the next NCLIN refer to the bounds on $A_L x$, and the last NCNLN refer to the bounds on $c(x)$. Possible values of ISTATE(*i*) follow:

- 0 Neither $r_j(x) \geq l_j$ nor $r_j(x) \leq u_j$ is expected to be active.
- 1 $r_j(x) \geq l_j$ is expected to be active.
- 2 $r_j(x) \leq u_j$ is expected to be active.
- 3 This may be used if $l_j = u_j$. Normally an equality constraint $r_j(x) = l_j = u_j$ is active at a solution.

The values 1, 2 or 3 all have the same effect when $BL(j) = BU(j)$. If necessary, E04WDF will override your specification of ISTATE, so that a poor choice will not cause the algorithm to fail.

On exit: describes the status of the constraints $l \leq r(x) \leq u$. For the *j*th lower or upper bound, $j = 1, 2, \dots, N + NCLIN + NCNLN$, the possible values of ISTATE(*j*) are as follows (see Figure 1). δ is the appropriate feasibility tolerance.

- 2 (Region 1) The lower bound is violated by more than δ .
- 1 (Region 5) The upper bound is violated by more than δ .
- 0 (Region 3) Both bounds are satisfied by more than δ .
- 1 (Region 2) The lower bound is active (to within δ).
- 2 (Region 4) The upper bound is active (to within δ).
- 3 (Region 2 = Region 4) The bounds are equal and the equality constraint is satisfied (to within δ).

These values of ISTATE are labelled in the printed solution according to Table 1.

Region	1	2	3	4	5	
$2 \equiv 4$						

ISTATE(<i>j</i>)	-2	1	0	2	-1	3
Printed solution	--	LL	FR	UL	++	EQ

Table 1

Labels used in the printed solution for the regions in Figure 1

- 14: CCON(max(1,NCNLN)) – REAL (KIND=nag_wp) array *Input/Output*
On entry: CCON need not be initialized if the (default) optional parameter **Cold Start** is used.
For a **Warm Start**, and if NCNLN > 0, CCON contains values of the nonlinear constraint functions c_i , for $i = 1, 2, \dots, \text{NCNLN}$, calculated in a previous call to E04WDF.
On exit: if NCNLN > 0, CCON(*i*) contains the value of the *i*th nonlinear constraint function c_i at the final iterate, for $i = 1, 2, \dots, \text{NCNLN}$.
If NCNLN = 0, the array CCON is not referenced.
- 15: CJAC(LDCJ,*) – REAL (KIND=nag_wp) array *Input/Output*
Note: the second dimension of the array CJAC must be at least N if NCNLN > 0, and at least 1 otherwise.
On entry: in general, CJAC need not be initialized before the call to E04WDF. However, if **Derivative Level** = 2 or 3, any constant elements of CJAC may be initialized. Such elements need not be reassigned on subsequent calls to CONFUN.
On exit: if NCNLN > 0, CJAC contains the Jacobian matrix of the nonlinear constraint functions at the final iterate, i.e., CJAC(*i,j*) contains the partial derivative of the *i*th constraint function with respect to the *j*th variable, for $i = 1, 2, \dots, \text{NCNLN}$ and $j = 1, 2, \dots, N$. (See the discussion of argument CJAC under CONFUN.)
If NCNLN = 0, the array CJAC is not referenced.
- 16: CLAMDA(N + NCLIN + NCNLN) – REAL (KIND=nag_wp) array *Input/Output*
On entry: need not be set if the (default) optional parameter **Cold Start** is used.
If the optional parameter **Warm Start** has been chosen, CLAMDA(*j*) must contain a multiplier estimate for each nonlinear constraint, with a sign that matches the status of the constraint specified by the ISTATE array, for $j = N + \text{NCLIN} + 1, \dots, N + \text{NCLIN} + \text{NCNLN}$. The remaining elements need not be set. If the *j*th constraint is defined as ‘inactive’ by the initial value of the ISTATE array (i.e., ISTATE(*j*) = 0), CLAMDA(*j*) should be zero; if the *j*th constraint is an inequality active at its lower bound (i.e., ISTATE(*j*) = 1), CLAMDA(*j*) should be non-negative; if the *j*th constraint is an inequality active at its upper bound (i.e., ISTATE(*j*) = 2), CLAMDA(*j*) should be non-positive. If necessary, the routine will modify CLAMDA to match these rules.
On exit: the values of the QP multipliers from the last QP subproblem. CLAMDA(*j*) should be non-negative if ISTATE(*j*) = 1 and non-positive if ISTATE(*j*) = 2.
- 17: OBJF – REAL (KIND=nag_wp) *Output*
On exit: the value of the objective function at the final iterate.
- 18: GRAD(N) – REAL (KIND=nag_wp) array *Output*
On exit: the gradient of the objective function (or its finite difference approximation) at the final iterate.

- 19: H(LDH,*) – REAL (KIND=nag_wp) array Input/Output
- Note:** the second dimension of the array H must be at least N unless the optional parameter **Hessian Limited Memory** is in effect. If **Hessian Limited Memory** is in effect, array H is not referenced.
- On entry:* H need not be initialized if the (default) optional parameter **Cold Start** is used, and will be set to the identity.
- If the optional parameter **Warm Start** has been chosen, H provides the initial approximation of the Hessian of the Lagrangian, i.e., $H(i, j) \approx \frac{\partial^2 \mathcal{L}(x, \lambda)}{\partial x_i \partial x_j}$, where $\mathcal{L}(x, \lambda) = F(x) - c(x)^T \lambda$ and λ is an estimate of the Lagrange multipliers. H must be a positive definite matrix.
- On exit:* contains the Hessian of the Lagrangian at the final estimate x .
- 20: X(N) – REAL (KIND=nag_wp) array Input/Output
- On entry:* an initial estimate of the solution.
- On exit:* the final estimate of the solution.
- 21: IW(LENIW) – INTEGER array Communication Array
- 22: LENIW – INTEGER Input
- On entry:* the dimension of the array IW as declared in the (sub)program from which E04WDF is called.
- Constraint:* LENIW \geq 600.
- 23: RW(LENRW) – REAL (KIND=nag_wp) array Communication Array
- 24: LENRW – INTEGER Input
- On entry:* the dimension of the array RW as declared in the (sub)program from which E04WDF is called.
- Constraint:* LENRW \geq 600.
- 25: IUSER(*) – INTEGER array User Workspace
- 26: RUSER(*) – REAL (KIND=nag_wp) array User Workspace
- IUSER and RUSER are not used by E04WDF, but are passed directly to CONFUN and OBJFUN and should be used to pass information to these routines.
- 27: IFAIL – INTEGER Input/Output
- On entry:* IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this argument you should refer to Section 3.4 in How to Use the NAG Library and its Documentation 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, because for this routine the values of the output arguments may be useful even if IFAIL \neq 0 on exit, the recommended value is -1. **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).
- E04WDF returns with IFAIL = 0 if the iterates have converged to a point x that satisfies the first-order Kuhn–Tucker (see Section 13.2) conditions to the accuracy requested by the **Major Optimality Tolerance**, i.e., the projected gradient and active constraint residuals are negligible at x .
- You should check whether the following four conditions are satisfied:

- (i) the final value of `rgNorm` (see Section 13.2) is significantly less than that at the starting point;
- (ii) during the final major iterations, the values of `Step` and `Minors` (see Section 13.1) are both one;
- (iii) the last few values of both `rgNorm` and `SumInf` (see Section 13.2) become small at a fast linear rate; and
- (iv) `condHz` (see Section 13.1) is small.

If all these conditions hold, x is almost certainly a local minimum of (1).

One caution about ‘Optimal solutions’. Some of the variables or slacks may lie outside their bounds more than desired, especially if scaling was requested. `Max Primal infeas` in the Print file refers to the largest bound infeasibility and which variable is involved. If it is too large, consider restarting with a smaller **Minor Feasibility Tolerance** (say 10 times smaller) and perhaps **Scale Option** = 0.

Similarly, `Max Dual infeas` in the Print file indicates which variable is most likely to be at a nonoptimal value. Broadly speaking, if

$$\text{Max Dual infeas} / \text{Max pi} = 10^{-d},$$

then the objective function would probably change in the d th significant digit if optimization could be continued. If d seems too large, consider restarting with a smaller **Major Optimality Tolerance**.

Finally, `Nonlinear constraint violn` in the Print file shows the maximum infeasibility for nonlinear rows. If it seems too large, consider restarting with a smaller **Major Feasibility Tolerance**.

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`).

Note: E04WDF may return useful information for one or more of the following detected errors or warnings.

Errors or warnings detected by the routine:

`IFAIL` = 1

On entry, `LENIW` = $\langle \text{value} \rangle$.
Constraint: `LENIW` \geq 600.

On entry, `LENRW` = $\langle \text{value} \rangle$.
Constraint: `LENRW` \geq 600.

The initialization routine E04WCF has not been called.

`IFAIL` = 2

Basis file dimensions do not match this problem.

On entry, bounds `BL` and `BU` for $\langle \text{value} \rangle$ are equal and infinite. `BL` = `BU` = $\langle \text{value} \rangle$ and `bigbnd` = $\langle \text{value} \rangle$.

On entry, bounds for $\langle \text{value} \rangle$ are inconsistent. `BL` = $\langle \text{value} \rangle$ and `BU` = $\langle \text{value} \rangle$.

On entry, `LDA` = $\langle \text{value} \rangle$ and `NCLIN` = $\langle \text{value} \rangle$.
Constraint: `LDA` \geq `NCLIN`.

On entry, `LDCJ` = $\langle \text{value} \rangle$ and `NCNLN` = $\langle \text{value} \rangle$.
Constraint: `LDCJ` \geq `NCNLN`.

On entry, LDH = $\langle value \rangle$ and N = $\langle value \rangle$.

Constraint: $LDH \geq N$.

On entry, N = $\langle value \rangle$.

Constraint: $N \geq \langle value \rangle$.

On entry, NCLIN = $\langle value \rangle$.

Constraint: $NCLIN \geq \langle value \rangle$.

On entry, NCNLN = $\langle value \rangle$.

Constraint: $NCNLN \geq \langle value \rangle$.

IFAIL = 3

The requested accuracy could not be achieved.

*A feasible solution has been found, but the requested accuracy in the dual infeasibilities could not be achieved. An abnormal termination has occurred, but E04WDF is within 10^{-2} of satisfying the **Major Optimality Tolerance**. Check that the **Major Optimality Tolerance** is not too small.*

IFAIL = 4

The linear constraints appear to be infeasible.

The problem appears to be infeasible. Infeasibilities have been minimized.

The problem appears to be infeasible. Nonlinear infeasibilities have been minimized.

The problem appears to be infeasible. The linear equality constraints could not be satisfied.

*When the constraints are linear, this message is based on a relatively reliable indicator of infeasibility. Feasibility is measured with respect to the upper and lower bounds on the variables and slacks. Among all the points satisfying the general constraints $Ax - s = 0$ (see (5) and (6) in Section 11.2), there is apparently no point that satisfies the bounds on x and s . Violations as small as the **Minor Feasibility Tolerance** are ignored, but at least one component of x or s violates a bound by more than the tolerance.*

When nonlinear constraints are present, infeasibility is much harder to recognize correctly. Even if a feasible solution exists, the current linearization of the constraints may not contain a feasible point. In an attempt to deal with this situation, when solving each QP subproblem, E04WDF is prepared to relax the bounds on the slacks associated with nonlinear rows.

*If a QP subproblem proves to be infeasible or unbounded (or if the Lagrange multiplier estimates for the nonlinear constraints become large), E04WDF enters so-called 'nonlinear elastic' mode. The subproblem includes the original QP objective and the sum of the infeasibilities – suitably weighted using the optional parameter **Elastic Weight**. In elastic mode, some of the bounds on the nonlinear rows are 'elastic' – i.e., they are allowed to violate their specific bounds. Variables subject to elastic bounds are known as elastic variables. An elastic variable is free to violate one or both of its original upper or lower bounds. If the original problem has a feasible solution and the elastic weight is sufficiently large, a feasible point eventually will be obtained for the perturbed constraints, and optimization can continue on the subproblem. If the nonlinear problem has no feasible solution, E04WDF will tend to determine a 'good' infeasible point if the elastic weight is sufficiently large. (If the elastic weight were infinite, E04WDF would locally minimize the nonlinear constraint violations subject to the linear constraints and bounds.)*

Unfortunately, even though E04WDF locally minimizes the nonlinear constraint violations, there may still exist other regions in which the nonlinear constraints are satisfied. Wherever possible, nonlinear constraints should be defined in such a way that feasible points are known to exist when the constraints are linearized.

IFAIL = 5

The problem appears to be unbounded. The constraint violation limit has been reached.

The problem appears to be unbounded. The objective function is unbounded.

The problem appears to be unbounded (or badly scaled).

For linear problems, unboundedness is detected by the simplex method when a nonbasic variable can be increased or decreased by an arbitrary amount without causing a basic variable to violate a bound. Consider adding an upper or lower bound to the variable. Also, examine the constraints that have nonzeros in the associated column, to see if they have been formulated as intended.

*Very rarely, the scaling of the problem could be so poor that numerical error will give an erroneous indication of unboundedness. Consider using the optional parameter **Scale Option**.*

For nonlinear problems, E04WDF monitors both the size of the current objective function and the size of the change in the variables at each step. If either of these is very large (as judged by the 'Unbounded' parameters (see Section 12)), the problem is terminated and declared unbounded. To avoid large function values, it may be necessary to impose bounds on some of the variables in order to keep them away from singularities in the nonlinear functions.

*The message may indicate an abnormal termination while enforcing the limit on the constraint violations. This exit implies that the objective is not bounded below in the feasible region defined by expanding the bounds by the value of the **Violation Limit**.*

IFAIL = 6

Iteration limit reached.

Major iteration limit reached.

The value of the optional parameter **Superbasics Limit** is too small.

*Either the **Iterations Limit** or the **Major Iterations Limit** was exceeded before the required solution could be found. Check the iteration log to be sure that progress was being made. If so, and if you caused a basis file to be saved by using the optional parameter **New Basis File**, consider restarting the run using the optional parameter **Old Basis File** to see whether further progress can be made. If you have no basis file available, you might rerun the problem after increasing the optional parameters **Minor Iterations Limit** and/or **Major Iterations Limit**.*

*If none of the above limits have been reached, this error may mean that the problem appears to be more nonlinear than anticipated. The current set of basic and superbasic variables have been optimized as much as possible and a pricing operation (where a nonbasic variable is selected to become superbasic) is necessary to continue, but it can't continue as the number of superbasic variables has already reached the limit specified by the optional parameter **Superbasics Limit**. In general, raise the **Superbasics Limit** s by a reasonable amount, bearing in mind the storage needed for the reduced Hessian. (The **Reduced Hessian Dimension** h will also increase to s unless specified otherwise, and the associated storage will be about $\frac{1}{2}s^2$ words.) In some cases you may have to set $h < s$ to conserve storage, but beware that the rate of convergence will probably fall off severely.*

IFAIL = 7

Numerical difficulties have been encountered and no further progress can be made.

Numerical difficulties have been encountered and no further progress can be made.

Several circumstances could lead to this exit.

1. The user-supplied subroutines OBJFUN or CONFUN could be returning accurate function values but inaccurate gradients (or vice versa). This is the most likely cause. Study the comments given for IFAIL = 8, and do your best to ensure that the coding is correct.
2. The function and gradient values could be consistent, but their precision could be too low. For example, accidental use of a real data type when double precision was intended would lead to a relative function precision of about 10^{-6} instead of something like 10^{-15} . The default **Major Optimality Tolerance** of 2×10^{-6} would need to be raised to about 10^{-3} for optimality to be declared (at a rather suboptimal point). Of course, it is better to revise the function coding to obtain as much precision as economically possible.

3. If function values are obtained from an expensive iterative process, they may be accurate to rather few significant figures, and gradients will probably not be available. One should specify

Function Precision t

Major Optimality Tolerance \sqrt{t}

but even then, if t is as large as 10^{-5} or 10^{-6} (only 5 or 6 significant figures), the same exit condition may occur. At present the only remedy is to increase the accuracy of the function calculation.

4. An LU factorization of the basis has just been obtained and used to recompute the basic variables x_B , given the present values of the superbasic and nonbasic variables. A step of ‘iterative refinement’ has also been applied to increase the accuracy of x_B . However, a row check has revealed that the resulting solution does not satisfy the current constraints $Ax - s = 0$ sufficiently well.

This probably means that the current basis is very ill-conditioned. If there are some linear constraints and variables, try **Scale Option** = 1 if scaling has not yet been used.

For certain highly structured basis matrices (notably those with band structure), a systematic growth may occur in the factor U . Consult the description of **Umax** and **Growth** in Section 13.4 and set the **LU Factor Tolerance** to 2.0 (or possibly even smaller, but not less than 1.0).

5. The first factorization attempt will have found the basis to be structurally or numerically singular. (Some diagonals of the triangular matrix U were respectively zero or smaller than a certain tolerance.) The associated variables are replaced by slacks and the modified basis is refactorized, but singularity persists. This must mean that the problem is badly scaled, or the **LU Factor Tolerance** is too much larger than 1.0. This is highly unlikely to occur.

IFAIL = 8

User-supplied function computes incorrect constraint derivatives.

User-supplied function computes incorrect objective derivatives.

If the message refers to the derivatives of the objective function, then a check has been made on some individual elements of the objective gradient array at the first point that satisfies the linear constraints. At least one component GRADj is being set to a value that disagrees markedly with its associated forward-difference estimate $\frac{\partial F}{\partial x_j}$. (The relative difference between the computed and estimated values is 1.0 or more.) This exit is a safeguard, since E04WDF will usually fail to make progress when the computed gradients are seriously inaccurate. In the process it may expend considerable effort before terminating with IFAIL = 7.

Check the function and gradient computation very carefully in OBJFUN. A simple omission could explain everything. If F or a component $\frac{\partial F}{\partial x_j}$ is very large, then give serious thought to scaling the function or the nonlinear variables.

*If you feel certain that the computed GRAD(j) is correct (and that the forward-difference estimate is therefore wrong), you can specify **Verify Level** = 0 to prevent individual elements from being checked. However, the optimization procedure may have difficulty.*

If the message refers to derivatives of the constraints, then at least one of the computed constraint derivatives is significantly different from an estimate obtained by forward-differencing the vector $c(x)$. Follow the advice given above, trying to ensure that the arrays CCON and CJAC are being set correctly in CONFUN.

IFAIL = 9

Unable to proceed into undefined region of user-supplied function.

User-supplied function is undefined at the first feasible point.

User-supplied function is undefined at the initial point.

*You have indicated that the problem functions are undefined by assigning the value $MODE = -1$ on exit from *OBJFUN* or *CONFUN*. E04WDF attempts to evaluate the problem functions closer to a point at which the functions are already known to be defined. This exit occurs if E04WDF is unable to find a point at which the functions are defined. This will occur in the case of:*

- undefined functions with no recovery possible;
- undefined functions at the first point;
- undefined functions at the first feasible point; or
- undefined functions when checking derivatives.

IFAIL = 10

User-supplied constraint function requested termination.

User-supplied objective function requested termination.

*You have indicated the wish to terminate solution of the current problem by setting $MODE$ to a value < -1 on exit from *OBJFUN* or *CONFUN*.*

IFAIL = 11

Internal error: memory allocation failed when attempting to allocate workspace sizes $\langle value \rangle$ and $\langle value \rangle$. Please contact NAG.

IFAIL = 12

Internal memory allocation was insufficient. Please contact NAG.

IFAIL = 13

An error has occurred in the basis package, perhaps indicating incorrect setup of arrays. Set the optional parameter **Print File** and examine the output carefully for further information.

IFAIL = 14

An unexpected error has occurred. Set the optional parameter **Print File** and examine the output carefully for further information.

IFAIL = -99

An unexpected error has been triggered by this routine. Please contact NAG.

See Section 3.9 in How to Use the NAG Library and its Documentation for further information.

IFAIL = -399

Your licence key may have expired or may not have been installed correctly.

See Section 3.8 in How to Use the NAG Library and its Documentation for further information.

IFAIL = -999

Dynamic memory allocation failed.

See Section 3.7 in How to Use the NAG Library and its Documentation for further information.

7 Accuracy

If IFAIL = 0 on exit, then the vector returned in the array X is an estimate of the solution to an accuracy of approximately **Major Optimality Tolerance**.

8 Parallelism and Performance

E04WDF makes calls to BLAS and/or LAPACK routines, which may be threaded within the vendor library used by this implementation. Consult the documentation for the vendor library for further information.

Please consult the X06 Chapter Introduction for information on how to control and interrogate the OpenMP environment used within this routine. Please also consult the Users' Note for your implementation for any additional implementation-specific information.

9 Further Comments

This section describes the final output produced by E04WDF. Intermediate and other output are given in Section 13.

9.1 The Final Output

If **Print File** > 0, the final output, including a listing of the status of every variable and constraint will be sent to the **Print File**. The following describes the output for each variable. A full stop (.) is printed for any numerical value that is zero.

Variable	gives the name (Variable) and index j , for $j = 1, 2, \dots, n$, of the variable.
State	gives the state of the variable (FR if neither bound is in the working set, EQ if a fixed variable, LL if on its lower bound, UL if on its upper bound, TF if temporarily fixed at its current value). If Value lies outside the upper or lower bounds by more than the Feasibility Tolerance , State will be ++ or -- respectively. (The latter situation can occur only when there is no feasible point for the bounds and linear constraints.) A key is sometimes printed before State. A <i>Alternative optimum possible</i> . The variable is active at one of its bounds, but its Lagrange multiplier is essentially zero. This means that if the variable were allowed to start moving away from its bound then there would be no change to the objective function. The values of the other free variables <i>might</i> change, giving a genuine alternative solution. However, if there are any degenerate variables (labelled D), the actual change might prove to be zero, since one of them could encounter a bound immediately. In either case the values of the Lagrange multipliers might also change. D <i>Degenerate</i> . The variable is free, but it is equal to (or very close to) one of its bounds. I <i>Infeasible</i> . The variable is currently violating one of its bounds by more than the Feasibility Tolerance .
Value	is the value of the variable at the final iteration.
Lower bound	is the lower bound specified for the variable. None indicates that $BL(j) \leq -bigbnd$.
Upper bound	is the upper bound specified for the variable. None indicates that $BU(j) \geq bigbnd$.
Lagr multiplier	is the Lagrange multiplier for the associated bound. This will be zero if State is FR unless $BL(j) \leq -bigbnd$ and $BU(j) \geq bigbnd$, in which case the entry will be blank. If x is optimal, the multiplier should be non-negative if State is LL and non-positive if State is UL.
Slack	is the difference between the variable Value and the nearer of its (finite) bounds $BL(j)$ and $BU(j)$. A blank entry indicates that the associated variable is not bounded (i.e., $BL(j) \leq -bigbnd$ and $BU(j) \geq bigbnd$).

The meaning of the output for linear and nonlinear constraints is the same as that given above for variables, with $BL(j)$ and $BU(j)$ replaced by $BL(n+j)$ and $BU(n+j)$ respectively, and with the following changes in the heading:

Linear constrnt gives the name (`lincon`) and index j , for $j = 1, 2, \dots, n_L$, of the linear constraint.
Nonlin constrnt gives the name (`nlncn`) and index $(j - n_L)$, for $j = n_L + 1, \dots, n_L + n_N$, of the nonlinear constraint.

Note that movement off a constraint (as opposed to a variable moving away from its bound) can be interpreted as allowing the entry in the `Slack` column to become positive.

Numerical values are output with a fixed number of digits; they are not guaranteed to be accurate to this precision.

10 Example

This example is based on Problem 71 in Hock and Schittkowski (1981) and involves the minimization of the nonlinear function

$$F(x) = x_1 x_4 (x_1 + x_2 + x_3) + x_3$$

subject to the bounds

$$\begin{aligned} 1 &\leq x_1 \leq 5 \\ 1 &\leq x_2 \leq 5 \\ 1 &\leq x_3 \leq 5 \\ 1 &\leq x_4 \leq 5, \end{aligned}$$

to the general linear constraint

$$x_1 + x_2 + x_3 + x_4 \leq 20,$$

and to the nonlinear constraints

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 + x_4^2 &\leq 40, \\ x_1 x_2 x_3 x_4 &\geq 25. \end{aligned}$$

The initial point, which is infeasible, is

$$x_0 = (1, 5, 5, 1)^T,$$

with $F(x_0) = 16$.

The optimal solution (to five figures) is

$$x^* = (1.0, 4.7430, 3.8211, 1.3794)^T,$$

and $F(x^*) = 17.014$. One bound constraint and both nonlinear constraints are active at the solution.

10.1 Program Text

```
!   E04WDF Example Program Text
!   Mark 26 Release. NAG Copyright 2016.

Module e04wdfe_mod

!       E04WDF Example Program Module:
!           Parameters and User-defined Routines

!       .. Use Statements ..
!       Use nag_library, Only: nag_wp
!       .. Implicit None Statement ..
!       Implicit None
!       .. Accessibility Statements ..
!       Private
!       Public                                :: confun, objfun
!       .. Parameters ..
```

```

Integer, Parameter, Public      :: leniw = 600, lenrw = 600, nin = 5,    &
                                nout = 6
Contains
Subroutine objfun(mode,n,x,objf,grad,nstate,iuser,ruser)
!   Routine to evaluate objective function and its 1st derivatives.

!   .. Scalar Arguments ..
Real (Kind=nag_wp), Intent (Out) :: objf
Integer, Intent (Inout)          :: mode
Integer, Intent (In)             :: n, nstate
!   .. Array Arguments ..
Real (Kind=nag_wp), Intent (Inout) :: grad(n), ruser(*)
Real (Kind=nag_wp), Intent (In) :: x(n)
Integer, Intent (Inout)          :: iuser(*)
!   .. Executable Statements ..
If (mode==0 .Or. mode==2) Then
  objf = x(1)*x(4)*(x(1)+x(2)+x(3)) + x(3)
End If

If (mode==1 .Or. mode==2) Then
  grad(1) = x(4)*(2.0E0_nag_wp*x(1)+x(2)+x(3))
  grad(2) = x(1)*x(4)
  grad(3) = x(1)*x(4) + 1.0E0_nag_wp
  grad(4) = x(1)*(x(1)+x(2)+x(3))
End If

Return

End Subroutine objfun
Subroutine confun(mode,ncnln,n,ldcj,needc,x,ccon,cjac,nstate,iuser,    &
  ruser)
!   Routine to evaluate the nonlinear constraints and their 1st
!   derivatives.

!   .. Scalar Arguments ..
Integer, Intent (In)          :: ldcj, n, ncnln, nstate
Integer, Intent (Inout)       :: mode
!   .. Array Arguments ..
Real (Kind=nag_wp), Intent (Out) :: ccon(max(1,ncnln))
Real (Kind=nag_wp), Intent (Inout) :: cjac(ldcj,n), ruser(*)
Real (Kind=nag_wp), Intent (In) :: x(n)
Integer, Intent (Inout)         :: iuser(*)
Integer, Intent (In)            :: needc(ncnln)
!   .. Intrinsic Procedures ..
Intrinsic                      :: max
!   .. Executable Statements ..
If (nstate==1) Then

!   First call to CONFUN. Set all Jacobian elements to zero.
!   Note that this will only work when 'Derivative Level = 3'
!   (the default; see Section 11.2).

  cjac(1:ncnln,1:n) = 0.0E0_nag_wp
End If

If (needc(1)>0) Then

  If (mode==0 .Or. mode==2) Then
    ccon(1) = x(1)**2 + x(2)**2 + x(3)**2 + x(4)**2
  End If

  If (mode==1 .Or. mode==2) Then
    cjac(1,1) = 2.0E0_nag_wp*x(1)
    cjac(1,2) = 2.0E0_nag_wp*x(2)
    cjac(1,3) = 2.0E0_nag_wp*x(3)
    cjac(1,4) = 2.0E0_nag_wp*x(4)
  End If

End If

If (needc(2)>0) Then

```

```

      If (mode==0 .Or. mode==2) Then
        ccon(2) = x(1)*x(2)*x(3)*x(4)
      End If

      If (mode==1 .Or. mode==2) Then
        cjac(2,1) = x(2)*x(3)*x(4)
        cjac(2,2) = x(1)*x(3)*x(4)
        cjac(2,3) = x(1)*x(2)*x(4)
        cjac(2,4) = x(1)*x(2)*x(3)
      End If

    End If

    Return

  End Subroutine confun
End Module e04wdfe_mod
Program e04wdfe

!      E04WDF Example Main Program

!      .. Use Statements ..
Use nag_library, Only: e04wcf, e04wdf, e04wgf, nag_wp
Use e04wdfe_mod, Only: confun, leniw, lenrw, nin, nout, objfun
!      .. Implicit None Statement ..
Implicit None
!      .. Local Scalars ..
Real (Kind=nag_wp)      :: objf
Integer                 :: i, ifail, lda, ldcj, ldh, majits, n, &
                        nclin, ncnln, sda, sdcjac
!      .. Local Arrays ..
Real (Kind=nag_wp), Allocatable :: a(:,,:), bl(:,), bu(:,), ccon(:,) &
                        cjac(:,,:), clamda(:,), grad(:,) &
                        h(:,,:), x(:,)
Real (Kind=nag_wp)      :: ruser(1), rw(lenrw)
Integer, Allocatable    :: istate(:,)
Integer                 :: iuser(1), iw(leniw)
!      .. Intrinsic Procedures ..
Intrinsic                :: max
!      .. Executable Statements ..
Write (nout,*) 'E04WDF Example Program Results'
Flush (nout)

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

Read (nin,*) n, nclin, ncnln
lda = max(1,nclin)

If (nclin>0) Then
  sda = n
Else
  sda = 1
End If

ldcj = max(1,ncnln)

If (ncnln>0) Then
  sdcjac = n
Else
  sdcjac = 1
End If

ldh = n
Allocate (istate(n+nclin+ncnln),a(lda,sda),bl(n+nclin+ncnln), &
  bu(n+nclin+ncnln),ccon(max(1,ncnln)),cjac(ldcj,sdcjac),clamda(n+nclin+ &
  ncnln),grad(n),h(ldh,n),x(n))

If (nclin>0) Then
  Read (nin,*)(a(i,1:sda),i=1,nclin)

```

```

End If

Read (nin,*) bl(1:(n+nclin+ncnln))
Read (nin,*) bu(1:(n+nclin+ncnln))
Read (nin,*) x(1:n)

! Call E04WCF to initialize E04WDF.

ifail = 0
Call e04wcf(iw,leniw,rw,lenrw,ifail)

! By default E04WDF does not print monitoring
! information. Set the print file unit or the summary
! file unit to get information.

ifail = 0
Call e04wgf('Print file',nout,iw,rw,ifail)

! Solve the problem.

ifail = 0
Call e04wdf(n,nclin,ncnln,lda,ldcj,ldh,a,bl,bu,confun,objfun,majits,      &
  istate,ccon,cjac,clamda,objf,grad,h,x,iw,leniw,rw,lenrw,iuser,ruser,  &
  ifail)

Write (nout,*)
Write (nout,99999) objf
Write (nout,99998) x(1:n)

99999 Format (1X,'Final objective value = ',F11.3)
99998 Format (1X,'Optimal X = ',7F9.2)
End Program e04wdfe

```

10.2 Program Data

E04WDF Example Program Data

```

4      1      2      3      4      5      6      7      : N, NCLIN and NCNLN
1.0    1.0    1.0    1.0      : Matrix A
1.0    1.0    1.0    1.0  -1.0E+25  -1.0E+25  25.0      : Lower bounds BL
5.0    5.0    5.0    5.0  20.0      40.0      1.0E+25      : Upper bounds BU
1.0    5.0    5.0    5.0    1.0      : Initial vector X

```

10.3 Program Results

E04WDF Example Program Results

Parameters
=====

Files

Solution file.....	0	Old basis file	0	(Print file).....	6
Insert file.....	0	New basis file	0	(Summary file).....	0
Punch file.....	0	Backup basis file.....	0		
Load file.....	0	Dump file.....	0		

Frequencies

Print frequency.....	100	Check frequency.....	60	Save new basis map.....	100
Summary frequency.....	100	Factorization frequency	50	Expand frequency.....	10000

QP subproblems

QPSolver Cholesky.....					
Scale tolerance.....	0.900	Minor feasibility tol..	1.00E-06	Iteration limit.....	10000
Scale option.....	0	Minor optimality tol..	1.00E-06	Minor print level.....	1
Crash tolerance.....	0.100	Pivot tolerance.....	2.04E-11	Partial price.....	1
Crash option.....	3	Elastic weight.....	1.00E+04	Prtl price section (A)	4
		New superbasics.....	99	Prtl price section (-I)	3

The SQP Method

Minimize.....		Cold start.....		Proximal Point method..	1
Nonlinear objectiv vars	4	Major optimality tol...	2.00E-06	Function precision....	1.72E-13
Unbounded step size....	1.00E+20	Superbasics limit.....	4	Difference interval....	4.15E-07
Unbounded objective....	1.00E+15	Reduced Hessian dim....	4	Central difference int.	5.57E-05

```

Major step limit..... 2.00E+00      Derivative linesearch..      Derivative level..... 3
Major iterations limit. 1000      Linesearch tolerance... 0.90000      Verify level..... 0
Minor iterations limit. 500      Penalty parameter..... 0.00E+00      Major Print Level..... 1

Hessian Approximation
-----
Full-Memory Hessian....      Hessian updates..... 99999999      Hessian frequency..... 99999999
                                          Hessian flush..... 99999999

Nonlinear constraints
-----
Nonlinear constraints.. 2      Major feasibility tol.. 1.00E-06      Violation limit..... 1.00E+06
Nonlinear Jacobian vars 4

Miscellaneous
-----
LU factor tolerance.... 1.10      LU singularity tol..... 2.04E-11      Timing level..... 0
LU update tolerance.... 1.10      LU swap tolerance..... 1.03E-04      Debug level..... 0
LU partial pivoting...      eps (machine precision) 1.11E-16      System information.... No

```

```

Matrix statistics
-----
                Total      Normal      Free      Fixed      Bounded
Rows            3          3          0          0          0
Columns         4          0          0          0          4

No. of matrix elements      12      Density      100.000
Biggest      1.0000E+00      (excluding fixed columns,
Smallest      0.0000E+00      free rows, and RHS)

```

```

No. of objective coefficients      0

Nonlinear constraints      2      Linear constraints      1
Nonlinear variables      4      Linear variables      0
Jacobian variables      4      Objective variables      4
Total constraints      3      Total variables      4

The user has defined      8      out of      8      constraint gradients.
The user has defined      4      out of      4      objective gradients.

```

Cheap test of user-supplied problem derivatives...

The constraint gradients seem to be OK.

--> The largest discrepancy was 1.84E-07 in constraint 6

The objective gradients seem to be OK.

```

Gradient projected in one direction 4.99993000077E+00
Difference approximation      4.99993303560E+00

```

Itns	Major	Minors	Step	nCon	Feasible	Optimal	MeritFunction	L+U	BSwap	nS	condHz	Penalty	
2	0	2		1	1.7E+00	2.8E+00	1.60000000E+01	7		2	1.0E+00		_ r
4	1	2	1.0E+00	2	1.3E-01	3.2E-01	1.7726188E+01	8		1	6.2E+00	8.3E-02	_n r
5	2	1	1.0E+00	3	3.7E-02	1.7E-01	1.7099571E+01	7		1	2.0E+00	8.3E-02	_s
6	3	1	1.0E+00	4	2.2E-02	1.1E-02	1.7014005E+01	7		1	1.8E+00	8.3E-02	_
7	4	1	1.0E+00	5	1.5E-04	6.0E-04	1.7014018E+01	7		1	1.8E+00	9.2E-02	_
8	5	1	1.0E+00	6	(3.3E-07)	2.3E-05	1.7014017E+01	7		1	1.9E+00	3.6E-01	_
9	6	1	1.0E+00	7	(4.2E-10)	(2.4E-08)	1.7014017E+01	7		1	1.9E+00	3.6E-01	_

```

E04WDM EXIT 0 -- finished successfully
E04WDM INFO 1 -- optimality conditions satisfied

```

```

Problem name      NLP
No. of iterations      9      Objective value      1.7014017287E+01
No. of major iterations      6      Linear objective      0.0000000000E+00
Penalty parameter      3.599E-01      Nonlinear objective      1.7014017287E+01
No. of calls to funobj      8      No. of calls to funcon      8
No. of superbasics      1      No. of basic nonlinears      2
No. of degenerate steps      0      Percentage      0.00
Max x      2 4.7E+00      Max pi      2 5.5E-01
Max Primal infeas      0 0.0E+00      Max Dual infeas      3 4.8E-08
Nonlinear constraint violn      2.7E-09

```

Variable	State	Value	Lower bound	Upper bound	Lagr multiplier	Slack
variable	1	LL	1.000000	5.000000	1.087871	.
variable	2	FR	4.743000	1.000000	.	0.2570
variable	3	FR	3.821150	1.000000	.	1.179

variable	4	FR	1.379408	1.000000	5.000000	.	0.3794
Linear constnt	State		Value	Lower bound	Upper bound	Lagr multiplier	Slack
lincon	1	FR	10.94356	None	20.00000	.	9.056
Nonlin constnt	State		Value	Lower bound	Upper bound	Lagr multiplier	Slack
nlncon	1	UL	40.00000	None	40.00000	-0.1614686	-0.2700E-08
nlncon	2	LL	25.00000	25.00000	None	0.5522937	-0.2215E-08
Final objective value =			17.014				
Optimal X =	1.00		4.74	3.82	1.38		

Note: the remainder of this document is intended for more advanced users. Section 11 contains a detailed description of the algorithm which may be needed in order to understand Sections 12 and 13. Section 12 describes the optional parameters which may be set by calls to E04WFF, E04WGF and/or E04WHF. Section 13 describes the quantities which can be requested to monitor the course of the computation.

11 Algorithmic Details

Here we summarise the main features of the SQP algorithm used in E04WDF and introduce some terminology used in the description of the subroutine and its arguments. The SQP algorithm is fully described in Gill *et al.* (2002).

11.1 Constraints and Slack Variables

The upper and lower bounds on the $n_L + n_N$ components of $\begin{pmatrix} A_L x \\ c(x) \end{pmatrix}$ are said to define the *general constraints* of the problem. E04WDF converts the general constraints to equalities by introducing a set of *slack variables* $s = (s_1, s_2, \dots, s_{n_L+n_N})^T$. For example, the linear constraint $5 \leq 2x_1 + 3x_2 \leq \infty$ is replaced by $2x_1 + 3x_2 - s_1 = 0$ together with the bounded slack $5 \leq s_1 \leq \infty$. The minimization problem (1) can therefore be written in the equivalent form

$$\underset{x,s}{\text{minimize}} F(x) \quad \text{subject to} \quad \begin{pmatrix} A_L x \\ c(x) \end{pmatrix} - s = 0, \quad l \leq \begin{pmatrix} x \\ s \end{pmatrix} \leq u. \quad (2)$$

The general constraints become the equalities $A_L x - s_L = 0$ and $c(x) - s_N = 0$, where s_L and s_N are the *linear* and *nonlinear* slacks.

11.2 Major Iterations

The basic structure of the SQP algorithm involves *major* and *minor* iterations. The major iterations generate a sequence of iterates $\{x_k\}$ that satisfy the linear constraints and converge to a point that satisfies the nonlinear constraints and the first-order conditions for optimality. At each iterate x_k a QP subproblem is used to generate a search direction towards the next iterate x_{k+1} . The constraints of the subproblem are formed from the linear constraints $A_L x - s_L = 0$ and the linearized constraint

$$c(x_k) + c'(x_k)(x - x_k) - s_N = 0, \quad (3)$$

where $c'(x_k)$ denotes the *Jacobian matrix*, whose elements are the first derivatives of $c(x)$ evaluated at x_k . The QP constraints therefore comprise the $n_L + n_N$ linear constraints

$$\begin{aligned} A_L x - s_L &= 0, \\ c'(x_k)x - s_N &= -c(x_k) + c'(x_k)x_k, \end{aligned} \quad (4)$$

where x and s are bounded above and below by u and l as before. If the $(n_L + n_N) \times n$ matrix A and $(n_L + n_N)$ -vector b are defined as

$$A = \begin{pmatrix} A_L \\ c'(x_k) \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ -c(x_k) + c'(x_k)x_k \end{pmatrix}, \quad (5)$$

then the QP subproblem can be written as

$$\underset{x,s}{\text{minimize}} q(x, x_k) = g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k(x - x_k) \quad \text{subject to } Ax - s = b, \quad l \leq \begin{pmatrix} x \\ s \end{pmatrix} \leq u, \quad (6)$$

where $q(x, x_k)$ is a quadratic approximation to a modified Lagrangian function (see Gill *et al.* (2002)). The matrix H_k is a quasi-Newton approximation to the Hessian of the Lagrangian. A BGFS update is applied after each major iteration. If some of the variables enter the Lagrangian linearly the Hessian will have some zero rows and columns. If the nonlinear variables appear first, then only the leading n_N rows and columns of the Hessian need to be approximated.

11.3 Minor Iterations

Solving the QP subproblem is itself an iterative procedure. Here, the iterations of the QP solver E04NQF form the *minor* iterations of the SQP method. E04NQF uses a reduced-Hessian active-set method implemented as a reduced-gradient method. At each minor iteration, the constraints $Ax - s = b$ are partitioned into the form

$$Bx_B + Sx_S + Nx_N = b, \quad (7)$$

where the *basis matrix* B is square and nonsingular, and the matrices S and N are the remaining columns of $(A \ -I)$. The vectors x_B , x_S and x_N are the associated *basic*, *superbasic* and *nonbasic* variables respectively; they are a permutation of the elements of x and s . At a QP subproblem, the basic and superbasic variables will lie somewhere between their bounds, while the nonbasic variables will normally be equal to one of their bounds. At each iteration, x_S is regarded as a set of independent variables that are free to move in any desired direction, namely one that will improve the value of the QP objective (or the sum of infeasibilities). The basic variables are then adjusted in order to ensure that (x, s) continues to satisfy $Ax - s = b$. The number of superbasic variables (n_S , say) therefore indicates the number of degrees of freedom remaining after the constraints have been satisfied. In broad terms, n_S is a measure of *how nonlinear* the problem is. In particular, n_S will always be zero for LP problems.

If it appears that no improvement can be made with the current definition of B , S and N , a nonbasic variable is selected to be added to S , and the process is repeated with the value of n_S increased by one. At all stages, if a basic or superbasic variable encounters one of its bounds, the variable is made nonbasic and the value of n_S is decreased by one.

Associated with each of the $n_L + n_N$ equality constraints $Ax - s = b$ are the *dual variables* π . Similarly, each variable in (x, s) has an associated *reduced gradient* d_j . The reduced gradients for the variables x are the quantities $g - A^T\pi$, where g is the gradient of the QP objective, and the reduced gradients for the slacks are the dual variables π . The QP subproblem is optimal if $d_j \geq 0$ for all nonbasic variables at their lower bounds, $d_j \leq 0$ for all nonbasic variables at their upper bounds, and $d_j = 0$ for other variables, including superbasics. In practice, an *approximate* QP solution $(\hat{x}_k, \hat{s}_k, \hat{\pi}_k)$ is found by relaxing these conditions.

11.4 The Merit Function

After a QP subproblem has been solved, new estimates of the solution are computed using a linesearch on the augmented Lagrangian merit function

$$\mathcal{M}(x, s, \pi) = F(x) - \pi^T(c(x) - s_N) + \frac{1}{2}(c(x) - s_N)^T D(c(x) - s_N), \quad (8)$$

where D is a diagonal matrix of penalty parameters ($D_{ii} \geq 0$), and π now refers to dual variables for the nonlinear constraints in (1). If (x_k, s_k, π_k) denotes the current solution estimate and $(\hat{x}_k, \hat{s}_k, \hat{\pi}_k)$ denotes the QP solution, the linesearch determines a step α_k ($0 < \alpha_k \leq 1$) such that the new point

$$\begin{pmatrix} x_{k+1} \\ s_{k+1} \\ \pi_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ s_k \\ \pi_k \end{pmatrix} + \alpha_k \begin{pmatrix} \hat{x}_k - x_k \\ \hat{s}_k - s_k \\ \hat{\pi}_k - \pi_k \end{pmatrix} \quad (9)$$

gives a *sufficient decrease* in the merit function \mathcal{M} . When necessary, the penalties in D are increased by the minimum-norm perturbation that ensures descent for \mathcal{M} (see Gill *et al.* (1992)). The value of s_N is

adjusted to minimize the merit function as a function of s before the solution of the QP subproblem (see Gill *et al.* (1986) and Eldersveld (1991)).

11.5 Treatment of Constraint Infeasibilities

E04WDF makes explicit allowance for infeasible constraints. First, infeasible *linear* constraints are detected by solving the linear program

$$\underset{x,v,w}{\text{minimize}} e^T(v+w) \quad \text{subject to } l \leq \begin{pmatrix} x \\ A_L x - v + w \end{pmatrix} \leq u, \quad v \geq 0, \quad w \geq 0, \quad (10)$$

where e is a vector of ones, and the nonlinear constraint bounds are temporarily excluded from l and u . This is equivalent to minimizing the sum of the general linear constraint violations subject to the bounds on x . (The sum is the ℓ_1 -norm of the linear constraint violations. In the linear programming literature, the approach is called *elastic programming*.)

The linear constraints are infeasible if the optimal solution of (10) has $v \neq 0$ or $w \neq 0$. E04WDF then terminates without computing the nonlinear functions.

Otherwise, all subsequent iterates satisfy the linear constraints. (Such a strategy allows linear constraints to be used to define a region in which the functions can be safely evaluated.) E04WDF proceeds to solve nonlinear problems as given, using search directions obtained from the sequence of QP subproblems (see (6)).

If a QP subproblem proves to be infeasible or unbounded (or if the dual variables π for the nonlinear constraints become large), E04WDF enters ‘elastic’ mode and thereafter solves the problem

$$\underset{x,v,w}{\text{minimize}} F(x) + \gamma e^T(v+w) \quad \text{subject to } l \leq \begin{pmatrix} x \\ A_L x \\ c(x) - v + w \end{pmatrix} \leq u, \quad v \geq 0, \quad w \geq 0, \quad (11)$$

where γ is a non-negative argument (the *elastic weight*), and $F(x) + \gamma e^T(v+w)$ is called a *composite objective* (the ℓ_1 penalty function for the nonlinear constraints).

The value of γ may increase automatically by multiples of 10 if the optimal v and w continue to be nonzero. If γ is sufficiently large, this is equivalent to minimizing the sum of the nonlinear constraint violations subject to the linear constraints and bounds.

The initial value of γ is controlled by the optional parameter **Elastic Weight**.

12 Optional Parameters

Several optional parameters in E04WDF define choices in the problem specification or the algorithm logic. In order to reduce the number of formal arguments of E04WDF these optional parameters have associated *default values* that are appropriate for most problems. Therefore, you need only specify those optional parameters whose values are to be different from their default values.

The remainder of this section can be skipped if you wish to use the default values for all optional parameters.

The following is a list of the optional parameters available. A full description of each optional parameter is provided in Section 12.1.

Backup Basis File

Central Difference Interval

Check Frequency

Cold Start

Crash Option

Crash Tolerance

Defaults

Derivative Level

Derivative Linesearch

Difference Interval
Dump File
Elastic Weight
Expand Frequency
Factorization Frequency
Feasibility Tolerance
Feasible Point
Function Precision
Hessian Frequency
Hessian Full Memory
Hessian Limited Memory
Hessian Updates
Infinite Bound Size
Insert File
Iterations Limit
Linesearch Tolerance
List
Load File
LU Complete Pivoting
LU Density Tolerance
LU Factor Tolerance
LU Partial Pivoting
LU Rook Pivoting
LU Singularity Tolerance
LU Update Tolerance
Major Feasibility Tolerance
Major Iterations Limit
Major Optimality Tolerance
Major Print Level
Major Step Limit
Maximize
Minimize
Minor Feasibility Tolerance
Minor Iterations Limit
Minor Print Level
New Basis File
New Superbasics Limit
Nolist
Nonderivative Linesearch
Old Basis File
Partial Price
Pivot Tolerance
Print File
Print Frequency
Proximal Point Method
Punch File
QPSolver CG

QPSolver Cholesky
QPSolver QN
Reduced Hessian Dimension
Save Frequency
Scale Option
Scale Print
Scale Tolerance
Solution File
Start Constraint Check At Variable
Start Objective Check At Variable
Stop Constraint Check At Variable
Stop Objective Check At Variable
Summary File
Summary Frequency
Superbasics Limit
Suppress Parameters
System Information No
System Information Yes
Timing Level
Unbounded Objective
Unbounded Step Size
Verify Level
Violation Limit
Warm Start

Optional parameters may be specified by calling one, or more, of the routines E04WEF, E04WFF and E04WGF before a call to E04WDF.

E04WEF reads options from an external options file, with *Begin* and *End* as the first and last lines respectively and each intermediate line defining a single optional parameter. For example,

```

Begin
    Print Level = 5
End

```

The call

```
CALL E04WEF (ISPECS, IW, RW, IFAIL)
```

can then be used to read the file on unit ISPECS. IFAIL will be zero on successful exit. E04WEF should be consulted for a full description of this method of supplying optional parameters.

E04WFF, E04WGF or E04WHF can be called to supply options directly, one call being necessary for each optional parameter. E04WFF, E04WGF or E04WHF should be consulted for a full description of this method of supplying optional parameters.

All optional parameters not specified by you are set to their default values. Optional parameters specified by you are unaltered by E04WDF (unless they define invalid values) and so remain in effect for subsequent calls to E04WDF, unless altered by you.

12.1 Description of the Optional Parameters

For each option, we give a summary line, a description of the optional parameter and details of constraints.

The summary line contains:

the keywords, where the minimum abbreviation of each keyword is underlined (if no characters of an optional qualifier are underlined, the qualifier may be omitted);

a parameter value, where the letters a , i and r denote options that take character, integer and real values respectively;

the default value, where the symbol ϵ is a generic notation for **machine precision** (see X02AJF), and ϵ_r denotes the relative precision of the objective function **Function Precision**, and *bigbnd* signifies the value of **Infinite Bound Size**.

Keywords and character values are case and white space insensitive.

Central Difference Interval

r

Default = $\epsilon_r^{\frac{1}{3}}$

When **Derivative Level** < 3, the central-difference interval r is used near an optimal solution to obtain more accurate (but more expensive) estimates of gradients. Twice as many function evaluations are required compared to forward differencing. The interval used for the j th variable is $h_j = r(1 + |x_j|)$. The resulting derivative estimates should be accurate to $O(r^2)$, unless the functions are badly scaled.

If you supply a value for this optional parameter, a small value between 0.0 and 1.0 is appropriate.

Check Frequency

i

Default = 60

Every i th minor iteration after the most recent basis factorization, a numerical test is made to see if the current solution x satisfies the general linear constraints (the linear constraints and the linearized nonlinear constraints, if any). The constraints are of the form $Ax - s = b$, where s is the set of slack variables. To perform the numerical test, the residual vector $r = b - Ax + s$ is computed. If the largest component of r is judged to be too large, the current basis is refactorized and the basic variables are recomputed to satisfy the general constraints more accurately. If $i \leq 0$, the value of $i = 99999999$ is used and effectively no checks are made.

Check Frequency = 1 is useful for debugging purposes, but otherwise this option should not be needed.

Cold Start

Default

Warm Start

This option controls the specification of the initial working set in the procedure for finding a feasible point for the linear constraints and bounds and in the first QP subproblem thereafter. With a **Cold Start**, the first working set is chosen by E04WDF based on the values of the variables and constraints at the initial point. Broadly speaking, the initial working set will include equality constraints and bounds or inequality constraints that violate or ‘nearly’ satisfy their bounds (to within **Crash Tolerance**).

With a **Warm Start**, you must set the ISTATE array and define CLAMDA and H as discussed in Section 5. ISTATE values associated with bounds and linear constraints determine the initial working set of the procedure to find a feasible point with respect to the bounds and linear constraints. ISTATE values associated with nonlinear constraints determine the initial working set of the first QP subproblem after such a feasible point has been found. E04WDF will override your specification of ISTATE if necessary, so that a poor choice of the working set will not cause a fatal error. For instance, any elements of ISTATE which are set to -2, -1 or 4 will be reset to zero, as will any elements which are set to 3 when the corresponding elements of BL and BU are not equal. A warm start will be advantageous if a good estimate of the initial working set is available – for example, when E04WDF is called repeatedly to solve related problems.

Crash Option

i

Default = 3

Crash Tolerance

r

Default = 0.1

If a **Cold Start** is specified, an internal Crash procedure is used to select an initial basis from certain rows and columns of the constraint matrix $(A \ -I)$. The optional parameter **Crash Option** i determines which rows and columns of A are eligible initially, and how many times the Crash procedure is called. Columns of $-I$ are used to pad the basis where necessary.

<i>i</i>	Meaning
0	The initial basis contains only slack variables: $B = I$.
1	The Crash procedure is called once, looking for a triangular basis in all rows and columns of A .
2	The Crash procedure is called twice (if there are nonlinear constraints). The first call looks for a triangular basis in linear rows, and the iteration proceeds with simplex iterations until the linear constraints are satisfied. The Jacobian is then evaluated for the first major iteration and the Crash procedure is called again to find a triangular basis in the nonlinear rows (retaining the current basis for linear rows).
3	The Crash procedure is called up to three times (if there are nonlinear constraints). The first two calls treat <i>linear equalities</i> and <i>linear inequalities</i> separately. As before, the last call treats nonlinear rows before the first major iteration.

If $i \geq 1$, certain slacks on inequality rows are selected for the basis first. (If $i \geq 2$, numerical values are used to exclude slacks that are close to a bound). The Crash procedure then makes several passes through the columns of A , searching for a basis matrix that is essentially triangular. A column is assigned to 'pivot' on a particular row if the column contains a suitably large element in a row that has not yet been assigned. (The pivot elements ultimately form the diagonals of the triangular basis.) For remaining unassigned rows, slack variables are inserted to complete the basis.

The **Crash Tolerance** r allows the starting Crash procedure to ignore certain 'small' nonzeros in each column of A . If a_{\max} is the largest element in column j , other nonzeros of a_{ij} in the columns are ignored if $|a_{ij}| \leq a_{\max} \times r$. (To be meaningful, r must be in the range $0 \leq r < 1$.)

When $r > 0.0$, the basis obtained by the Crash procedure may not be strictly triangular, but it is likely to be nonsingular and almost triangular. The intention is to obtain a starting basis containing more columns of A and fewer (arbitrary) slacks. A feasible solution may be reached sooner on some problems.

For example, suppose the first m columns of A form the matrix shown under **LU Factor Tolerance**; i. e., a tridiagonal matrix with entries $-1, 4, -1$. To help the Crash procedure choose all m columns for the initial basis, we would specify a **Crash Tolerance** of r for some value of $r > 0.5$.

Defaults

This special keyword may be used to reset all optional parameters to their default values.

<u>Derivative Level</u>	<i>i</i>	Default = 3
--------------------------------	----------	-------------

Optional parameter **Derivative Level** specifies which nonlinear function gradients are known analytically and will be supplied to E04WDF by user-supplied subroutines OBJFUN and CONFUN.

<i>i</i>	Meaning
3	All objective and constraint gradients are known.
2	All constraint gradients are known, but some or all components of the objective gradient are unknown.
1	The objective gradient is known, but some or all of the constraint gradients are unknown.
0	Some components of the objective gradient are unknown and some of the constraint gradients are unknown.

The value $i = 3$ should be used whenever possible. It is the most reliable and will usually be the most efficient.

If $i = 0$ or 2 , E04WDF will *estimate* the missing components of the objective gradient, using finite differences. This may simplify the coding of OBJFUN. However, it could increase the total run-time substantially (since a special call to OBJFUN is required for each missing element), and there is less assurance that an acceptable solution will be located. If the nonlinear variables are not well scaled, it may be necessary to specify a non-default optional parameter **Difference Interval**.

If $i = 0$ or 1 , E04WDF will estimate missing elements of the Jacobian. For each column of the Jacobian, one call to CONFUN is needed to estimate all missing elements in that column, if any.

At times, central differences are used rather than forward differences. (This is not under your control.)

Derivative Linesearch

Default

Nonderivative Linesearch

At each major iteration a linesearch is used to improve the merit function. Optional parameter **Derivative Linesearch** uses safeguarded cubic interpolation and requires both function and gradient values to compute estimates of the step α_k . If some analytic derivatives are not provided, or optional parameter **Nonderivative Linesearch** is specified, E04WDF employs a linesearch based upon safeguarded quadratic interpolation, which does not require gradient evaluations.

A nonderivative linesearch can be slightly less robust on difficult problems, and it is recommended that the default be used if the functions and derivatives can be computed at approximately the same cost. If the gradients are very expensive relative to the functions, a nonderivative linesearch may give a significant decrease in computation time.

If **Nonderivative Linesearch** is selected, E04WDF signals the evaluation of the linesearch by calling OBJFUN with $\text{MODE} = 0$. If the potential saving provided by a nonderivative linesearch is to be realised, it is essential that OBJFUN be coded so that derivatives are not computed when $\text{MODE} = 0$.

Difference Interval

 r Default = $\sqrt{\epsilon_r}$

This alters the interval r used to estimate gradients by forward differences. It does so in the following circumstances:

- in the interval ('cheap') phase of verifying the problem derivatives;
- for verifying the problem derivatives;
- for estimating missing derivatives.

In all cases, a derivative with respect to x_j is estimated by perturbing that component of x to the value $x_j + r(1 + |x_j|)$, and then evaluating $F(x)$ or $c(x)$ at the perturbed point. The resulting gradient estimates should be accurate to $O(r)$ unless the functions are badly scaled. Judicious alteration of r may sometimes lead to greater accuracy.

If you supply a value for this optional parameter, a small value between 0.0 and 1.0 is appropriate.

Dump File

 i_1

Default = 0

Load File

 i_2

Default = 0

Optional parameters **Dump File** and **Load File** are similar to optional parameters **Punch File** and **Insert File**, but they record solution information in a manner that is more direct and more easily modified. A full description of information recorded in optional parameters **Dump File** and **Load File** is given in Gill *et al.* (2005a).

If $i_1 > 0$, the last solution obtained will be output to the file with unit number i_1 .

If $i_2 > 0$, the **Load File**, containing basis information, will be read. The file will usually have been output previously as a **Dump File**. The file will not be accessed if optional parameters **Old Basis File** or **Insert File** are specified.

Elastic Weight

 r Default = 10^4

This keyword determines the initial weight γ associated with the problem (11) (see Section 11.5).

At major iteration k , if elastic mode has not yet started, a scale factor $\sigma_k = 1 + \|g(x_k)\|_\infty$ is defined from the current objective gradient. Elastic mode is then started if the QP subproblem is infeasible, or the QP dual variables are larger in magnitude than $\sigma_k r$. The QP is resolved in elastic mode with $\gamma = \sigma_k r$.

Thereafter, major iterations continue in elastic mode until they converge to a point that is optimal for (11) (see Section 11.5). If the point is feasible for equation (1) ($v = w = 0$), it is declared locally

optimal. Otherwise, γ is increased by a factor of 10 and major iterations continue. If γ has already reached a maximum allowable value, equation (1) is declared locally infeasible.

Expand Frequency i Default = 10000

This option is part of the anti-cycling procedure designed to make progress even on highly degenerate problems.

For linear models, the strategy is to force a positive step at every iteration, at the expense of violating the bounds on the variables by a small amount. Suppose that the optional parameter **Minor Feasibility Tolerance** is δ . Over a period of i iterations, the tolerance actually used by E04WDF increases from 0.5δ to δ (in steps of $0.5\delta/i$).

For nonlinear models, the same procedure is used for iterations in which there is only one superbasic variable. (Cycling can occur only when the current solution is at a vertex of the feasible region.) Thus, zero steps are allowed if there is more than one superbasic variable, but otherwise positive steps are enforced.

Increasing i helps reduce the number of slightly infeasible nonbasic variables (most of which are eliminated during a resetting procedure). However, it also diminishes the freedom to choose a large pivot element (see optional parameter **Pivot Tolerance**).

Factorization Frequency i Default = 50

At most i basis changes will occur between factorizations of the basis matrix.

With linear programs, the basis factors are usually updated every iteration. The default i is reasonable for typical problems. Higher values up to $i = 100$ (say) may be more efficient on well-scaled problems.

When the objective function is nonlinear, fewer basis updates will occur as an optimum is approached. The number of iterations between basis factorizations will therefore increase. During these iterations a test is made regularly (according to the optional parameter **Check Frequency**) to ensure that the general constraints are satisfied. If necessary the basis will be refactorized before the limit of i updates is reached.

Function Precision r Default = $\epsilon^{0.8}$

The *relative function precision* ϵ_r is intended to be a measure of the relative accuracy with which the functions can be computed. For example, if $F(x)$ is computed as 1000.56789 for some relevant x and if the first 6 significant digits are known to be correct, the appropriate value for ϵ_r would be $1.0\text{E}-6$.

(Ideally the functions $F(x)$ or $c_i(x)$ should have magnitude of order 1. If all functions are substantially less than 1 in magnitude, ϵ_r should be the *absolute* precision. For example, if $F(x) = 1.23456789\text{E}-4$ at some point and if the first 6 significant digits are known to be correct, the appropriate value for ϵ_r would be $1.0\text{E}-10$.)

The default value of ϵ_r is appropriate for simple analytic functions.

In some cases the function values will be the result of extensive computation, possibly involving a costly iterative procedure that can provide few digits of precision. Specifying an appropriate **Function Precision** may lead to savings, by allowing the linesearch procedure to terminate when the difference between function values along the search direction becomes as small as the absolute error in the values.

Hessian Full Memory Default if $n \leq 75$
Hessian Limited Memory Default if $n > 75$

These options select the method for storing and updating the approximate Hessian. (E04WDF uses a quasi-Newton approximation to the Hessian of the Lagrangian. A BFGS update is applied after each major iteration.)

If **Hessian Full Memory** is specified, the approximate Hessian is treated as a dense matrix and the BFGS updates are applied explicitly. This option is most efficient when the number of variables n is not too large (say, less than 75). In this case, the storage requirement is fixed and one can expect 1-step Q-superlinear convergence to the solution.

Hessian Limited Memory should be used on problems where n is very large. In this case a limited-memory procedure is used to update a diagonal Hessian approximation H_r a limited number of times. (Updates are accumulated as a list of vector pairs. They are discarded at regular intervals after H_r has been reset to their diagonal.)

Hessian Frequency i Default = 99999999

If optional parameter **Hessian Full Memory** is in effect and i BFGS updates have already been carried out, the Hessian approximation is reset to the identity matrix. (For certain problems, occasional resets may improve convergence, but in general they should not be necessary.)

Hessian Full Memory and **Hessian Frequency** = 10 have a similar effect to **Hessian Limited Memory** and **Hessian Updates** = 10 (except that the latter retains the current diagonal during resets).

Hessian Updates i Default = **Hessian Frequency** if **Hessian Full Memory**, 10 otherwise

If optional parameter **Hessian Limited Memory** is in effect and i BFGS updates have already been carried out, all but the diagonal elements of the accumulated updates are discarded and the updating process starts again.

Broadly speaking, the more updates stored, the better the quality of the approximate Hessian. However, the more vectors stored, the greater the cost of each QP iteration. The default value is likely to give a robust algorithm without significant expense, but faster convergence can sometimes be obtained with significantly fewer updates (e.g., $i = 5$).

Infinite Bound Size r Default = 10^{20}

If $r > 0$, r defines the ‘infinite’ bound *bigbnd* in the definition of the problem constraints. Any upper bound greater than or equal to *bigbnd* will be regarded as $+\infty$ (and similarly any lower bound less than or equal to $-bigbnd$ will be regarded as $-\infty$). If $r < 0$, the default value is used.

Iterations Limit i Default = $\max(10000, 10 \max(n, n_L + n_N))$

The value of i specifies the maximum number of minor iterations allowed (i.e., iterations of the simplex method or the QP algorithm), summed over all major iterations. (See also the description of the optional parameter **Minor Iterations Limit**.)

Linesearch Tolerance r Default = 0.9

This tolerance, r , controls the accuracy with which a step length will be located along the direction of search each iteration. At the start of each linesearch a target directional derivative for the merit function is identified. This parameter determines the accuracy to which this target value is approximated, and it must be a value in the range $0.0 \leq r \leq 1.0$.

The default value $r = 0.9$ requests just moderate accuracy in the linesearch.

If the nonlinear functions are cheap to evaluate, a more accurate search may be appropriate; try $r = 0.1$, 0.01 or 0.001.

If the nonlinear functions are expensive to evaluate, a less accurate search may be appropriate. *If all gradients are known*, try $r = 0.99$. (The number of major iterations might increase, but the total number of function evaluations may decrease enough to compensate.)

If not all gradients are known, a moderately accurate search remains appropriate. Each search will require only 1–5 function values (typically), but many function calls will then be needed to estimate missing gradients for the next iteration.

Nolist Default
List

For E04WDF, normally each optional parameter specification is printed as it is supplied. Optional parameter **Nolist** may be used to suppress the printing and optional parameter **List** may be used to turn on printing.

<u>LU Density Tolerance</u>	r_1	Default = 0.6
<u>LU Singularity Tolerance</u>	r_2	Default = ϵ^2

The density tolerance, r_1 , is used during *LU* factorization of the basis matrix B . Columns of L and rows of U are formed one at a time, and the remaining rows and columns of the basis are altered appropriately. At any stage, if the density of the remaining matrix exceeds r_1 , the Markowitz strategy for choosing pivots is terminated, and the remaining matrix is factored by a dense *LU* procedure. Raising the density tolerance towards 1.0 may give slightly sparser *LU* factors, with a slight increase in factorization time.

The singularity tolerance, r_2 , helps guard against ill-conditioned basis matrices. After B is refactorized, the diagonal elements of U are tested as follows: if $|u_{jj}| \leq r_2$ or $|u_{jj}| < r_2 \max_i |u_{ij}|$, the j th column of the basis is replaced by the corresponding slack variable. (This is most likely to occur after a restart.)

<u>LU Factor Tolerance</u>	r_1	Default = 1.10
<u>LU Update Tolerance</u>	r_2	Default = 1.10

The values of r_1 and r_2 affect the stability of the basis factorization $B = LU$, during refactorization and updates respectively. The lower triangular matrix L is a product of matrices of the form

$$\begin{pmatrix} 1 & \\ \mu & 1 \end{pmatrix}$$

where the multipliers μ will satisfy $|\mu| \leq r_i$. The default values of r_1 and r_2 usually strike a good compromise between stability and sparsity. They must satisfy $r_1, r_2 \geq 1.0$.

For large and relatively dense problems, $r_1 = 10.0$ or 5.0 (say) may give a useful improvement in stability without impairing sparsity to a serious degree.

For certain very regular structures (e.g., band matrices) it may be necessary to reduce r_1 and/or r_2 in order to achieve stability. For example, if the columns of A include a sub-matrix of the form

$$\begin{pmatrix} 4 & -1 & & & & \\ -1 & 4 & -1 & & & \\ & -1 & 4 & -1 & & \\ & & \cdots & \cdots & \cdots & \\ & & & -1 & 4 & -1 \\ & & & & -1 & 4 \end{pmatrix},$$

one should set both r_1 and r_2 to values in the range $1.0 \leq r_i < 4.0$.

<u>LU Partial Pivoting</u>	Default
<u>LU Complete Pivoting</u>	
<u>LU Rook Pivoting</u>	

The *LU* factorization implements a Markowitz-type search for pivots that locally minimize the fill-in subject to a threshold pivoting stability criterion. The default option is to use threshold partial pivoting. The optional parameters **LU Rook Pivoting** and **LU Complete Pivoting** are more expensive than partial pivoting but are more stable and better at revealing rank, as long as **LU Factor Tolerance** is not too large (say < 2.0). When numerical difficulties are encountered, E04WDF automatically reduces the *LU* tolerance towards 1.0 and switches (if necessary) to rook or complete pivoting, before reverting to the default or specified options at the next refactorization (with **System Information Yes**, relevant messages are output to the **Print File**).

<u>Major Feasibility Tolerance</u>	r	Default = $\max(10^{-6}, \sqrt{\epsilon})$
---	-----	--

This tolerance, r , specifies how accurately the nonlinear constraints should be satisfied. The default value is appropriate when the linear and nonlinear constraints contain data to about that accuracy.

Let v_{\max} be the maximum nonlinear constraint violation, normalized by the size of the solution, which is required to satisfy

$$v_{\max}/ = \max_i v_i / \|x\| \leq r, \quad (12)$$

where v_i is the violation of the i th nonlinear constraint ($i = 1 : n_L$).

In the major iteration log (see Section 13.2, v_{\max} appears as the quantity labelled ‘Feasible’. If some of the problem functions are known to be of low accuracy, a larger **Major Feasibility Tolerance** may be appropriate.

Major Optimality Tolerance r Default = $2 \max(10^{-6}, \sqrt{\epsilon})$

This tolerance, r , specifies the final accuracy of the dual variables. On successful termination, E04WDF will have computed a solution (x, s, π) such that

$$c_{\max} = \max_j c_j / \|\pi\| \leq r, \quad (13)$$

where c_j is an estimate of the complementarity slackness for variable j where $j = 1 : n + n_L + n_N$. The values c_i are computed from the final QP solution using the reduced gradients $d_j = g_j - \pi^T a_j$ (where g_j is the j th component of the objective gradient, a_j is the associated column of the constraint matrix $(A \ -I)$, and π is the set of QP dual variables):

$$c_j = \begin{cases} d_j \min(x_j - l_j, 1) & \text{if } d_j \geq 0; \\ -d_j \min(u_j - x_j, 1) & \text{if } d_j < 0. \end{cases} \quad (14)$$

In the **Print File**, c_{\max} appears as the quantity labelled ‘Optimal’.

Major Iterations Limit i Default = $\max(1000, 3 \max(n, n_L + n_N))$

This is the maximum number of major iterations allowed. It is intended to guard against an excessive number of linearizations of the constraints. If $i = 0$, optimality and feasibility are checked.

Major Print Level i Default = 000001

This controls the amount of output to the optional parameters **Print File** and **Summary File** at each major iteration. **Major Print Level** = 0 suppresses most output, except for error messages. **Major Print Level** = 1 gives normal output for linear and nonlinear problems, and **Major Print Level** = 11 gives additional details of the Jacobian factorization that commences each major iteration.

In general, the value being specified may be thought of as a binary number of the form

Major Print Level *JFDXbs*

where each letter stands for a digit that is either 0 or 1 as follows:

- s* a single line that gives a summary of each major iteration. (This entry in *JFDXbs* is not strictly binary since the summary line is printed whenever $JFDXbs \geq 1$);
- b* basis statistics, i.e., information relating to the basis matrix whenever it is refactorized. (This output is always provided if $JFDXbs \geq 10$);
- X* x_k , the nonlinear variables involved in the objective function or the constraints. These appear under the heading ‘Jacobian variables’;
- D* π_k , the dual variables for the nonlinear constraints. These appear under the heading ‘Multiplier estimates’;
- F* $c(x_k)$, the values of the nonlinear constraint functions;
- J* $J(x_k)$, the Jacobian matrix. This appears under the heading ‘ x and Jacobian’.

To obtain output of any items *JFDXbs*, set the corresponding digit to 1, otherwise to 0.

If $J = 1$, the Jacobian matrix will be output column-wise at the start of each major iteration. Column j will be preceded by the value of the corresponding variable x_j and a key to indicate whether the

variable is basic, superbasic or nonbasic. (Hence if $J = 1$, there is no reason to specify $X = 1$ unless the objective contains more nonlinear variables than the Jacobian.) A typical line of output is

```
3 1.250000E+01 BS 1 1.00000E+00 4 2.00000E+00
```

which would mean that x_3 is basic at value 12.5, and the third column of the Jacobian has elements of 1.0 and 2.0 in rows 1 and 4.

Major Step Limit

r

Default = 2.0

This parameter limits the change in x during a linesearch. It applies to all nonlinear problems, once a ‘feasible solution’ or ‘feasible subproblem’ has been found. A linesearch determines a step α over the range $0 < \alpha \leq \beta$, where β is 1 if there are nonlinear constraints, or is the step to the nearest upper or lower bound on x if all the constraints are linear. Normally, the first step length tried is $\alpha_1 = \min(1, \beta)$.

1. In some cases, such as $f(x) = ae^{bx}$ or $f(x) = ax^b$, even a moderate change in the components of x can lead to floating-point overflow. The parameter r is therefore used to define a limit $\bar{\beta} = r(1 + \|x\|)/\|p\|$ (where p is the search direction), and the first evaluation of $f(x)$ is at the potentially smaller step length $\alpha_1 = \min(1, \bar{\beta}, \beta)$.
2. Wherever possible, upper and lower bounds on x should be used to prevent evaluation of nonlinear functions at meaningless points. The optional parameter **Major Step Limit** provides an additional safeguard. The default value $r = 2.0$ should not affect progress on well behaved problems, but setting $r = 0.1$ or 0.01 may be helpful when rapidly varying functions are present. A ‘good’ starting point may be required. An important application is to the class of nonlinear least squares problems.
3. In cases where several local optima exist, specifying a small value for r may help locate an optimum near the starting point.

Minimize

Default

Maximize

Feasible Point

The keywords **Minimize** and **Maximize** specify the required direction of optimization. It applies to both linear and nonlinear terms in the objective.

The keyword **Feasible Point** means ‘Ignore the objective function, while finding a feasible point for the linear and nonlinear constraints’. It can be used to check that the nonlinear constraints are feasible without altering the call to E04WDF.

Minor Feasibility Tolerance

Feasibility Tolerance

r

Default = $\max\{10^{-6}, \sqrt{\epsilon}\}$

E04WDF tries to ensure that all variables eventually satisfy their upper and lower bounds to within this tolerance, r . This includes slack variables. Hence, general linear constraints should also be satisfied to within r .

Feasibility with respect to nonlinear constraints is judged by the optional parameter **Major Feasibility Tolerance** (not by r).

If the bounds and linear constraints cannot be satisfied to within r , the problem is declared *infeasible*. If the corresponding sum of infeasibilities is quite small, it may be appropriate to raise r by a factor of 10 or 100. Otherwise, some error in the data should be suspected.

Nonlinear functions will be evaluated only at points that satisfy the bounds and linear constraints. If there are regions where a function is undefined, every attempt should be made to eliminate these regions from the problem.

For example, if $f(x) = \sqrt{x_1} + \log(x_2)$, it is essential to place lower bounds on both variables. If $r = 1.0E-6$, the bounds $x_1 \geq 10^{-5}$ and $x_2 \geq 10^{-4}$ might be appropriate. (The log singularity is more serious. In general, keep x as far away from singularities as possible.)

If **Scale Option** ≥ 1 , feasibility is defined in terms of the *scaled* problem (since it is then more likely to be meaningful).

In reality, E04WDF uses r as a feasibility tolerance for satisfying the bounds on x and s in each QP subproblem. If the sum of infeasibilities cannot be reduced to zero, the QP subproblem is declared infeasible. E04WDF is then in *elastic mode* thereafter (with only the linearized nonlinear constraints defined to be elastic). See the description of the optional parameter **Elastic Weight**.

Minor Iterations Limit i Default = 500

If the number of minor iterations for the optimality phase of the QP subproblem exceeds i , then all nonbasic QP variables that have not yet moved are frozen at their current values and the reduced QP is solved to optimality.

Note that more than i minor iterations may be necessary to solve the reduced QP to optimality. These extra iterations are necessary to ensure that the terminated point gives a suitable direction for the linesearch.

In the major iteration log (see Section 13.2) a τ at the end of a line indicates that the corresponding QP was artificially terminated using the limit i .

Compare with the optional parameter **Iterations Limit**, which defines an independent *absolute* limit on the *total* number of minor iterations (summed over all QP subproblems).

Minor Print Level i Default = 1

This controls the amount of output to the **Print File** and the **Summary File** during solution of the QP subproblems. The value of i has the following effect:

i	Output
0	No minor iteration output except error messages.
≥ 1	A single line of output at each minor iteration (controlled by optional parameters Print Frequency and Summary Frequency).
≥ 10	Basis factorization statistics generated during the periodic refactorization of the basis (see the optional parameter Factorization Frequency). Statistics for the <i>first factorization</i> each major iteration are controlled by the optional parameter Major Print Level .

New Basis File	i_1	Default = 0
Backup Basis File	i_2	Default = 0
Save Frequency	i_3	Default = 100

New Basis File and **Backup Basis File** are sometimes referred to as basis maps. They contain the most compact representation of the state of each variable. They are intended for restarting the solution of a problem at a point that was reached by an earlier run. For nontrivial problems, it is advisable to save basis maps at the end of a run, in order to restart the run if necessary.

If $i_1 > 0$, a basis map will be saved on the file associated with unit i_1 every i_3 th iteration. The first record of the file will contain the word `PROCEEDING` if the run is still in progress. A basis map will also be saved at the end of a run, with some other word indicating the final solution status.

Use of $i_2 > 0$ is intended as a safeguard against losing the results of a long run. Suppose that a **New Basis File** is being saved every 100 (**Save Frequency**) iterations, and that E04WDF is about to save such a basis at iteration 2000. It is conceivable that the run may be interrupted during the next few milliseconds (in the middle of the save). In this case the Basis file will be corrupted and the run will have been essentially wasted.

To eliminate this risk, both a **New Basis File** and a **Backup Basis File** may be specified. The following would be suitable for the above example:

```
Backup Basis File 11
New Basis File 12
```

The current basis will then be saved every 100 iterations, first on the file associated with unit 12 and then immediately on the file associated with unit 11. If the run is interrupted at iteration 2000 during the save on the file associated with unit 12, there will still be a usable basis on the file associated with unit 11 (corresponding to iteration 1900).

Note that a new basis will be saved in **New Basis File** at the end of a run if it terminates normally, but it will not be saved in **Backup Basis File**. In the above example, if an optimum solution is found at iteration 2050 (or if the iteration limit is 2050), the final basis in the file associated with unit 12 will correspond to iteration 2050, but the last basis saved in the file associated with unit 11 will be the one for iteration 2000.

A full description of information recorded in **New Basis File** and **Backup Basis File** is given in Gill *et al.* (2005a).

New Superbasics Limit i Default = 99

This option causes early termination of the QP subproblems if the number of free variables has increased significantly since the first feasible point. If the number of new superbasics is greater than i , the nonbasic variables that have not yet moved are frozen and the resulting smaller QP is solved to optimality.

In the major iteration log (see Section 13.1), a τ at the end of a line indicates that the QP was terminated early in this way.

Old Basis File i Default = 0

If $i > 0$, the basis maps information will be obtained from this file. The file will usually have been output previously as a **New Basis File** or **Backup Basis File**. A full description of information recorded in **New Basis File** and **Backup Basis File** is given in Gill *et al.* (2005a).

The file will not be acceptable if the number of rows or columns in the problem has been altered.

Partial Price i Default = 1

This parameter is recommended for large problems that have significantly more variables than constraints. It reduces the work required for each ‘pricing’ operation (where a nonbasic variable is selected to become superbasic). When $i = 1$, all columns of the constraint matrix $(A \ -I)$ are searched. Otherwise, A and I are partitioned to give i roughly equal segments A_j and I_j , for $j = 1, 2, \dots, i$. If the previous pricing search was successful on A_{j-1} and I_{j-1} , the next search begins on the segments A_j and I_j . (All subscripts here are modulo i .) If a reduced gradient is found that is larger than some dynamic tolerance, the variable with the largest such reduced gradient (of appropriate sign) is selected to become superbasic. If nothing is found, the search continues on the next segments A_{j+1} and I_{j+1} , and so on.

For time-stage models having t time periods, **Partial Price** t (or $t/2$ or $t/3$) may be appropriate.

Pivot Tolerance r Default = ϵ^2

During the solution of QP subproblems, the pivot tolerance is used to prevent columns entering the basis if they would cause the basis to become almost singular.

When x changes to $x + \alpha p$ for some search direction p , a ‘ratio test’ determines which component of x reaches an upper or lower bound first. The corresponding element of p is called the pivot element. Elements of p are ignored (and therefore cannot be pivot elements) if they are smaller than the pivot tolerance r .

It is common for two or more variables to reach a bound at essentially the same time. In such cases, the **Minor Feasibility Tolerance** (say, t) provides some freedom to maximize the pivot element and thereby improve numerical stability. Excessively small values of t should therefore not be specified. To a lesser extent, the **Expand Frequency** (say, f) also provides some freedom to maximize the pivot element. Excessively large values of f should therefore not be specified.

Print File i Default = 0

If $i > 0$, the following information is output to a file associated with unit i during the solution of each problem:

- a listing of the optional parameters;
- some statistics about the problem;

- the amount of storage available for the *LU* factorization of the basis matrix;
- notes about the initial basis resulting from a Crash procedure or a Basis file;
- the iteration log;
- basis factorization statistics;
- the exit IFAIL condition and some statistics about the solution obtained;
- the printed solution, if requested.

These items are described in Sections 9 and 13. Further brief output may be directed to the **Summary File**.

Print Frequency *i* Default = 100

If $i > 0$, one line of the iteration log will be printed every i th iteration. A value such as $i = 10$ is suggested for those interested only in the final solution. If $i \leq 0$, the value of $i = 99999999$ is used and effectively no checks are made.

Proximal Point Method *i* Default = 1

$i = 1$ or 2 specifies minimization of $\|x - x_0\|_1$ or $\frac{1}{2}\|x - x_0\|_2^2$ when the starting point x_0 is changed to satisfy the linear constraints (where x_0 refers to nonlinear variables).

Punch File i_1 Default = 0
Insert File i_2 Default = 0

The **Punch File** from a previous run may be used as an **Insert File** for a later run on the same problem. A full description of information recorded in **Insert File** and **Punch File** is given in Gill *et al.* (2005a).

If $i_1 > 0$, the final solution obtained will be output to the file. For linear programs, this format is compatible with various commercial systems.

If $i_2 > 0$ the **Insert File** containing basis information will be read from unit i_2 . The file will usually have been output previously as a **Punch File**. The file will not be accessed if **Old Basis File** is specified.

QPSolver Cholesky Default
QPSolver CG
QPSolver QN

Specifies the active-set algorithm used to solve subproblem (11) (see Section 11.5). **QPSolver Cholesky** holds the full Cholesky factor R of the reduced Hessian $Z^T H Z$. As the QP iterations proceed, the dimension of R changes with the number of superbasic variables. If the number of superbasic variables needs to increase beyond the value of **Reduced Hessian Dimension**, the reduced Hessian cannot be stored and the solver switches to **QPSolver CG**. The Cholesky solver is reactivated if the number of superbasics stabilizes at a value less than **Reduced Hessian Dimension**.

QPSolver QN solves the QP using a quasi-Newton method. In this case, R is the factor of a quasi-Newton approximate Hessian.

QPSolver CG uses an active-set method similar to **QPSolver QN**, but uses the conjugate-gradient method to solve all systems involving the reduced Hessian.

The Cholesky QP solver is the most robust, but may require a significant amount of computation if there are many superbasics.

The quasi-Newton QP solver does not require computation of the exact R at the start of the subproblem in (11). It may be appropriate when the number of superbasics is large but relatively few iterations are needed to reach a solution (e.g., if E04WDF is called with a Warm Start).

The conjugate-gradient QP solver is appropriate for problems with many degrees of freedom (say, more than 2000 superbasics).

Reduced Hessian Dimension i Default = min(2000, n)

This specifies that an i by i triangular matrix R (to define the reduced Hessian according to $R^T R = Z^T H Z$) is to be available for use by the Cholesky QP solver.

Scale Option i Default = 0

Scale Tolerance r Default = 0.9

Scale Print

Three scale options are available as follows:

- | i | Meaning |
|-----|--|
| 0 | No scaling. This is recommended if it is known that x and the constraint matrix never have very large elements (say, larger than 100). |
| 1 | The constraints and variables are scaled by an iterative procedure that attempts to make the matrix coefficients as close as possible to 1.0 (see Fourer (1982)). This will sometimes improve the performance of the solution procedures. |
| 2 | The constraints and variables are scaled by the iterative procedure. Also, a certain additional scaling is performed that may be helpful if the right-hand side b or the solution x is large. This takes into account columns of $(A \ -I)$ that are fixed or have positive lower bounds or negative upper bounds. |

Optional parameter **Scale Tolerance** affects how many passes might be needed through the constraint matrix. On each pass, the scaling procedure computes the ratio of the largest and smallest nonzero coefficients in each column:

$$\rho_j = \max_i |a_{ij}| / \min_i |a_{ij}| \quad (a_{ij} \neq 0).$$

If $\max_j \rho_j$ is less than r times its previous value, another scaling pass is performed to adjust the row and column scales. Raising r from 0.9 to 0.99 (say) usually increases the number of scaling passes through A . At most 10 passes are made. The value of r should lie in the range $0 < r < 1$.

Scale Print causes the row scales $r(i)$ and column scales $c(j)$ to be printed to **Print File**, if **System Information Yes** has been specified. The scaled matrix coefficients are $\bar{a}_{ij} = a_{ij}c(j)/r(i)$, and the scaled bounds on the variables and slacks are $\bar{l}_j = l_j/c(j)$, $\bar{u}_j = u_j/c(j)$, where $c(j) = r(j - n)$ if $j > n$.

Solution File i Default = 0

If $i > 0$, the final solution will be output to file i (whether optimal or not). All numbers are printed in 1pe16.6 format.

To see more significant digits in the printed solution, it will sometimes be useful to make i refer to **Print File**.

Start Objective Check At Variable i Default = 1

Stop Objective Check At Variable i Default = n

Start Constraint Check At Variable i Default = 1

Stop Constraint Check At Variable i Default = n

These keywords take effect only if **Verify Level** > 0. They may be used to control the verification of gradient elements computed by function OBJFUN and/or Jacobian elements computed by function CONFUN. For example, if the first 30 elements of the objective gradient appeared to be correct in an earlier run, so that only element 31 remains questionable, it is reasonable to specify **Start Objective Check At Variable** = 31. If the first 30 variables appear linearly in the objective, so that the corresponding gradient elements are constant, the above choice would also be appropriate.

Summary File i_1

Default = 0

Summary Frequency i_2

Default = 100

If $i_1 > 0$, a brief log will be output to the file associated with unit i_1 , including one line of information every i_2 th iteration. In an interactive environment, it is useful to direct this output to the terminal, to allow a run to be monitored online. (If something looks wrong, the run can be manually terminated.) Further details are given in Section 13.6.

Superbasics Limit i Default = n

This option places a limit on the storage allocated for superbasic variables. Ideally, i should be set slightly larger than the ‘number of degrees of freedom’ expected at an optimal solution.

For nonlinear problems, the number of degrees of freedom is often called the ‘number of independent variables’. Normally, i need not be greater than $n_N + 1$, where n_N is the number of nonlinear variables. For many problems, i may be considerably smaller than n_N . This will save storage if n_N is very large.

Suppress Parameters

Normally E04WDF prints the options file as it is being read, and then prints a complete list of the available keywords and their final values. The optional parameter **Suppress Parameters** tells E04WDF not to print the full list.

System Information No

Default

System Information Yes

This option prints additional information on the progress of major and minor iterations, and Crash statistics. See Section 13.

Timing Level i

Default = 0

If $i > 0$, some timing information will be output to the Print file, if **Print File** > 0 .

Unbounded Objective r_1 Default = $1.0\text{E} + 15$ **Unbounded Step Size** r_2 Default = *bigbnd*

These parameters are intended to detect unboundedness in nonlinear problems. During a linesearch, F is evaluated at points of the form $x + \alpha p$, where x and p are fixed and α varies. If $|F|$ exceeds r_1 or α exceeds r_2 , iterations are terminated with the exit message IFAIL = 5.

If singularities are present, unboundedness in $F(x)$ may be manifested by a floating-point overflow (during the evaluation of $F(x + \alpha p)$), before the test against r_1 can be made.

Unboundedness in x is best avoided by placing finite upper and lower bounds on the variables.

Verify Level i

Default = 0

This option refers to finite difference checks on the derivatives computed by the user-supplied routines. Derivatives are checked at the first point that satisfies all bounds and linear constraints.

 i **Meaning**

- 0 Only a ‘cheap’ test will be performed, requiring two calls to CONFUN.
- 1 Individual gradients will be checked (with a more reliable test). A key of the form OK or Bad? indicates whether or not each component appears to be correct.
- 2 Individual columns of the problem Jacobian will be checked.
- 3 Options 2 and 1 will both occur (in that order).
- 1 Derivative checking is disabled.

Verify Level = 3 should be specified whenever a new function routine is being developed. Missing derivatives are not checked, so they result in no overhead.

Violation Limit r

Default = 1.0E + 6

This keyword defines an absolute limit on the magnitude of the maximum constraint violation, r , after the linesearch. On completion of the linesearch, the new iterate x_{k+1} satisfies the condition

$$v_i(x_{k+1}) \leq r \max(1, v_i(x_0)),$$

where x_0 is the point at which the nonlinear constraints are first evaluated and $v_i(x)$ is the i th nonlinear constraint violation $v_i(x) = \max(0, l_i - c_i(x), c_i(x) - u_i)$.

The effect of this violation limit is to restrict the iterates to lie in an *expanded* feasible region whose size depends on the magnitude of r . This makes it possible to keep the iterates within a region where the objective is expected to be well defined and bounded below. If the objective is bounded below for all values of the variables, then r may be any large positive value.

13 Description of Monitoring Information

E04WDF produces monitoring information, statistical information and information about the solution. Section 9.1 contains the final output information sent to unit **Print File**. This section contains other output information.

13.1 Major Iteration Log

This section describes the output to unit **Print File** if **Major Print Level** > 0. One line of information is output every k th major iteration, where k is **Print Frequency**.

Label	Description
Itns	is the cumulative number of minor iterations.
Major	is the current major iteration number.
Minors	is the number of iterations required by both the feasibility and optimality phases of the QP subproblem. Generally, Minors will be 1 in the later iterations, since theoretical analysis predicts that the correct active set will be identified near the solution (see Section 11).
Step	is the step length α taken along the current search direction p . The variables x have just been changed to $x + \alpha p$. On reasonably well-behaved problems, the unit step will be taken as the solution is approached.
nCon or nObj	nCon is the number of times CONFUN has been called to evaluate the nonlinear problem functions. Evaluations needed for the estimation of the derivatives by finite differences are not included. nCon is printed as a guide to the amount of work required for the linesearch. If n_N , the number of nonlinear constraints, is zero, nCon does not appear, but is replaced by nObj. This quantity is the number of calls made to OBJFUN.
Feasible	is the value of v_{\max} (see (12)), the maximum component of the scaled nonlinear constraint residual (see the description of the optional parameter Major Feasibility Tolerance). The solution is regarded as acceptably feasible if Feasible is less than the Major Feasibility Tolerance . In this case, the entry is contained in parentheses. If the constraints are linear, all iterates are feasible and this entry is not printed.
Optimal	is the value of c_{\max} (see (13)), the maximum complementary gap (see the description of the optional parameter Major Optimality Tolerance). It is an estimate of the degree of nonoptimality of the reduced costs. Both Feasible and Optimal are small in the neighbourhood of a solution.
MeritFunction or Objective	is the value of the augmented Lagrangian merit function (see (8)). This function will decrease at each iteration unless it was necessary to increase the penalty

parameters (see Section 11.4). As the solution is approached, MeritFunction will converge to the value of the objective at the solution.

In elastic mode, the merit function is a composite function involving the constraint violations weighted by the elastic weight.

If the constraints are linear, this item is labelled **Objective**, the value of the objective function. It will decrease monotonically to its optimal value.

L+U is the number of nonzeros representing the basis factors L and U on completion of the QP subproblem.

If nonlinear constraints are present, the basis factorization $B = LU$ is computed at the start of the first minor iteration. At this stage, $L+U = \text{lenL} + \text{lenU}$, where lenL (see Section 13.4) is the number of subdiagonal elements in the columns of a lower triangular matrix and lenU (see Section 13.4) is the number of diagonal and superdiagonal elements in the rows of an upper-triangular matrix.

As columns of B are replaced during the minor iterations, $L+U$ may fluctuate up or down but, in general, will tend to increase. As the solution is approached and the minor iterations decrease towards zero, $L+U$ will reflect the number of nonzeros in the LU factors at the start of the QP subproblem.

If the constraints are linear, refactorization is subject only to the **Factorization Frequency**, and $L+U$ will tend to increase between factorizations.

BSwap is the number of columns of the basis matrix B that were swapped with columns of S to improve the condition of B . The swaps are determined by an LU factorization of the rectangular matrix $B_S = (B \ S)^T$ with stability being favoured more than sparsity.

nS is the current number of superbasic variables.

condHz is an estimate of the condition number of $R^T R$, itself an estimate of $Z^T H Z$, the reduced Hessian of the Lagrangian. The condition number is the square of the ratio of the largest and smallest diagonals of the upper triangular matrix R , this being a lower bound on the condition number of $R^T R$. **condHz** gives a rough indication of whether or not the optimization procedure is having difficulty. If ϵ is the relative *machine precision* being used, the SQP algorithm will make slow progress if **condHz** becomes as large as $\epsilon^{-1/2} \approx 10^8$, and will probably fail to find a better solution if **condHz** reaches $\epsilon^{-3/4} \approx 10^{12}$.

To guard against high values of **condHz**, attention should be given to the scaling of the variables and the constraints. In some cases it may be necessary to add upper or lower bounds to certain variables to keep them a reasonable distance from singularities in the nonlinear functions or their derivatives.

Penalty is the Euclidean norm of the vector of penalty parameters used in the augmented Lagrangian merit function (not printed if there are no nonlinear constraints).

The summary line may include additional code characters that indicate what happened during the course of the major iteration. These will follow the separator ‘_’ in the output.

Label	Description
c	central differences have been used to compute the unknown components of the objective and constraint gradients. A switch to central differences is made if either the linesearch gives a small step, or x is close to being optimal. In some cases, it may be necessary to re-solve the QP subproblem with the central difference gradient and Jacobian.
d	during the linesearch it was necessary to decrease the step in order to obtain a maximum constraint violation conforming to the value of Violation Limit .
D	you set $\text{MODE} = -1$ on exit from OBJFUN, indicating that the linesearch needed to be done with a smaller value of the step length α .

l	the norm wise change in the variables was limited by the value of the Major Step Limit . If this output occurs repeatedly during later iterations, it may be worthwhile increasing the value of Major Step Limit .
i	if E04WDF is not in elastic mode, an i signifies that the QP subproblem is infeasible. This event triggers the start of nonlinear elastic mode, which remains in effect for all subsequent iterations. Once in elastic mode, the QP subproblems are associated with the elastic problem (11) (see Section 11.5). If E04WDF is already in elastic mode, an i indicates that the minimizer of the elastic subproblem does not satisfy the linearized constraints. (In this case, a feasible point for the usual QP subproblem may or may not exist.)
M	an extra evaluation of the problem functions was needed to define an acceptable positive definite quasi-Newton update to the Lagrangian Hessian. This modification is only done when there are nonlinear constraints.
m	this is the same as M except that it was also necessary to modify the update to include an augmented Lagrangian term.
n	no positive definite BFGS update could be found. The approximate Hessian is unchanged from the previous iteration.
R	the approximate Hessian has been reset by discarding all but the diagonal elements. This reset will be forced periodically by the Hessian Frequency and Hessian Updates keywords. However, it may also be necessary to reset an ill-conditioned Hessian from time to time.
r	the approximate Hessian was reset after ten consecutive major iterations in which no BFGS update could be made. The diagonals of the approximate Hessian are retained if at least one update has been done since the last reset. Otherwise, the approximate Hessian is reset to the identity matrix.
s	a self-scaled BFGS update was performed. This update is used when the Hessian approximation is diagonal, and hence always follows a Hessian reset.
t	the minor iterations were terminated because of the Minor Iterations Limit .
T	the minor iterations were terminated because of the New Superbasics Limit .
u	the QP subproblem was unbounded.
w	a weak solution of the QP subproblem was found.
z	the Superbasics Limit was reached.

13.2 Minor Iteration Log

If **Minor Print Level** > 0, one line of information is output to the Print file every k th minor iteration, where k is the specified **Print Frequency**. A heading is printed before the first such line following a basis factorization. The heading contains the items described below. In this description, a pricing operation is the process by which a nonbasic variable is selected to become superbasic (in addition to those already in the superbasic set). The selected variable is denoted by jq. Variable jq often becomes basic immediately. Otherwise it remains superbasic, unless it reaches its opposite bound and returns to the nonbasic set.

If **Partial Price** is in effect, variable jq is selected from A_{pp} or I_{pp} , the ppth segments of the constraint matrix $(A \ -I)$.

Label	Description
Itn	the current iteration number.
LPmult or QPmult	is the reduced cost (or reduced gradient) of the variable jq selected by the pricing procedure at the start of the present iteration. Algebraically, the reduced gradient is $d_j = g_j - \pi^T a_j$ for $j = jq$, where g_j is the gradient of the current objective function, π is the

	vector of dual variables for the QP subproblem, and a_j is the j th column of $\begin{pmatrix} A & -I \end{pmatrix}$.
	Note that the reduced cost is the 1-norm of the reduced-gradient vector at the start of the iteration, just after the pricing procedure.
LPstep or QPstep	is the step length α taken along the current search direction p . The variables x have just been changed to $x + \alpha p$. Write Step to stand for LPStep or QPStep, depending on the problem. If a variable is made superbasic during the current iteration ($+SBS > 0$), Step will be the step to the nearest bound. During the solution of (11), the step can be greater than one only if the reduced Hessian is not positive definite.
nInf	is the number of infeasibilities <i>after</i> the present iteration. This number will not increase unless the iterations are in elastic mode.
SumInf	is the sum of infeasibilities after the present iteration, if $nInf > 0$. The value usually decreases at each nonzero Step, but if it decreases by 2 or more, SumInf may occasionally increase.
rgNorm	is the norm of the reduced-gradient vector at the start of the iteration. (It is the norm of the vector with elements d_j for variables j in the superbasic set.) During the solution of subproblem (11) this norm will be approximately zero after a unit step. (The heading is not printed if the problem is linear.)
LPObjective, QPObjective or Elastic QPobj	the QP objective function after the present iteration. In elastic mode, the heading is changed to Elastic QPobj. In either case, the value printed decreases monotonically.
+SBS	is the variable j_q selected by the pricing operation to be added to the superbasic set.
-SBS	is the superbasic variable chosen to become nonbasic.
-BS	is the basis variable removed (if any) to become nonbasic.
Pivot	if column a_q replaces the r th column of the basis B , Pivot is the r th element of a vector y satisfying $By = a_q$. Wherever possible, Step is chosen to avoid extremely small values of Pivot (since they cause the basis to be nearly singular). In rare cases, it may be necessary to increase the Pivot Tolerance to exclude very small elements of y from consideration during the computation of Step.
L+U	is the number of nonzeros representing the basis factors L and U . Immediately after a basis factorization $B = LU$, $L+U$ is $\text{len}L + \text{len}U$, the number of subdiagonal elements in the columns of a lower triangular matrix and the number of diagonal and superdiagonal elements in the rows of an upper-triangular matrix. Further nonzeros are added to L when various columns of B are later replaced. As columns of B are replaced, the matrix U is maintained explicitly (in sparse form). The value of L will steadily increase, whereas the value of U may fluctuate up or down. Thus the value of $L+U$ may fluctuate up or down (in general, it will tend to increase).
ncp	is the number of compressions required to recover storage in the data structure for U . This includes the number of compressions needed during the previous basis factorization.
nS	is the current number of superbasic variables. (The heading is not printed if the problem is linear.)

condHz see Section 13.1. (The heading is not printed if the problem is linear.)

13.3 Crash Statistics

If **Major Print Level** ≥ 10 and **System Information Yes** has been specified, the following items are output to the Print file when **Cold Start** and no Backup Basis file is loaded. They refer to the number of columns that the Crash procedure selects during selected passes through A while searching for a triangular basis matrix.

Label	Description
Slacks	is the number of slacks selected initially.
Free cols	is the number of free columns in the basis.
Preferred	is the number of 'preferred' columns in the basis (i.e., $ISTATE(j) = 3$ for some $j \leq n$).
Unit	is the number of unit columns in the basis.
Double	is the number of columns in the basis containing 2 nonzeros.
Triangle	is the number of triangular columns in the basis.
Pad	is the number of slacks used to pad the basis (to make it a nonsingular triangle).

13.4 Basis Factorization Statistics

If **Major Print Level** ≥ 10 , the first seven items in the list below are output to the Print file whenever the basis B or the rectangular matrix $B_S = (B \ S)^T$ is factorized before solution of the next QP subproblem. See Section 12.1 for a full description of an optional parameter.

Gaussian elimination is used to compute a sparse LU factorization of B or B_S , where PLP^T and PUQ are lower and upper triangular matrices for some permutation matrices P and Q . Stability is ensured as described under the optional parameter **LU Factor Tolerance**.

If **Minor Print Level** ≥ 10 , the same items are printed during the QP solution whenever the current B is factorized. In addition, if **System Information Yes** has been specified, the entries from **Elms** onwards are also printed.

Label	Description														
Factor	the number of factorizations since the start of the run.														
Demand	a code giving the reason for the present factorization, as follows:														
	<table> <tr> <th>Code</th><th>Meaning</th></tr> <tr> <td>0</td><td>First LU factorization.</td></tr> <tr> <td>1</td><td>The number of updates reached the Factorization Frequency.</td></tr> <tr> <td>2</td><td>The nonzeros in the updated factors have increased significantly.</td></tr> <tr> <td>7</td><td>Not enough storage to update factors.</td></tr> <tr> <td>10</td><td>Row residuals too large (see the description of the optional parameter Check Frequency).</td></tr> <tr> <td>11</td><td>Ill-conditioning has caused inconsistent results.</td></tr> </table>	Code	Meaning	0	First LU factorization.	1	The number of updates reached the Factorization Frequency .	2	The nonzeros in the updated factors have increased significantly.	7	Not enough storage to update factors.	10	Row residuals too large (see the description of the optional parameter Check Frequency).	11	Ill-conditioning has caused inconsistent results.
Code	Meaning														
0	First LU factorization.														
1	The number of updates reached the Factorization Frequency .														
2	The nonzeros in the updated factors have increased significantly.														
7	Not enough storage to update factors.														
10	Row residuals too large (see the description of the optional parameter Check Frequency).														
11	Ill-conditioning has caused inconsistent results.														
Itn	is the current minor iteration number.														
Nonlin	is the number of nonlinear variables in the current basis B .														
Linear	is the number of linear variables in B .														
Slacks	is the number of slack variables in B .														
B BR BS or BT factorize	is the type of LU factorization.														

B	periodic factorization of the basis B .
BR	more careful rank-revealing factorization of B using threshold rook pivoting. This occurs mainly at the start, if the first basis factors seem singular or ill-conditioned. Followed by a normal B factorize.
BS	B_S is factorized to choose a well-conditioned B from the current $(B \ S)$. Followed by a normal B factorize.
BT	same as BS except the current B is tried first and accepted if it appears to be not much more ill-conditioned than after the previous BS factorize.
m	is the number of rows in B or B_S .
n	is the number of columns in B or B_S . Preceded by '=' or '>' respectively.
Elms	is the number of nonzero elements in B or B_S .
Amax	is the largest nonzero in B or B_S .
Density	is the percentage nonzero density of B or B_S .
Merit/MerRP/MerCP	is the average Markowitz merit count for the elements chosen to be the diagonals of PUQ . Each merit count is defined to be $(c-1)(r-1)$ where c and r are the number of nonzeros in the column and row containing the element at the time it is selected to be the next diagonal. Merit is the average of n such quantities. It gives an indication of how much work was required to preserve sparsity during the factorization. If LU Complete Pivoting or LU Rook Pivoting has been selected, this heading is changed to MerCP, respectively MerRP.
lenL	is the number of nonzeros in L .
L+U	is the number of nonzeros representing the basis factors L and U . Immediately after a basis factorization $B = LU$, $L+U$ is $\text{lenL} + \text{lenU}$, the number of subdiagonal elements in the columns of a lower triangular matrix and the number of diagonal and superdiagonal elements in the rows of an upper-triangular matrix. Further nonzeros are added to L when various columns of B are later replaced. As columns of B are replaced, the matrix U is maintained explicitly (in sparse form). The value of L will steadily increase, whereas the value of U may fluctuate up or down. Thus the value of $L+U$ may fluctuate up or down (in general, it will tend to increase).
Cmpressns	is the number of times the data structure holding the partially factored matrix needed to be compressed to recover unused storage. Ideally this number should be zero. If it is more than 3 or 4, the amount of workspace available to E04WDF should be increased for efficiency.
Incres	is the percentage increase in the number of nonzeros in L and U relative to the number of nonzeros in B or B_S .
Utri	is the number of triangular rows of B or B_S at the top of U .
lenU	the number of nonzeros in U , including its diagonals.
Ltol	is the largest subdiagonal element allowed in L . This is the specified LU Factor Tolerance or a smaller value that is currently being used for greater stability.
Umax	the maximum nonzero element in U .
Ugrwth	is the ratio U_{\max}/A_{\max} , which ideally should not be substantially larger than 10.0 or 100.0. If it is orders of magnitude larger, it may be advisable to reduce the LU Factor Tolerance to 5.0, 4.0, 3.0 or 2.0, say (but bigger than 1.0). As long as L_{\max} is not large (say 5.0 or less), $\max(A_{\max}, U_{\max})/D_{\min}$ gives an estimate of the condition number B . If this is extremely large, the basis is nearly singular. Slacks are used to replace suspect columns of B and the modified basis is refactored.
Ltri	is the number of triangular columns of B or B_S at the left of L .

dense1	is the number of columns remaining when the density of the basis matrix being factorized reached 0.3.
Lmax	is the actual maximum subdiagonal element in L (bounded by Ltol).
Akmax	is the largest nonzero generated at any stage of the LU factorization. (Values much larger than Amax indicate instability.) Akmax is not printed if LU Partial Pivoting is selected.
Agrowth	is the ratio Akmax/Amax. Values much larger than 100 (say) indicate instability. Growth is not printed if LU Partial Pivoting is selected.
bump	is the size of the block to be factorized nontrivially after the triangular rows and columns of B or B_S have been removed.
dense2	is the number of columns remaining when the density of the basis matrix being factorized reached 0.6. (The Markowitz pivot strategy searches fewer columns at that stage.)
DUmax	is the largest diagonal of PUQ .
DUmin	is the smallest diagonal of PUQ .
condU	the ratio DUmax/DUmin, which estimates the condition number of U (and of B if Ltol is less than 5.0, say).

13.5 The Solution File

At the end of a run, the final solution may be output as a Solution file, according to **Solution File**. Some header information appears first to identify the problem and the final state of the optimization procedure. A ROWS section and a COLUMNS section then follow, giving one line of information for each row and column. The format used is similar to certain commercial systems, though there is no industry standard.

In general, numerical values are output with format `f16.5`. The maximum record length is 111 characters, including the first (carriage-control) character.

To reduce clutter, a full stop (.) is printed for any numerical value that is exactly zero. The values ± 1 are also printed specially as 1.0 and -1.0 . Infinite bounds ($\pm 10^{20}$ or larger) are printed as None.

A Solution file is intended to be read from disk by a self-contained program that extracts and saves certain values as required for possible further computation. Typically, the first 14 records would be ignored. Each subsequent record may be read using

```
format(i8, 2x, 2a4, 1x, a1, 1x, a3, 5e16.6, i7)
```

adapted to suit the occasion. The end of the ROWS section is marked by a record that starts with a 1 and is otherwise blank. If this and the next 4 records are skipped, the COLUMNS section can then be read under the same format. (There should be no need for backspace statements.)

13.5.1 The ROWS section

General linear constraints take the form $l \leq Ax \leq u$. The i th constraint is therefore of the form

$$\alpha \leq \nu_i x \leq \beta,$$

where ν_i is the i th row of A .

Internally, the constraints take the form $Ax - s = 0$, where s is the set of slack variables (which happen to satisfy the bounds $l \leq s \leq u$). For the i th constraint it is the slack variable s_i that is directly available, and it is sometimes convenient to refer to its state. Nonlinear constraints $\alpha \leq c_i(x) + \nu_i x \leq \beta$ are treated similarly, except that the row activity and degree of infeasibility are computed directly from $c_i(x) + \nu_i x$ rather than s_i . A fullstop (.) is printed for any numerical value that is exactly zero.

Label	Description
Number	is the value of $n + i$. (This is used internally to refer to s_i in the intermediate output.)

Row	gives the name of the i th row.
State	the state of the i th row relative to the bounds α and β . The various states possible are as follows: <ul style="list-style-type: none"> LL the row is at its lower limit, α. UL the row is at its upper limit, β. EQ the limits are the same ($\alpha = \beta$). FR s_i is nonbasic and currently zero, even though it is free to take any value between its bounds α and β. BS s_i is basic. SBS s_i is superbasic. <p>A key is sometimes printed before State. Note that unless the optional parameter Scale Option = 0 is specified, the tests for assigning a key are applied to the variables of the scaled problem.</p> <ul style="list-style-type: none"> A <i>Alternative optimum possible</i>. The variable is nonbasic, but its reduced gradient is essentially zero. This means that if the variable were allowed to start moving away from its bound, there would be no change in the value of the objective function. The values of the other free variables <i>might</i> change, giving a genuine alternative solution. However, if there are any degenerate variables (labelled D), the actual change might prove to be zero, since one of them could encounter a bound immediately. In either case, the values of the Lagrange multipliers <i>might</i> also change. D <i>Degenerate</i>. The variable is basic or superbasic, but it is equal (or very close) to one of its bounds. I <i>Infeasible</i>. The variable is basic or superbasic and is currently violating one of its bounds by more than the value of the Feasibility Tolerance. N <i>Not precisely optimal</i>. The variable is nonbasic or superbasic. If the value of the reduced gradient for the variable exceeds the value of the optional parameter Major Optimality Tolerance, the solution would not be declared optimal because the reduced gradient for the variable would not be considered negligible.
Activity	is the value of $v_i x$ (or $c_i(x) + v_i x$ for nonlinear rows) at the final iterate.
Slack Activity	is the value by which the row differs from its nearest bound. (For the free row (if any), it is set to Activity.)
Lower Limit	is α , the lower bound on the row.
Upper Limit	is β , the upper bound on the row.
Dual Activity	is the value of the dual variable π_i (the Lagrange multiplier for the i th constraint). The full vector π always satisfies $B^T \pi = g_B$, where B is the current basis matrix and g_B contains the associated gradients for the current objective function. For FP problems, π_i is set to zero.
i	gives the index i of the i th row.

13.5.2 The COLUMNS section

Let the j th component of x be the variable x_j and assume that it satisfies the bounds $\alpha \leq x_j \leq \beta$. A fullstop (.) is printed for any numerical value that is exactly zero.

Label	Description
Number	is the column number j . (This is used internally to refer to x_j in the intermediate output.)

Column	gives the name of x_j .
State	the state of x_j relative to the bounds α and β . The various states possible are as follows: LL x_j is nonbasic at its lower limit, α . UL x_j is nonbasic at its upper limit, β . EQ x_j is nonbasic and fixed at the value $\alpha = \beta$. FR x_j is nonbasic at some value strictly between its bounds: $\alpha < x_j < \beta$. BS x_j is basic. Usually $\alpha < x_j < \beta$. SBS x_j is superbasic. Usually $\alpha < x_j < \beta$. A key is sometimes printed before State. Note that unless the optional parameter Scale Option = 0 is specified, the tests for assigning a key are applied to the variables of the scaled problem. A <i>Alternative optimum possible</i> . The variable is nonbasic, but its reduced gradient is essentially zero. This means that if the variable were allowed to start moving away from its bound, there would be no change in the value of the objective function. The values of the other free variables <i>might</i> change, giving a genuine alternative solution. However, if there are any degenerate variables (labelled D), the actual change might prove to be zero, since one of them could encounter a bound immediately. In either case, the values of the Lagrange multipliers <i>might</i> also change. D <i>Degenerate</i> . The variable is basic or superbasic, but it is equal (or very close) to one of its bounds. I <i>Infeasible</i> . The variable is basic or superbasic and is currently violating one of its bounds by more than the value of the Feasibility Tolerance . N <i>Not precisely optimal</i> . The variable is nonbasic or superbasic. If the value of the reduced gradient for the variable exceeds the value of the optional parameter Major Optimality Tolerance , the solution would not be declared optimal because the reduced gradient for the variable would not be considered negligible.
Activity	is the value of x_j at the final iterate.
Obj Gradient	is the value of g_j at the final iterate. For FP problems, g_j is set to zero.
Lower Limit	is the lower bound specified for the variable. None indicates that $BL(j) \leq -bigbnd$.
Upper Limit	is the upper bound specified for the variable. None indicates that $BU(j) \geq bigbnd$.
Reduced Gradnt	is the value of the reduced gradient $d_j = g_j - \pi^T a_j$ where a_j is the j th column of the constraint matrix. For FP problems, d_j is set to zero.
m + j	is the value of $m + j$.

Note that movement off a constraint (as opposed to a variable moving away from its bound) can be interpreted as allowing the entry in the Slack Activity column to become positive.

Numerical values are output with a fixed number of digits; they are not guaranteed to be accurate to this precision.

13.6 The Summary File

If **Summary File** > 0, the following information is output to the unit number associated with **Summary File**. (It is a brief summary of the output directed to unit **Print File**):

- the optional parameters supplied via the option setting routines, if any;
 - the Basis file loaded, if any;
 - a brief major iteration log (see Section 13.1);
 - a brief minor iteration log (see Section 13.2);
 - the exit condition, IFAIL;
 - a summary of the final iterate.
-