NAG Library Routine Document

E04UGF/E04UGA

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 9 of this document. If, however, you wish to reset some or all of the settings please refer to Section 10 for a detailed description of the algorithm, to Section 11 for a detailed description of the specification of the optional parameters and to Section 12 for a detailed description of the monitoring information produced by the routine.*

1 Purpose

E04UGF/E04UGA solves sparse nonlinear programming problems.

E04UGA is a version of E04UGF that has additional parameters in order to make it safe for use in multithreaded applications (see Section 5). The initialization routine E04WBF **must** have been called before calling E04UGA.

2 Specification

2.1 Specification for E04UGF

```
 SUBROUTINE E04UGF(CONFUN, OBJFUN, N, M, NCNLN, NONLN, NJNLN, IOBJ, NNZ,
1                  A, HA, KA, BL, BU, START, NNAME, NAMES, NS, XS,
2                  ISTATE, CLAMDA, MINIZ, MINZ, NINF, SINF, OBJ, IZ,
3                  LENIZ, Z, LENZ, IUSER, RUSER, IFAIL)

 INTEGER           N, M, NCNLN, NONLN, NJNLN, IOBJ, NNZ, HA(NNZ),
1                  KA(N+1), NNAME, NS, ISTATE(N+M), MINIZ, MINZ, NINF,
2                  IZ(LENIZ), LENIZ, LENZ, IUSER(*), IFAIL
 double precision  A(NNZ), BL(N+M), BU(N+M), XS(N+M), CLAMDA(N+M), SINF,
1                  OBJ, Z(LENZ), RUSER(*)
 CHARACTER*1       START
 CHARACTER*8       NAMES(NNAME)
 EXTERNAL          CONFUN, OBJFUN
```

2.2 Specification for E04UGA

```
 SUBROUTINE E04UGA(CONFUN, OBJFUN, N, M, NCNLN, NONLN, NJNLN, IOBJ, NNZ,
1                  A, HA, KA, BL, BU, START, NNAME, NAMES, NS, XS,
2                  ISTATE, CLAMDA, MINIZ, MINZ, NINF, SINF, OBJ, IZ,
3                  LENIZ, Z, LENZ, IUSER, RUSER, LWSAV, IWSAV, RWSAV,
4                  IFAIL)

 INTEGER           N, M, NCNLN, NONLN, NJNLN, IOBJ, NNZ, HA(NNZ),
1                  KA(N+1), NNAME, NS, ISTATE(N+M), MINIZ, MINZ, NINF,
2                  IZ(LENIZ), LENIZ, LENZ, IUSER(*), IWSAV(550), IFAIL
 double precision  A(NNZ), BL(N+M), BU(N+M), XS(N+M), CLAMDA(N+M), SINF,
1                  OBJ, Z(LENZ), RUSER(*), RWSAV(550)
 LOGICAL           LWSAV(20)
 CHARACTER*1       START
 CHARACTER*8       NAMES(NNAME)
 EXTERNAL          CONFUN, OBJFUN
```

Before calling E04UGA, or either of the option setting routines E04UHA or E04UJA, E04WBF **must** be called. The specification for E04WBF is:

```
 SUBROUTINE E04WBF(RNAME, CWSAV, LCWSAV, LWSAV, LLWSAV, IWSAV, LIWSAV,
1                  RWSAV, LRWSAV, IFAIL)

 INTEGER           LCWSAV, LLWSAV, IWSAV(LIWSAV), LIWSAV, LRWSAV, IFAIL
```

double precision	RWSAV(LRWSAV)
LOGICAL	LWSAV(LLWSAV)
CHARACTER*6	RNAME
CHARACTER*80	CWSAV(LCWSAV)

E04WBF should be called with RNAME = 'E04UGA'. LCWSAV, LLWSAV, LIWSAV and LRWSAV, the declared lengths of CWSAV, LWSAV, IWSAV and RWSAV respectively, must satisfy:

LCWSAV ≥ 1

LLWSAV ≥ 20

LIWSAV ≥ 550

LRWSAV ≥ 550

The contents of the arrays CWSAV, LWSAV, IWSAV and RWSAV **must not** be altered between calling routines E04UGA, E04UHA, E04UJA and E04WBF.

3 Description

E04UGF/E04UGA is designed to solve a class of nonlinear programming problems that are assumed to be stated in the following general form:

$$\underset{x\in R^n}{\text{minimize}}\, f(x) \qquad \text{subject to} \qquad l \leq \begin{pmatrix} x \\ F(x) \\ Gx \end{pmatrix} \leq u, \tag{1}$$

where $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ is a set of variables, $f(x)$ is a smooth scalar objective function, l and u are constant lower and upper bounds, $F(x)$ is a vector of smooth nonlinear constraint functions $\{F_i(x)\}$ and G is a *sparse* matrix.

The constraints involving F and Gx are called the *general constraints*. Note that upper and lower bounds are specified for all variables and constraints. This form allows full generality in specifying various types of constraint. In particular, the jth constraint can be defined as an *equality* by setting $l_j = u_j$. 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.)

E04UGF/E04UGA converts the upper and lower bounds on the m elements of F and Gx to equalities by introducing a set of *slack variables* s, where $s = (s_1, s_2, \ldots, s_m)^{\mathrm{T}}$. 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 problem defined by (1) can therefore be re-written in the following equivalent form:

$$\underset{x\in R^n, s\in R^m}{\text{minimize}}\, f(x) \qquad \text{subject to} \qquad \{\, Gx \,\} - s = 0, \qquad l \leq \begin{pmatrix} x \\ s \end{pmatrix} \leq u. \tag{2}$$

Since the slack variables s are subject to the same upper and lower bounds as the elements of F and Gx, the bounds on F and Gx can simply be thought of as bounds on the combined vector (x, s). The elements of x and s are partitioned into *basic*, *nonbasic* and *superbasic variables* defined as follows:

– a basic variable (x_j say) is the jth variable associated with the jth column of the basis matrix B;

– a nonbasic variable is a variable that is temporarily fixed at its current value (usually its upper or lower bound);

– a superbasic variable is a nonbasic variable which is not at one of its bounds that is free to move in any desired direction (namely one that will improve the value of the objective function or reduce the sum of infeasibilities).

For example, in the simplex method (see Gill *et al.* (1981)) the elements of x can be partitioned at each vertex into a set of m basic variables (all non-negative) and a set of $(n - m)$ nonbasic variables (all zero). This is equivalent to partitioning the columns of the constraint matrix as $(\, B \quad N \,)$, where B contains the m columns that correspond to the basic variables and N contains the $(n - m)$ columns that correspond to the nonbasic variables. Note that B is square and nonsingular.

The optional parameter Maximize may be used to specify an alternative problem in which $f(x)$ is maximized. If the objective function is nonlinear and all the constraints are linear, F is absent and the problem is said to be *linearly constrained*. In general, the objective and constraint functions are *structured* in the sense that they are formed from sums of linear and nonlinear functions. This structure can be exploited by the routine during the solution process as follows.

Consider the following nonlinear optimization problem with four variables (u, v, z, w):

$$\underset{u,v,z,w}{\text{minimize}} \qquad (u+v+z)^2 + 3z + 5w$$

subject to the constraints

$$\begin{aligned} u^2 + v^2 + z \quad &= 2 \\ u^4 + v^4 \quad + w &= 4 \\ 2u + 4v \quad &\geq 0 \end{aligned}$$

and to the bounds

$$\begin{aligned} z &\geq 0 \\ w &\geq 0. \end{aligned}$$

This problem has several characteristics that can be exploited by the routine:

– the objective function is nonlinear. It is the sum of a *nonlinear* function of the variables (u, v, z) and a *linear* function of the variables (z, w);

– the first two constraints are nonlinear. The third is linear;

– each nonlinear constraint function is the sum of a *nonlinear* function of the variables (u, v) and a *linear* function of the variables (z, w).

The nonlinear terms are defined by OBJFUN and CONFUN (see Section 5), which involve only the appropriate subset of variables.

For the objective, we define the function $f(u, v, z) = (u + v + z)^2$ to include only the nonlinear part of the objective. The three variables (u, v, z) associated with this function are known as the *nonlinear objective variables*. The number of them is given by NONLN (see Section 5) and they are the only variables needed in OBJFUN. The linear part $3z + 5w$ of the objective is stored in row IOBJ (see Section 5) of the (constraint) Jacobian matrix A (see below).

Thus, if x' and y' denote the nonlinear and linear objective variables, respectively, the objective may be re-written in the form

$$f(x') + c^{\mathrm{T}}x' + d^{\mathrm{T}}y',$$

where $f(x')$ is the nonlinear part of the objective and c and d are constant vectors that form a row of A. In this example, $x' = (u, v, z)$ and $y' = w$.

Similarly for the constraints, we define a vector function $F(u, v)$ to include just the nonlinear terms. In this example, $F_1(u, v) = u^2 + v^2$ and $F_2(u, v) = u^4 + v^4$, where the two variables (u, v) are known as the *nonlinear Jacobian variables*. The number of them is given by NJNLN (see Section 5) and they are the only variables needed in CONFUN. Thus, if x'' and y'' denote the nonlinear and linear Jacobian variables, respectively, the constraint functions and the linear part of the objective have the form

$$\begin{pmatrix} F(x'') + A_2 y'' \\ A_3 x'' + A_4 y'' \end{pmatrix}, \tag{3}$$

where $x'' = (u, v)$ and $y'' = (z, w)$ in this example. This ensures that the Jacobian is of the form

$$A = \begin{pmatrix} J(x'') & A_2 \\ A_3 & A_4 \end{pmatrix},$$

where $J(x'') = \dfrac{\partial F(x'')}{\partial x}$. Note that $J(x'')$ *always* appears in the *top left-hand corner* of A.

The inequalities $l_1 \leq F\left(x''\right) + A_2y'' \leq u_1$ and $l_2 \leq A_3x'' + A_4y'' \leq u_2$ implied by the constraint functions in (3) are known as the *nonlinear* and *linear* constraints, respectively. The nonlinear constraint vector $F\left(x''\right)$ in (3) and (optionally) its partial derivative matrix $J\left(x''\right)$ are set in CONFUN. The matrices A_2, A_3 and A_4 contain any (constant) linear terms. Along with the sparsity pattern of $J\left(x''\right)$ they are stored in the arrays A, HA and KA (see Section 5).

In general, the vectors x' and x'' have different dimensions, but they *always overlap*, in the sense that the shorter vector is always the beginning of the other. In the above example, the nonlinear Jacobian variables (u, v) are an ordered subset of the nonlinear objective variables (u, v, z). In other cases it could be the other way round (whichever is the most convenient), but the first way keeps $J\left(x''\right)$ as small as possible.

Note that the nonlinear objective function $f\left(x'\right)$ may involve either a subset or superset of the variables appearing in the nonlinear constraint functions $F\left(x''\right)$. Thus, NONLN $\leq$ NJNLN (or vice-versa). Sometimes the objective and constraints really involve *disjoint sets of nonlinear variables*. In such cases the variables should be ordered so that NONLN $>$ NJNLN and $x' = \left(x'', x'''\right)$, where the objective is nonlinear in just the last vector x'''. The first NJNLN elements of the gradient array OBJGRD should also be set to zero in OBJFUN. This is illustrated in Section 9.

If all elements of the constraint Jacobian are known (i.e., the optional parameter Derivative Level $= 2$ or 3), any constant elements may be assigned their correct values in A, HA and KA. The corresponding elements of the constraint Jacobian array FJAC need not be reset in CONFUN. This includes values that are identically zero as constraint Jacobian elements are assumed to be zero unless specified otherwise. It must be emphasised that, if Derivative Level $= 0$ or 1, unassigned elements of FJAC are *not* treated as constant; they are estimated by finite differences, at nontrivial expense.

If there are no nonlinear constraints in (1) and $f(x)$ is linear or quadratic, then it may be more efficient to use E04NQF to solve the resulting linear or quadratic programming problem, or one of E04MFF/E04MFA, E04NCF/E04NCA or E04NFF/E04NFA if G is a *dense* matrix. If the problem is dense and does have nonlinear constraints then one of E04UFF/E04UFA, E04USF/E04USA or E04WDF (as appropriate) should be used instead.

You must supply an initial estimate of the solution to (1), together with versions of OBJFUN and CONFUN that define $f\left(x'\right)$ and $F\left(x''\right)$, respectively, and as many first partial derivatives as possible. Note that if there are any nonlinear constraints, then the *first* call to CONFUN will precede the *first* call to OBJFUN.

E04UGF/E04UGA is based on the SNOPT package described in Gill *et al.* (2002), which in turn utilizes routines from the MINOS package (see Murtagh and Saunders (1995)). It incorporates a sequential quadratic programming (SQP) method that obtains search directions from a sequence of quadratic programming (QP) subproblems. Each QP subproblem minimizes a quadratic model of a certain Lagrangian function subject to a linearization of the constraints. An augmented Lagrangian merit function is reduced along each search direction to ensure convergence from any starting point. Further details can be found in Section 10.

Throughout this document the symbol ϵ is used to represent the ***machine precision*** (see X02AJF).

4 References

Conn A R (1973) Constrained optimization using a nondifferentiable penalty function *SIAM J. Numer. Anal.* **10** 760–779

Eldersveld S K (1991) Large-scale sequential quadratic programming algorithms *PhD Thesis* Department of Operations Research, Stanford University, Stanford

Fletcher R (1984) An l_1 penalty method for nonlinear constraints *Numerical Optimization 1984* (eds P T Boggs, R H Byrd and R B Schnabel) 26–40 SIAM Philadelphia

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, 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 (1989) A practical anti-cycling procedure for linearly constrained optimization *Math. Programming* **45** 437–474

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

Murtagh B A and Saunders M A (1995) MINOS 5.4 Users' Guide *Report SOL 83-20R* Department of Operations Research, Stanford University

Ortega J M and Rheinboldt W C (1970) *Iterative Solution of Nonlinear Equations in Several Variables* Academic Press

Powell M J D (1974) Introduction to constrained optimization *Numerical Methods for Constrained Optimization* (eds P E Gill and W Murray) 1–28 Academic Press

5 Parameters

1: CONFUN – SUBROUTINE, supplied by the NAG Library or the user. *External Procedure*

CONFUN must calculate the vector $F(x)$ of nonlinear constraint functions and (optionally) its Jacobian $\left(=\dfrac{\partial F}{\partial x}\right)$ for a specified n_1'' ($\leq n$) element vector x. If there are no nonlinear constraints (i.e., NCNLN $= 0$), CONFUN will never be called by E04UGF/E04UGA and CONFUN may be the dummy routine E04UGM. E04UGM is included in the NAG Library. If there are nonlinear constraints, the first call to CONFUN will occur before the first call to OBJFUN.

The specification of CONFUN is:

```
      SUBROUTINE CONFUN(MODE, NCNLN, NJNLN, NNZJAC, X, F, FJAC, NSTATE,
     1                  IUSER, RUSER)
      INTEGER            MODE, NCNLN, NJNLN, NNZJAC, NSTATE, IUSER(*)
      double precision   X(NJNLN), F(NCNLN), FJAC(NNZJAC), RUSER(*)
```

1: MODE – INTEGER *Input/Output*

On entry: indicates which values must be assigned during each call of CONFUN. Only the following values need be assigned:

MODE $= 0$
F.

MODE $= 1$
All available elements of FJAC.

MODE $= 2$
F and all available elements of FJAC.

On exit: you may set to a negative value as follows:

MODE ≤ -2
The solution to the current problem is terminated and in this case E04UGF/E04UGA will terminate with IFAIL set to MODE.

MODE $= -1$

The nonlinear constraint functions cannot be calculated at the current x. E04UGF/E04UGA will then terminate with IFAIL $= -1$ unless this occurs during the linesearch; in this case, the linesearch will shorten the step and try again.

2: NCNLN – INTEGER *Input*

On entry: n_N, the number of nonlinear constraints. These must be the first NCNLN constraints in the problem.

3: NJNLN – INTEGER *Input*

On entry: n_1'', the number of nonlinear variables. These must be the first NJNLN variables in the problem.

4: NNZJAC – INTEGER *Input*

On entry: the number of nonzero elements in the constraint Jacobian. Note that NNZJAC will usually be less than NCNLN $\times$ NJNLN.

5: X(NJNLN) – ***double precision*** array *Input*

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

6: F(NCNLN) – ***double precision*** array *Output*

On exit: if MODE $= 0$ or 2, F(i) must contain the value of the ith nonlinear constraint function at x.

7: FJAC(NNZJAC) – ***double precision*** array *Input/Output*

On entry: the elements of FJAC are set to special values which enable E04UGF/E04UGA to detect whether they are changed by CONFUN.

On exit: if MODE $= 1$ or 2, FJAC must return the available elements of the constraint Jacobian evaluated at x. These elements must be stored in exactly the same positions as implied by the definitions of the arrays A, HA and KA. If optional parameter Derivative Level $= 2$ or 3, the value of any constant Jacobian element not defined by CONFUN will be obtained directly from A. Note that the routine does not perform any internal checks for consistency (except indirectly via the optional parameter Verify Level), so great care is essential.

8: NSTATE – INTEGER *Input*

On entry: if NSTATE $= 1$, then E04UGF/E04UGA is calling CONFUN for the first time. This parameter setting allows you to save computation time if certain data must be read or calculated only once.

If NSTATE ≥ 2, then E04UGF/E04UGA is calling CONFUN for the last time. This parameter setting allows you to perform some additional computation on the final solution. In general, the last call to CONFUN is made with NSTATE $= 2 +$ IFAIL (see Section 6).

Otherwise, NSTATE $= 0$.

9: IUSER$(*)$ – INTEGER array *User Workspace*

10: RUSER$(*)$ – ***double precision*** array *User Workspace*

CONFUN is called from E04UGF/E04UGA with the parameters IUSER and RUSER as supplied to E04UGF/E04UGA. You are free to use the arrays IUSER and RUSER to supply information to CONFUN as an alternative to using COMMON global variables.

CONFUN must be declared as EXTERNAL in the (sub)program from which E04UGF/E04UGA is called. Parameters denoted as *Input* must **not** be changed by this procedure.

2: OBJFUN – SUBROUTINE, supplied by the NAG Library or the user. *External Procedure*

OBJFUN must calculate the nonlinear part of the objective function $f(x)$ and (optionally) its gradient $\left(=\dfrac{\partial f}{\partial x}\right)$ for a specified n_1' ($\leq n$) element vector x. If there are no nonlinear objective variables (i.e., NONLN $= 0$), OBJFUN will never be called by E04UGF/E04UGA and OBJFUN may be the dummy routine E04UGN. E04UGN is included in the NAG Library.

The specification of OBJFUN is:

```
SUBROUTINE OBJFUN(MODE, NONLN, X, OBJF, OBJGRD, NSTATE, IUSER,
1                 RUSER)
INTEGER           MODE, NONLN, NSTATE, IUSER(*)
double precision  X(NONLN), OBJF, OBJGRD(NONLN), RUSER(*)
```

1: MODE – INTEGER *Input/Output*

On entry: indicates 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 OBJGRD.

MODE $= 2$
OBJF and all available elements of OBJGRD.

On exit: you may set to a negative value as follows:

MODE ≤ -2
The solution to the current problem is terminated and in this case E04UGF/E04UGA will terminate with IFAIL set to MODE.

MODE $= -1$
The nonlinear part of the objective function cannot be calculated at the current x. E04UGF/E04UGA will then terminate with IFAIL $= -1$ unless this occurs during the linesearch; in this case, the linesearch will shorten the step and try again.

2: NONLN – INTEGER *Input*

On entry: n_1', the number of nonlinear objective variables. These must be the first NONLN variables in the problem.

3: X(NONLN) – ***double precision*** array *Input*

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

4: OBJF – ***double precision*** *Output*

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

5: OBJGRD(NONLN) – ***double precision*** array *Input/Output*

On entry: the elements of OBJGRD are set to special values which enable E04UGF/E04UGA to detect whether they are changed by OBJFUN.

On exit: if MODE $= 1$ or 2, OBJGRD must return the available elements of the gradient evaluated at x.

6: NSTATE – INTEGER *Input*

On entry: if NSTATE $= 1$, E04UGF/E04UGA is calling OBJFUN for the first time. This parameter setting allows you to save computation time if certain data must be read or calculated only once.

If NSTATE ≥ 2, E04UGF/E04UGA is calling OBJFUN for the last time. This parameter setting allows you to perform some additional computation on the final solution. In general, the last call to OBJFUN is made with NSTATE $= 2 +$ IFAIL (see Section 6).

Otherwise, NSTATE $= 0$.

7: IUSER$(*)$ – INTEGER array *User Workspace*

8: RUSER$(*)$ – ***double precision*** array *User Workspace*

OBJFUN is called from E04UGF/E04UGA with the parameters IUSER and RUSER as supplied to E04UGF/E04UGA. You are free to use the arrays IUSER and RUSER to supply information to OBJFUN as an alternative to using COMMON global variables.

OBJFUN must be declared as EXTERNAL in the (sub)program from which E04UGF/E04UGA is called. Parameters denoted as *Input* must **not** be changed by this procedure.

3: N – INTEGER *Input*

On entry: n, the number of variables (excluding slacks). This is the number of columns in the full Jacobian matrix A.

Constraint: N ≥ 1.

4: M – INTEGER *Input*

On entry: m, the number of general constraints (or slacks). This is the number of rows in A, including the free row (if any; see IOBJ). Note that A must contain at least one row. If your problem has no constraints, or only upper and lower bounds on the variables, then you must include a dummy 'free' row consisting of a single (zero) element subject to 'infinite' upper and lower bounds. Further details can be found under the descriptions for IOBJ, NNZ, A, HA, KA, BL and BU.

Constraint: M ≥ 1.

5: NCNLN – INTEGER *Input*

On entry: n_N, the number of nonlinear constraints.

Constraint: $0 \leq$ NCNLN $\leq$ M.

6: NONLN – INTEGER *Input*

On entry: n_1', the number of nonlinear objective variables. If the objective function is nonlinear, the leading n_1' columns of A belong to the nonlinear objective variables. (See also the description for NJNLN.)

Constraint: $0 \leq$ NONLN $\leq$ N.

7: NJNLN – INTEGER *Input*

On entry: n_1'', the number of nonlinear Jacobian variables. If there are any nonlinear constraints, the leading n_1'' columns of A belong to the nonlinear Jacobian variables. If $n_1' > 0$ and $n_1'' > 0$, the nonlinear objective and Jacobian variables overlap. The total number of nonlinear variables is given by $\bar{n} = \max(n_1', n_1'')$.

Constraints:

if NCNLN $= 0$, NJNLN $= 0$;
if NCNLN > 0, $1 \leq$ NJNLN $\leq$ N.

8: IOBJ – INTEGER *Input*

On entry: if IOBJ $>$ NCNLN, row IOBJ of A is a free row containing the nonzero elements of the linear part of the objective function.

IOBJ $= 0$
There is no free row.

IOBJ $= -1$
There is a dummy 'free' row.

Constraints:

if IOBJ > 0, NCNLN $<$ IOBJ $\leq$ M;
otherwise IOBJ ≥ -1.

9: NNZ – INTEGER *Input*

On entry: the number of nonzero elements in A (including the Jacobian for any nonlinear constraints). If IOBJ $= -1$, set NNZ $= 1$.

Constraint: $1 \leq$ NNZ $\leq$ N $\times$ M.

10: A(NNZ) – ***double precision*** array *Input/Output*

On entry: the nonzero elements of the Jacobian matrix A, ordered by increasing column index. Since the constraint Jacobian matrix $J(x'')$ must always appear in the top left-hand corner of A, those elements in a column associated with any nonlinear constraints must come before any elements belonging to the linear constraint matrix G and the free row (if any; see IOBJ).

In general, A is partitioned into a nonlinear part and a linear part corresponding to the nonlinear variables and linear variables in the problem. Elements in the nonlinear part may be set to any value (e.g., zero) because they are initialized at the first point that satisfies the linear constraints and the upper and lower bounds.

If Derivative Level $= 2$ or 3, the nonlinear part may also be used to store any constant Jacobian elements. Note that if CONFUN does not define the constant Jacobian element FJAC(i) then the missing value will be obtained directly from A(j) for some $j \geq i$.

If Derivative Level $= 0$ or 1, unassigned elements of FJAC are *not* treated as constant; they are estimated by finite differences, at nontrivial expense.

The linear part must contain the nonzero elements of G and the free row (if any). If IOBJ $= -1$, set A$(1) = 0$. Elements with the same row and column indices are not allowed. (See also the descriptions for HA and KA.)

On exit: elements in the nonlinear part corresponding to nonlinear Jacobian variables are overwritten.

11: HA(NNZ) – INTEGER array *Input*

On entry: HA(i) must contain the row index of the nonzero element stored in A(i), for $i = 1, 2, \ldots,$ NNZ. The row indices for a column may be supplied in any order subject to the condition that those elements in a column associated with any nonlinear constraints must appear before those elements associated with any linear constraints (including the free row, if any). Note that CONFUN must define the Jacobian elements in the same order. If IOBJ $= -1$, set HA$(1) = 1$.

Constraint: $1 \leq$ HA$(i) \leq$ M, for $i = 1, 2, \ldots,$ NNZ.

12: KA(N + 1) – INTEGER array *Input*

On entry: KA(j) must contain the index in A of the start of the jth column, for $j = 1, 2, \ldots,$ N. To specify the jth column as empty, set KA$(j) =$ KA$(j + 1)$. Note that the first and last elements of KA must be such that KA$(1) = 1$ and KA(N + 1) $=$ NNZ $+ 1$. If IOBJ $= -1$, set KA$(j) = 2$, for $j = 2, 3, \ldots,$ N.

Constraints:

$\mathrm{KA}(1) = 1$;
$\mathrm{KA}(j) \geq 1$, for $j = 2, 3, \ldots, \mathrm{N}$;
$\mathrm{KA}(\mathrm{N}+1) = \mathrm{NNZ} + 1$;
$0 \leq \mathrm{KA}(j+1) - \mathrm{KA}(j) \leq \mathrm{M}$, for $j = 1, 2, \ldots, \mathrm{N}$.

13: BL(N + M) – ***double precision*** array *Input*

On entry: l, the lower bounds for all the variables and general constraints, in the following order. The first N elements of BL must contain the bounds on the variables x, the next NCNLN elements the bounds for the nonlinear constraints $F(x)$ (if any) and the next (M − NCNLN) elements the bounds for the linear constraints Gx and the free row (if any). To specify a nonexistent lower bound (i.e., $l_j = -\infty$), set $\mathrm{BL}(j) \leq -$*bigbnd*. To specify the jth constraint as an *equality*, set $\mathrm{BL}(j) = \mathrm{BU}(j) = \beta$, say, where $|\beta| <$ *bigbnd*. If IOBJ $= -1$, set $\mathrm{BL}(\mathrm{N} + \mathrm{abs}(\mathrm{IOBJ})) \leq -$ *bigbnd*.

Constraint: if NCNLN $<$ IOBJ $\leq$ M or IOBJ $= -1$, $\mathrm{BL}(\mathrm{N} + \mathrm{abs}(\mathrm{IOBJ})) \leq -$*bigbnd*

(See also the description for BU.)

14: BU(N + M) – ***double precision*** array *Input*

On entry: u, the upper bounds for all the variables and general constraints, in the following order. The first N elements of BU must contain the bounds on the variables x, the next NCNLN elements the bounds for the nonlinear constraints $F(x)$ (if any) and the next (M − NCNLN) elements the bounds for the linear constraints Gx and the free row (if any). To specify a nonexistent upper bound (i.e., $u_j = +\infty$), set $\mathrm{BU}(j) \geq$ *bigbnd*. To specify the jth constraint as an *equality*, set $\mathrm{BU}(j) = \mathrm{BL}(j) = \beta$, say, where $|\beta| <$ *bigbnd*. If IOBJ $= -1$, set $\mathrm{BU}(\mathrm{N} + \mathrm{abs}(\mathrm{IOBJ})) \geq$ *bigbnd*.

Constraints:

if NCNLN $<$ IOBJ $\leq$ M or IOBJ $= -1$, $\mathrm{BU}(\mathrm{N} + \mathrm{abs}(\mathrm{IOBJ})) \geq$ *bigbnd*;
$\mathrm{BL}(j) \leq \mathrm{BU}(j)$, for $j = 1, 2, \ldots, \mathrm{N} + \mathrm{M}$;
if $\mathrm{BL}(j) = \mathrm{BU}(j) = \beta$, $|\beta| <$ *bigbnd*.

15: START – CHARACTER*1 *Input*

On entry: indicates how a starting basis is to be obtained.

START = 'C'
An internal Crash procedure will be used to choose an initial basis.

START = 'W'
A basis is already defined in ISTATE and NS (probably from a previous call).

Constraint: START = 'C' or 'W'.

16: NNAME – INTEGER *Input*

On entry: the number of column (i.e., variable) and row (i.e., constraint) names supplied in NAMES.

NNAME = 1
There are no names. Default names will be used in the printed output.

NNAME = N + M
All names must be supplied.

Constraint: NNAME = 1 or N + M.

17: NAMES(NNAME) – CHARACTER*8 array *Input*

On entry: specifies the column and row names to be used in the printed output.

If NNAME = 1, NAMES is not referenced and the printed output will use default names for the columns and rows.

If NNAME = N + M, the first N elements must contain the names for the columns, the next NCNLN elements must contain the names for the nonlinear rows (if any) and the next (M – NCNLN) elements must contain the names for the linear rows (if any) to be used in the printed output. Note that the name for the free row or dummy 'free' row must be stored in NAMES(N + abs(IOBJ)).

18: NS – INTEGER *Input/Output*

On entry: n_S, the number of superbasics. It need not be specified if START = 'C', but must retain its value from a previous call when START = 'W'.

On exit: the final number of superbasics.

19: XS(N + M) – ***double precision*** array *Input/Output*

On entry: the initial values of the variables and slacks (x, s). (See the description for ISTATE.)

On exit: the final values of the variables and slacks (x, s).

20: ISTATE(N + M) – INTEGER array *Input/Output*

On entry: if START = 'C', the first N elements of ISTATE and XS must specify the initial states and values, respectively, of the variables x. (The slacks s need not be initialized.) An internal Crash procedure is then used to select an initial basis matrix B. The initial basis matrix will be triangular (neglecting certain small elements in each column). It is chosen from various rows and columns of $(A \quad -I)$. Possible values for ISTATE(j) are as follows:

ISTATE(*j*)	**State of XS(*j*) during Crash procedure**
0 or 1	Eligible for the basis
2	Ignored
3	Eligible for the basis (given preference over 0 or 1)
4 or 5	Ignored

If nothing special is known about the problem, or there is no wish to provide special information, you may set ISTATE(j) = 0 and XS(j) = 0.0, for $j = 1, 2, \ldots,$ N. All variables will then be eligible for the initial basis. Less trivially, to say that the jth variable will probably be equal to one of its bounds, set ISTATE(j) = 4 and XS(j) = BL(j) or ISTATE(j) = 5 and XS(j) = BU(j) as appropriate.

Following the Crash procedure, variables for which ISTATE(j) = 2 are made superbasic. Other variables not selected for the basis are then made nonbasic at the value XS(j) if BL(j) ≤ XS(j) ≤ BU(j), or at the value BL(j) or BU(j) closest to XS(j).

If START = 'W', ISTATE and XS must specify the initial states and values, respectively, of the variables and slacks (x, s). If the routine has been called previously with the same values of N and M, ISTATE already contains satisfactory information.

Constraints:

if START = 'C', $0 \leq$ ISTATE(j) ≤ 5, for $j = 1, 2, \ldots,$ N;
if START = 'W', $0 \leq$ ISTATE(j) ≤ 3, for $j = 1, 2, \ldots,$ N + M.

On exit: the final states of the variables and slacks (x, s). The significance of each possible value of ISTATE(j) is as follows:

ISTATE(*j*)	**State of variable *j***	**Normal value of XS(*j*)**
0	Nonbasic	BL(j)
1	Nonbasic	BU(j)
2	Superbasic	Between BL(j) and BU(j)
3	Basic	Between BL(j) and BU(j)

If NINF $= 0$, basic and superbasic variables may be outside their bounds by as much as the value of the optional parameter Minor Feasibility Tolerance. Note that if scaling is specified, the optional parameter Minor Feasibility Tolerance applies to the variables of the *scaled* problem. In this case, the variables of the original problem may be as much as 0.1 outside their bounds, but this is unlikely unless the problem is very badly scaled.

Very occasionally some nonbasic variables may be outside their bounds by as much as the optional parameter Minor Feasibility Tolerance and there may be some nonbasic variables for which XS(j) lies strictly between its bounds.

If NINF > 0, some basic and superbasic variables may be outside their bounds by an arbitrary amount (bounded by SINF if scaling was not used).

21: CLAMDA(N + M) – ***double precision*** array *Input/Output*

On entry: if NCNLN > 0, CLAMDA(j) must contain a Lagrange multiplier estimate for the jth nonlinear constraint $F_j(x)$, for $j = \mathrm{N}+1, \mathrm{N}+2, \ldots, \mathrm{N}+\mathrm{NCNLN}$. If nothing special is known about the problem, or there is no wish to provide special information, you may set CLAMDA$(j) = 0.0$. The remaining elements need not be set.

On exit: a set of Lagrange multipliers for the bounds on the variables (*reduced costs*) and the general constraints (*shadow costs*). More precisely, the first N elements contain the multipliers for the bounds on the variables, the next NCNLN elements contain the multipliers for the nonlinear constraints $F(x)$ (if any) and the next (M − NCNLN) elements contain the multipliers for the linear constraints Gx and the free row (if any).

22: MINIZ – INTEGER *Output*

On exit: the minimum value of LENIZ required to start solving the problem. If IFAIL $= 12$, E04UGF/E04UGA may be called again with LENIZ suitably larger than MINIZ. (The bigger the better, since it is not certain how much workspace the basis factors need.)

23: MINZ – INTEGER *Output*

On exit: the minimum value of LENZ required to start solving the problem. If IFAIL $= 13$, E04UGF/E04UGA may be called again with LENZ suitably larger than MINZ. (The bigger the better, since it is not certain how much workspace the basis factors need.)

24: NINF – INTEGER *Output*

On exit: the number of constraints that lie outside their bounds by more than the value of the optional parameter Minor Feasibility Tolerance.

If the *linear* constraints are infeasible, the sum of the infeasibilities of the linear constraints is minimized subject to the upper and lower bounds being satisfied. In this case, NINF contains the number of elements of Gx that lie outside their upper or lower bounds. Note that the nonlinear constraints are not evaluated.

Otherwise, the sum of the infeasibilities of the *nonlinear* constraints is minimized subject to the linear constraints and the upper and lower bounds being satisfied. In this case, NINF contains the number of elements of $F(x)$ that lie outside their upper or lower bounds.

25: SINF – ***double precision*** *Output*

On exit: the sum of the infeasibilities of constraints that lie outside their bounds by more than the value of the optional parameter Minor Feasibility Tolerance.

26: OBJ – ***double precision*** *Output*

On exit: the value of the objective function.

27: IZ(LENIZ) – INTEGER array *Workspace*
28: LENIZ – INTEGER *Input*

On entry: the dimension of the array IZ as declared in the (sub)program from which E04UGF/E04UGA is called.

Constraint: LENIZ $\geq \max(500, \mathrm{N} + \mathrm{M})$.

29: Z(LENZ) – ***double precision*** array *Workspace*
30: LENZ – INTEGER *Input*

On entry: the dimension of the array Z as declared in the (sub)program from which E04UGF/E04UGA is called.

Constraint: LENZ ≥ 500.

The amounts of workspace provided (i.e., LENIZ and LENZ) and required (i.e., MINIZ and MINZ) are (by default) output on the current advisory message unit (as defined by X04ABF). Since the minimum values of LENIZ and LENZ required to start solving the problem are returned in MINIZ and MINZ respectively, you may prefer to obtain appropriate values from the output of a preliminary run with LENIZ set to $\max(500, \mathrm{N} + \mathrm{M})$ and/or LENZ set to 500. (E04UGF/E04UGA will then terminate with IFAIL $= 15$ or 16.)

31: IUSER(*) – INTEGER array *User Workspace*

Note: the dimension of the array IUSER must be at least 1.

IUSER is not used by E04UGF/E04UGA, but is passed directly to user-supplied subroutines CONFUN and OBJFUN and may be used to pass information to those routines.

32: RUSER(*) – ***double precision*** array *User Workspace*

Note: the dimension of the array RUSER must be at least 1.

RUSER is not used by E04UGF/E04UGA, but is passed directly to user-supplied subroutines CONFUN and OBJFUN and may be used to pass information to those routines.

33: IFAIL – INTEGER *Input/Output*

Note: *for E04UGA, IFAIL does not occur in this position in the parameter list. See the additional parameters described below.*

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction for details.

On exit: IFAIL $= 0$ unless the routine detects an error (see Section 6).

For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, because for this routine the values of the output parameters 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.**

E04UGF/E04UGA returns with IFAIL $= 0$ if the iterates have converged to a point x that satisfies the first-order Kuhn–Karesh–Tucker conditions (see Section 8.1) to the accuracy requested by the optional parameters Major Feasibility Tolerance (default value $= \sqrt{\epsilon}$) and Major Optimality Tolerance (default value $= \sqrt{\epsilon}$).

Note: *the following are additional parameters for specific use with E04UGA. Users of E04UGF therefore need not read the remainder of this description.*

33: LWSAV(20) – LOGICAL array *Communication Array*
34: IWSAV(550) – INTEGER array *Communication Array*
35: RWSAV(550) – ***double precision*** array *Communication Array*

The arrays LWSAV, IWSAV and RWSAV **must not** be altered between calls to any of the routines E04WBF, E04UGA, E04UHA or E04UJA.

36: IFAIL – INTEGER *Input/Output*

Note: see the parameter description for IFAIL above.

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: E04UGF/E04UGA may return useful information for one or more of the following detected errors or warnings.

Errors or warnings detected by the routine:

IFAIL < 0

A negative value of IFAIL indicates an exit from E04UGF/E04UGA because you set MODE < 0 in OBJFUN or CONFUN. The value of IFAIL will be the same as your setting of MODE.

IFAIL $= 1$

The problem is infeasible. The general constraints cannot all be satisfied simultaneously to within the values of the optional parameters Major Feasibility Tolerance (default value $= \sqrt{\epsilon}$) and Minor Feasibility Tolerance (default value $= \sqrt{\epsilon}$).

IFAIL $= 2$

The problem is unbounded (or badly scaled). The objective function is not bounded below (or above in the case of maximization) in the feasible region because a nonbasic variable can apparently be increased or decreased by an arbitrary amount without causing a basic variable to violate a bound. Add an upper or lower bound to the variable (whose index is printed by default by E04UGF) and rerun E04UGF/E04UGA.

IFAIL $= 3$

The problem may be unbounded. Check that the values of the optional parameters Unbounded Objective (default value $= 10^{15}$) and Unbounded Step Size (default value $= \max(bigbnd, 10^{20})$) are not too small. This exit also implies that the objective function is not bounded below (or above in the case of maximization) in the feasible region defined by expanding the bounds by the value of the optional parameter Violation Limit (default value $= 10.0$).

IFAIL $= 4$

Too many iterations. The values of the optional parameters Major Iteration Limit (default value $= 1000$) and/or Iteration Limit (default value $= 10000$) are too small.

IFAIL $= 5$

Feasible solution found, but requested accuracy could not be achieved. Check that the value of the optional parameter Major Optimality Tolerance (default value $= \sqrt{\epsilon}$) is not too small (say, $< \epsilon$).

IFAIL $= 6$

The value of the optional parameter Superbasics Limit (default value $= \min(500, \bar{n} + 1)$) is too small.

IFAIL $= 7$

An input parameter is invalid.

IFAIL = 8

The user-supplied derivatives of the objective function computed by OBJFUN appear to be incorrect. Check that OBJFUN has been coded correctly and that all relevant elements of the objective gradient have been assigned their correct values.

IFAIL = 9

The user-supplied derivatives of the nonlinear constraint functions computed by CONFUN appear to be incorrect. Check that CONFUN has been coded correctly and that all relevant elements of the nonlinear constraint Jacobian have been assigned their correct values.

IFAIL = 10

The current point cannot be improved upon. Check that OBJFUN and CONFUN have been coded correctly and that they are consistent with the value of the optional parameter Derivative Level (default value = 3).

IFAIL = 11

Numerical error in trying to satisfy the linear constraints (or the linearized nonlinear constraints). The basis is very ill-conditioned.

IFAIL = 12

Not enough integer workspace for the basis factors. Increase LENIZ and rerun E04UGF/E04UGA.

IFAIL = 13

Not enough real workspace for the basis factors. Increase LENZ and rerun E04UGF/E04UGA.

IFAIL = 14

The basis is singular after 15 attempts to factorize it (and adding slacks where necessary). Either the problem is badly scaled or the value of the optional parameter LU Factor Tolerance (default value = 5.0 or 100.0) is too large.

IFAIL = 15

Not enough integer workspace to start solving the problem. Increase LENIZ to at least MINIZ and rerun E04UGF/E04UGA.

IFAIL = 16

Not enough real workspace to start solving the problem. Increase LENZ to at least MINZ and rerun E04UGF/E04UGA.

IFAIL = 17

An unexpected error has occurred. Please contact NAG.

7 Accuracy

If the value of the optional parameter Major Optimality Tolerance is set to 10^{-d} (default value $= \sqrt{\epsilon}$) and IFAIL = 0 on exit, then the final value of $f(x)$ should have approximately d correct significant digits.

8 Further Comments

This section contains a description of the printed output.

8.1 Major Iteration Printout

This section describes the intermediate printout and final printout produced by the major iterations of E04UGF/E04UGA. The intermediate printout is a subset of the monitoring information produced by the routine at every iteration (see Section 12). You can control the level of printed output (see the description of the optional parameter Major Print Level). Note that the intermediate printout and final printout are produced only if Major Print Level ≥ 10 (the default for E04UGF, by default no output is produced by E04UGA).

The following line of summary output (< 80 characters) is produced at every major iteration. In all cases, the values of the quantities printed are those in effect *on completion* of the given iteration.

`Maj` is the major iteration count.

`Mnr` is the number of minor iterations required by the feasibility and optimality phases of the QP subproblem. Generally, `Mnr` will be 1 in the later iterations, since theoretical analysis predicts that the correct active set will be identified near the solution (see Section 10).

Note that `Mnr` may be greater than the optional parameter Minor Iteration Limit if some iterations are required for the feasibility phase.

`Step` is the step α_k taken along the computed search direction. On reasonably well-behaved problems, the unit step (i.e., $\alpha_k = 1$) will be taken as the solution is approached.

`Merit Function` is the value of the augmented Lagrangian merit function (6) at the current iterate. This function will decrease at each iteration unless it was necessary to increase the penalty parameters (see Section 8.1). As the solution is approached, `Merit Function` will converge to the value of the objective function at the solution.

In elastic mode (see Section 10.2) then the merit function is a composite function involving the constraint violations weighted by the value of the optional parameter Elastic Weight.

If there are no nonlinear constraints present then this entry contains `Objective`, the value of the objective function $f(x)$. In this case, $f(x)$ will decrease monotonically to its optimal value.

`Feasibl` is the value of *rowerr*, the largest element of the scaled nonlinear constraint residual vector defined in the description of the optional parameter Major Feasibility Tolerance. The solution is regarded as 'feasible' if `Feasibl` is less than (or equal to) the optional parameter Major Feasibility Tolerance. `Feasibl` will be approximately zero in the neighbourhood of a solution.

If there are no nonlinear constraints present, all iterates are feasible and this entry is not printed.

`Optimal` is the value of *maxgap*, the largest element of the maximum complementarity gap vector defined in the description of the optional parameter Major Optimality Tolerance. The Lagrange multipliers are regarded as 'optimal' if `Optimal` is less than (or equal to) the optional parameter Major Optimality Tolerance. `Optimal` will be approximately zero in the neighbourhood of a solution.

`Cond Hz` is an estimate of the condition number of the reduced Hessian of the Lagrangian (not printed if NCNLN and NONLN are both zero). It is the square of the ratio between the largest and smallest diagonal elements of the upper triangular matrix R. This constitutes a lower bound on the condition number of the matrix $R^{\mathrm{T}}R$ that approximates the reduced Hessian. The larger this number, the more difficult the problem.

`PD` is a two-letter indication of the status of the convergence tests involving the feasibility and optimality of the iterates defined in the descriptions of the optional parameters Major Feasibility Tolerance and Major Optimality Tolerance. Each

letter is `T` if the test is satisfied and `F` otherwise. The tests indicate whether the values of `Feasibl` and `Optimal` are sufficiently small. For example, `TF` or `TT` is printed if there are no nonlinear constraints present (since all iterates are feasible). If either indicator is `F` when E04UGF/E04UGA terminates with IFAIL $= 0$, you should check the solution carefully.

`M` is printed if an extra evaluation of user-supplied subroutines OBJFUN and CONFUN was needed in order to define an acceptable positive-definite quasi-Newton update to the Hessian of the Lagrangian. This modification is only performed when there are nonlinear constraints present.

`m` is printed if, in addition, it was also necessary to modify the update to include an augmented Lagrangian term.

`s` is printed if a self-scaled BFGS (Broyden–Fletcher–Goldfarb–Shanno) update was performed. This update is always used when the Hessian approximation is diagonal and hence always follows a Hessian reset.

`S` is printed if, in addition, it was also necessary to modify the self-scaled update in order to maintain positive-definiteness.

`n` is printed if no positive-definite BFGS update could be found, in which case the approximate Hessian is unchanged from the previous iteration.

`r` is printed if the approximate Hessian was reset after 10 consecutive major iterations in which no BFGS update could be made. The diagonal elements of the approximate Hessian are retained if at least one update has been performed since the last reset. Otherwise, the approximate Hessian is reset to the identity matrix.

`R` is printed if the approximate Hessian has been reset by discarding all but its diagonal elements. This reset will be forced periodically by the values of the optional parameters Hessian Frequency and Hessian Updates. However, it may also be necessary to reset an ill-conditioned Hessian from time to time.

`l` is printed if the change in the norm of the variables was greater than the value defined by the optional parameter Major Step Limit. If this output occurs frequently during later iterations, it may be worthwhile increasing the value of Major Step Limit.

`c` is printed if central differences have been used to compute the unknown elements 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.

`u` is printed if the QP subproblem was unbounded.

`t` is printed if the minor iterations were terminated after the number of iterations specified by the value of the optional parameter Minor Iteration Limit was reached.

`i` is printed if the QP subproblem was infeasible when the routine was not in elastic mode. 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 (8) (see Section 10.2). It is also printed if the minimizer of the elastic subproblem does not satisfy the linearized constraints when the routine is already in elastic mode. (In this case, a feasible point for the usual QP subproblem may or may not exist.)

`w` is printed if a weak solution of the QP subproblem was found.

The final printout includes a listing of the status of every variable and constraint.

The following describes the printout for each variable. A full stop (.) is printed for any numerical value that is zero.

`Variable` gives the name of the variable. If NNAME $= 1$, a default name is assigned to the jth variable for $j = 1, 2, \ldots, n$. If NNAME $=$ N $+$ M, the name supplied in NAMES(j) is assigned to the jth variable.

`State` gives the state of the variable (`LL` if nonbasic on its lower bound, `UL` if nonbasic on its upper bound, `EQ` if nonbasic and fixed, `FR` if nonbasic and strictly between its bounds, `BS` if basic and `SBS` if superbasic).

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` *Alternative optimum possible*. 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 current value, there would be no change in the value of the objective function. The values of the basic and superbasic variables *might* change, giving a genuine alternative solution. The values of the Lagrange multipliers *might* also change.

`D` *Degenerate*. The variable is basic, but it is equal to (or very close to) one of its bounds.

`I` *Infeasible*. The variable is basic and is currently violating one of its bounds by more than the value of the optional parameter Minor Feasibility Tolerance.

`N` *Not precisely optimal*. The variable is nonbasic. Its reduced gradient is larger than the value of the optional parameter Major 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 Mult` is the Lagrange multiplier for the associated bound. This will be zero if `State` is `FR`. If x is optimal, the multiplier should be non-negative if `State` is `LL`, nonpositive if `State` is `UL` and zero if `State` is `BS` or `SBS`.

`Residual` 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 printout for general constraints is the same as that given above for variables, with 'variable' replaced by 'constraint', n replaced by m, NAMES(j) replaced by NAMES$(n + j)$, BL(j) and BU(j) are replaced by BL$(n + j)$ and BU$(n + j)$ respectively. The heading is changed as follows:

`Constrnt` gives the name of the general constraint.

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

8.2 Minor Iteration Printout

This section describes the printout produced by the minor iterations of E04UGF/E04UGA, which involve solving a QP subproblem at every major iteration. (Further details can be found in Section 8.1.) The printout is a subset of the monitoring information produced by the routine at every iteration (see Section 12). You can control the level of printed output (see the description of the optional parameter Minor Print Level). Note that the printout is produced only if Minor Print Level ≥ 1 (default value $= 0$, which produces no output).

The following line of summary output (< 80 characters) is produced at every minor iteration. In all cases, the values of the quantities printed are those in effect *on completion* of the given iteration of the QP subproblem.

`Itn`	is the iteration count.
`Step`	is the step taken along the computed search direction.
`Ninf`	is the number of infeasibilities. This will not increase unless the iterations are in elastic mode. `Ninf` will be zero during the optimality phase.
`Sinf`	is the value of the sum of infeasibilities if `Ninf` is nonzero. This will be zero during the optimality phase.
`Objective`	is the value of the current QP objective function when `Ninf` is zero and the iterations are not in elastic mode. The switch to elastic mode is indicated by a change in the heading to `Composite Obj`.
`Composite Obj`	is the value of the composite objective function (9) when the iterations are in elastic mode. This function will decrease monotonically at each iteration.
`Norm rg`	is the Euclidean norm of the reduced gradient of the QP objective function. During the optimality phase, this norm will be approximately zero after a unit step.

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

9 Example

This is a reformulation of Problem 74 in Hock and Schittkowski (1981) and involves the minimization of the nonlinear function

$$f(x) = 10^{-6}x_3^3 + \tfrac{2}{3} \times 10^{-6}x_4^3 + 3x_3 + 2x_4$$

subject to the bounds

$$\begin{aligned} -0.55 \leq x_1 &\leq 0.55, \\ -0.55 \leq x_2 &\leq 0.55, \\ 0 \leq x_3 &\leq 1200, \\ 0 \leq x_4 &\leq 1200, \end{aligned}$$

to the nonlinear constraints

$$\begin{aligned} 1000\sin(-x_1 - 0.25) &+ 1000\sin(-x_2 - 0.25) - x_3 &&= -894.8, \\ 1000\sin(x_1 - 0.25) &+ 1000\sin(x_1 - x_2 - 0.25) - x_4 &&= -894.8, \\ 1000\sin(x_2 - 0.25) &+ 1000\sin(x_2 - x_1 - 0.25) &&= -1294.8, \end{aligned}$$

and to the linear constraints

$$\begin{aligned} -x_1 + x_2 &\geq -0.55, \\ x_1 - x_2 &\geq -0.55. \end{aligned}$$

The initial point, which is infeasible, is

$$x_0 = (\,0, \quad 0, \quad 0, \quad 0\,)^{\mathrm{T}},$$

and $f(x_0) = 0$.

The optimal solution (to five figures) is

$$x^* = (0.11887, -0.39623, 679.94, 1026.0)^{\mathrm{T}},$$

and $f(x^*) = 5126.4$. All the nonlinear constraints are active at the solution.

The document for E04UHF/E04UHA includes an example program to solve Problem 45 from Hock and Schittkowski (1981) using some of the optional parameters described in Section 11.

9.1 Program Text

Note: *the following program illustrates the use of E04UGF. An equivalent program illustrating the use of E04UGA is available with the supplied Library and is also available from the NAG web site.*

```
*     E04UGF Example Program Text
*     Mark 20 Revised. NAG Copyright 2001.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          IDUMMY
      PARAMETER        (IDUMMY=-11111)
      INTEGER          MMAX, NMAX, NNZMAX, LENIZ, LENZ
      PARAMETER        (MMAX=100,NMAX=100,NNZMAX=300,LENIZ=5000,
     +                 LENZ=5000)
*     .. Local Scalars ..
      DOUBLE PRECISION OBJ, SINF
      INTEGER          I, ICOL, IFAIL, IOBJ, J, JCOL, M, MINIZ, MINZ, N,
     +                 NCNLN, NINF, NJNLN, NNAME, NNZ, NONLN, NS
      CHARACTER        START
*     .. Local Arrays ..
      DOUBLE PRECISION A(NNZMAX), BL(NMAX+MMAX), BU(NMAX+MMAX),
     +                 CLAMDA(NMAX+MMAX), RUSER(1), XS(NMAX+MMAX),
     +                 Z(LENZ)
      INTEGER          HA(NNZMAX), ISTATE(NMAX+MMAX), IUSER(1),
     +                 IZ(LENIZ), KA(NMAX+1)
      CHARACTER*8      NAMES(NMAX+MMAX)
*     .. External Subroutines ..
      EXTERNAL         CONFUN, E04UGF, OBJFUN
*     .. Executable Statements ..
      WRITE (NOUT,*) 'E04UGF Example Program Results'
*     Skip heading in data file.
      READ (NIN,*)
      READ (NIN,*) N, M
      IF (N.LE.NMAX .AND. M.LE.MMAX) THEN
*
*        Read NCNLN, NONLN and NJNLN from data file.
*
         READ (NIN,*) NCNLN, NONLN, NJNLN
*
*        Read NNZ, IOBJ, START, NNAME and NAMES from data file.
*
         READ (NIN,*) NNZ, IOBJ, START, NNAME
         IF (NNAME.EQ.N+M) READ (NIN,*) (NAMES(I),I=1,N+M)
*
*        Initialize KA.
*
         DO 20 I = 1, N + 1
            KA(I) = IDUMMY
   20    CONTINUE
*
*        Read the matrix A from data file. Set up KA.
*
         JCOL = 1
         KA(JCOL) = 1
         DO 60 I = 1, NNZ
*
*           Element ( HA( I ), ICOL ) is stored in A( I ).
*
            READ (NIN,*) A(I), HA(I), ICOL
*
            IF (ICOL.LT.JCOL) THEN
*
*              Elements not ordered by increasing column index.
*
               WRITE (NOUT,99999) 'Element in column', ICOL,
     +           ' found after element in column', JCOL, '. Problem',
     +           ' abandoned.'
               GO TO 100
            ELSE IF (ICOL.EQ.JCOL+1) THEN
*
```

```
*              Index in A of the start of the ICOL-th column equals I.
*
               KA(ICOL) = I
               JCOL = ICOL
            ELSE IF (ICOL.GT.JCOL+1) THEN
*
*              Index in A of the start of the ICOL-th column equals I,
*              but columns JCOL+1,JCOL+2,...,ICOL-1 are empty. Set the
*              corresponding elements of KA to I.
*
               DO 40 J = JCOL + 1, ICOL - 1
                  KA(J) = I
   40          CONTINUE
               KA(ICOL) = I
               JCOL = ICOL
            END IF
   60    CONTINUE
*
         KA(N+1) = NNZ + 1
*
         IF (N.GT.ICOL) THEN
*
*           Columns N,N-1,...,ICOL+1 are empty. Set the corresponding
*           elements of KA accordingly.
*
            DO 80 I = N, ICOL + 1, -1
               IF (KA(I).EQ.IDUMMY) KA(I) = KA(I+1)
   80       CONTINUE
         END IF
*
*        Read BL, BU, ISTATE, XS and CLAMDA from data file.
*
         READ (NIN,*) (BL(I),I=1,N+M)
         READ (NIN,*) (BU(I),I=1,N+M)
         IF (START.EQ.'C') THEN
            READ (NIN,*) (ISTATE(I),I=1,N)
         ELSE IF (START.EQ.'W') THEN
            READ (NIN,*) (ISTATE(I),I=1,N+M)
         END IF
         READ (NIN,*) (XS(I),I=1,N)
         IF (NCNLN.GT.0) READ (NIN,*) (CLAMDA(I),I=N+1,N+NCNLN)
*
*        Solve the problem.
*
         IFAIL = 1
*
         CALL E04UGF(CONFUN,OBJFUN,N,M,NCNLN,NONLN,NJNLN,IOBJ,NNZ,A,HA,
     +               KA,BL,BU,START,NNAME,NAMES,NS,XS,ISTATE,CLAMDA,
     +               MINIZ,MINZ,NINF,SINF,OBJ,IZ,LENIZ,Z,LENZ,IUSER,
     +               RUSER,IFAIL)
*
         IF (IFAIL.NE.0) THEN
            WRITE (NOUT,99998) IFAIL
         END IF
*
      ELSE
         WRITE (NOUT,99997)
      END IF
*
  100 CONTINUE
*
99999 FORMAT (/1X,A,I5,A,I5,A,A)
99998 FORMAT (1X,/1X,' ** E04UGF returned with IFAIL = ',I5)
99997 FORMAT (1X,' At least one of N or NCLIN is too large')
      END
*
      SUBROUTINE CONFUN(MODE,NCNLN,NJNLN,NNZJAC,X,F,FJAC,NSTATE,IUSER,
     +                  RUSER)
*     Computes the nonlinear constraint functions and their Jacobian.
*     .. Scalar Arguments ..
      INTEGER          MODE, NCNLN, NJNLN, NNZJAC, NSTATE
```

```
*     .. Array Arguments ..
      DOUBLE PRECISION F(NCNLN), FJAC(NNZJAC), RUSER(*), X(NJNLN)
      INTEGER          IUSER(*)
*     .. Intrinsic Functions ..
      INTRINSIC        COS, SIN
*     .. Executable Statements ..
      IF (MODE.EQ.0 .OR. MODE.EQ.2) THEN
         F(1) = 1000.0D+0*SIN(-X(1)-0.25D+0) + 1000.0D+0*SIN(-X(2)
     +          -0.25D+0)
         F(2) = 1000.0D+0*SIN(X(1)-0.25D+0) + 1000.0D+0*SIN(X(1)-X(2)
     +          -0.25D+0)
         F(3) = 1000.0D+0*SIN(X(2)-X(1)-0.25D+0) + 1000.0D+0*SIN(X(2)
     +          -0.25D+0)
      END IF
*
      IF (MODE.EQ.1 .OR. MODE.EQ.2) THEN
*
*        Nonlinear Jacobian elements for column 1.
*
         FJAC(1) = -1000.0D+0*COS(-X(1)-0.25D+0)
         FJAC(2) = 1000.0D+0*COS(X(1)-0.25D+0) + 1000.0D+0*COS(X(1)-X(2)
     +             -0.25D+0)
         FJAC(3) = -1000.0D+0*COS(X(2)-X(1)-0.25D+0)
*
*        Nonlinear Jacobian elements for column 2.
*
         FJAC(4) = -1000.0D+0*COS(-X(2)-0.25D+0)
         FJAC(5) = -1000.0D+0*COS(X(1)-X(2)-0.25D+0)
         FJAC(6) = 1000.0D+0*COS(X(2)-X(1)-0.25D+0) + 1000.0D+0*COS(X(2)
     +             -0.25D+0)
      END IF
*
      END
*
      SUBROUTINE OBJFUN(MODE,NONLN,X,OBJF,OBJGRD,NSTATE,IUSER,RUSER)
*     Computes the nonlinear part of the objective function and its
*     gradient
*     .. Scalar Arguments ..
      DOUBLE PRECISION OBJF
      INTEGER          MODE, NONLN, NSTATE
*     .. Array Arguments ..
      DOUBLE PRECISION OBJGRD(NONLN), RUSER(*), X(NONLN)
      INTEGER          IUSER(*)
*     .. Executable Statements ..
      IF (MODE.EQ.0 .OR. MODE.EQ.2) OBJF = 1.0D-6*X(3)**3 + 2.0D-6*X(4)
     +    **3/3.0D+0
*
      IF (MODE.EQ.1 .OR. MODE.EQ.2) THEN
         OBJGRD(1) = 0.0D+0
         OBJGRD(2) = 0.0D+0
         OBJGRD(3) = 3.0D-6*X(3)**2
         OBJGRD(4) = 2.0D-6*X(4)**2
      END IF
*
      END
```

9.2 Program Data

```
E04UGF Example Program Data
 4   6              :Values of N and M
 3   4   2          :Values of NCNLN, NONLN and NJNLN
14   6  'C'  10     :Values of NNZ, IOBJ, START and NNAME
'Varble 1'  'Varble 2'  'Varble 3'  'Varble 4'  'NlnCon 1'
'NlnCon 2'  'NlnCon 3'  'LinCon 1'  'LinCon 2'  'Free Row' :End of NAMES
 1.0E+25   1   1
 1.0E+25   2   1
 1.0E+25   3   1
     1.0   5   1
    -1.0   4   1
 1.0E+25   1   2
```

```
 1.0E+25   2   2
 1.0E+25   3   2
    -1.0   5   2
     1.0   4   2
     3.0   6   3
    -1.0   1   3
    -1.0   2   4
     2.0   6   4    :End of matrix A
-0.55     -0.55      0.0       0.0     -894.8   -894.8   -1294.8  -0.55
-0.55     -1.0E+25  :End of BL
 0.55      0.55      1200.0    1200.0  -894.8   -894.8   -1294.8   1.0E+25
 1.0E+25   1.0E+25  :End of BU
 0    0    0    0   :End of ISTATE
 0.0  0.0  0.0  0.0 :End of XS
 0.0  0.0  0.0      :End of CLAMDA
```

9.3 Program Results

```
 E04UGF Example Program Results

 *** E04UGF

 Parameters
 ----------

 Frequencies.
 Check frequency.........        60       Expand frequency.......     10000
 Factorization frequency.        50

 QP subproblems.
 Scale tolerance.........  9.00E-01       Minor feasibility tol..  1.05E-08
 Scale option............         1       Minor optimality tol...  1.05E-08
 Partial price...........         1       Crash tolerance........  1.00E-01
 Pivot tolerance.........  2.04E-11       Minor print level......         0
 Crash option............         0       Elastic weight.........  1.00E+02

 The SQP method.
 Minimize................
 Nonlinear objective vars         4       Major optimality tol...  1.05E-08
 Function precision......  1.72E-13       Unbounded step size....  1.00E+20
 Superbasics limit.......         4       Forward difference int.  4.15E-07
 Unbounded objective.....  1.00E+15       Central difference int.  5.56E-05
 Major step limit........  2.00E+00       Derivative linesearch..
 Derivative level........         3       Major iteration limit..      1000
 Linesearch tolerance....  9.00E-01       Verify level...........         0
 Minor iteration limit...       500       Major print level......        10
 Infinite bound size.....  1.00E+20       Iteration limit........     10000

 Hessian approximation.
 Hessian full memory.....                 Hessian updates........  99999999
 Hessian frequency.......  99999999

 Nonlinear constraints.
 Nonlinear constraints...         3       Major feasibility tol..  1.05E-08
 Nonlinear Jacobian vars.         2       Violation limit........  1.00E+01

 Miscellaneous.
 Variables...............         4       Linear constraints.....         3
 Nonlinear variables.....         4       Linear variables.......         0
 LU factor tolerance.....  5.00E+00       LU singularity tol.....  2.04E-11
 LU update tolerance.....  5.00E+00       LU density tolerance...  6.00E-01
 eps (machine precision).  1.11E-16       Monitoring file........        -1
 COLD start..............                 Infeasible exit........

 Workspace provided is                 IZ(    5000),  Z(    5000).
 To start solving the problem we need  IZ(     628),  Z(     758).

 Itn      0 -- Scale option reduced from               1 to               0.

 Itn      0 -- Feasible linear rows.
```

```
Itn      0 -- Norm(x-x0) minimized. Sum of infeasibilities = 0.00E+00.

confun  sets       6   out of       6   constraint gradients.
objfun  sets       4   out of       4   objective  gradients.

Cheap test on confun...

The Jacobian seems to be OK.

The largest discrepancy was    4.41E-08  in constraint     2.

Cheap test on objfun...

The objective gradients seem to be OK.
Gradient projected in two directions   0.00000000000E+00   0.00000000000E+00
Difference approximations              1.74111992322E-19   4.48742248252E-21

Itn      0 -- All-slack basis B = I selected.

Itn      7 -- Large multipliers.
              Elastic mode started with weight =  2.0E+02.


 Maj  Mnr    Step Merit Function Feasibl Optimal Cond Hz PD
   0   12 0.0E+00   3.199952E+05 1.7E+00 8.0E-01 2.1E+06 FF  R   i
   1    2 1.0E+00   2.463016E+05 1.2E+00 3.2E+03 4.5E+00 FF s
   2    1 1.0E+00   1.001802E+04 3.3E-02 9.2E+01 4.5E+00 FF
   3    1 1.0E+00   5.253418E+03 6.6E-04 2.5E+01 4.8E+00 FF
   4    1 1.0E+00   5.239444E+03 2.0E-06 2.8E+01 1.0E+02 FF
   5    1 1.0E+00   5.126208E+03 6.0E-04 5.9E-01 1.1E+02 FF
   6    1 1.0E+00   5.126498E+03 4.7E-07 2.9E-02 1.0E+02 FF
   7    1 1.0E+00   5.126498E+03 5.9E-10 1.5E-03 1.1E+02 TF
   8    1 1.0E+00   5.126498E+03 1.2E-12 7.6E-09 1.1E+02 TT

Exit from NP problem after     8 major iterations,
                              21 minor iterations.

Variable State     Value       Lower Bound  Upper Bound   Lagr Mult  Residual

Varble 1    BS   0.118876     -0.55000      0.55000     -1.2529E-07  0.4311
Varble 2    BS  -0.396234     -0.55000      0.55000      1.9243E-08  0.1538
Varble 3    BS    679.945           .        1200.0      1.7001E-10   520.1
Varble 4   SBS    1026.07           .        1200.0     -2.1918E-10   173.9

Constrnt State     Value       Lower Bound  Upper Bound   Lagr Mult  Residual

NlnCon 1    EQ   -894.800      -894.80      -894.80      -4.387      3.3644E-09
NlnCon 2    EQ   -894.800      -894.80      -894.80      -4.106      6.0049E-10
NlnCon 3    EQ   -1294.80      -1294.8      -1294.8      -5.463      3.3551E-09
LinCon 1    BS  -0.515110     -0.55000          None           .      3.4890E-02
LinCon 2    BS   0.515110     -0.55000          None           .       1.065
Free Row    BS    4091.97          None          None     -1.000       4092.

Exit E04UGF - Optimal solution found.

Final objective value =     5126.498
```

Note: *the remainder of this document is intended for more advanced users. Section 10 contains a detailed description of the algorithm which may be needed in order to understand Sections 11 and 12. Section 11 describes the optional parameters which may be set by calls to E04UHF/E04UHA and/or E04UJF/E04UJA. Section 12 describes the quantities which can be requested to monitor the course of the computation*.

10 Algorithmic Details

This section contains a detailed description of the method used by E04UGF/E04UGA.

10.1 Overview

Here we briefly summarize the main features of the method and introduce some terminology. Where possible, explicit reference is made to the names of variables that are parameters of the routine or appear in the printed output. Further details can be found in Gill *et al.* (2002).

At a solution of (1), some of the constraints will be *active*, i.e., satisfied exactly. Let

$$r(x) = \begin{pmatrix} x \\ F(x) \\ Gx \end{pmatrix}$$

and $\mathcal{G}$ denote the set of indices of $r(x)$ corresponding to active constraints at an arbitrary point x. Let $r_j'(x)$ denote the usual *derivative* of $r_j(x)$, which is the row vector of first partial derivatives of $r_j(x)$ (see Ortega and Rheinboldt (1970)). The vector $r_j'(x)$ comprises the jth row of $r'(x)$ so that

$$r'(x) = \begin{pmatrix} I \\ J(x) \\ G \end{pmatrix},$$

where $J(x)$ is the Jacobian of $F(x)$.

A point x is a *first-order Kuhn–Karesh–Tucker (KKT) point* for (1) (see Powell (1974)) if the following conditions hold:

(a) x is feasible;

(b) there exists a vector λ (*the Lagrange multiplier vector for the bound and general constraints*) such that

$$g(x) = r'(x)^{\mathrm{T}}\lambda = \begin{pmatrix} I & J(x)^{\mathrm{T}} & G^{\mathrm{T}} \end{pmatrix}\lambda, \tag{4}$$

where g is the gradient of f evaluated at x;

(c) the Lagrange multiplier λ_j associated with the jth constraint satisfies $\lambda_j = 0$ if $l_j < r_j(x) < u_j$; $\lambda_j \geq 0$ if $l_j = r_j(x)$; $\lambda_j \leq 0$ if $r_j(x) = u_j$; and λ_j can have any value if $l_j = u_j$.

An equivalent statement of the condition (4) is

$$Z^{\mathrm{T}}g(x) = 0,$$

where Z is a matrix defined as follows. Consider the set N of vectors orthogonal to the gradients of the active constraints, i.e.,

$$N = \{z \quad | \quad r_j'(x)z = 0 \quad \text{for all} \quad j \in \mathcal{G}\}.$$

The columns of Z may then be taken as any basis for the vector space N. The vector $Z^{\mathrm{T}}g$ is termed the *reduced gradient* of f at x. Certain additional conditions must be satisfied in order for a first-order KKT point to be a solution of (1) (see Powell (1974)).

The basic structure of E04UGF/E04UGA 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 x^* that satisfies the first-order KKT optimality conditions. At each iterate 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 $Gx - s_L = 0$ and the nonlinear constraint linearization

$$F(x_k) + F'(x_k)(x - x_k) - s_N = 0,$$

where $F'(x_k)$ denotes the *Jacobian matrix*, whose rows are the first partial derivatives of $F(x)$ evaluated at the point x_k. The QP constraints therefore comprise the m linear constraints

$$\begin{aligned} F'(x_k)x - s_N &= -F(x_k) + F'(x_k)x_k, \\ Gx - s_L &= 0, \end{aligned}$$

where x and $s = (s_N, s_L)^{\mathrm{T}}$ are bounded above and below by u and l as before. If the m by n matrix A and m element vector b are defined as

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

then the QP subproblem can be written as

$$\underset{x,s}{\text{minimize}}\, q(x) \quad \text{subject to} \quad Ax - s = b, \quad l \leq \begin{Bmatrix} x \\ s \end{Bmatrix} \leq u, \tag{5}$$

where $q(x)$ is a quadratic approximation to a modified Lagrangian function (see Gill *et al.* (2002)).

The linear constraint matrix A is stored in the arrays A, HA and KA (see Section 5). This allows you to specify the sparsity pattern of nonzero elements in $F'(x)$ and G and to identify any nonzero elements that remain constant throughout the minimization.

Solving the QP subproblem is itself an iterative procedure, with the *minor* iterations of an SQP method being the iterations of the QP method. At each minor iteration, the constraints $Ax - s = b$ are (conceptually) partitioned into the form

$$Bx_B + Sx_S + Nx_N = b,$$

where the *basis matrix* B is square and nonsingular. The elements of x_B, x_S and x_N are called the *basic*, *superbasic* and *nonbasic* variables respectively; they are a permutation of the elements of x and s. At a QP solution, the basic and superbasic variables will lie somewhere between their bounds, while the nonbasic variables will be equal to one of their upper or lower bounds. At each minor 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 function $q(x)$ or sum of infeasibilities (as appropriate). 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 if there are no nonlinear constraints in (1) and $f(x)$ is linear.

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 decreased by one.

Associated with each of the m equality constraints $Ax - s = b$ is a *dual variable* π_i. Similarly, each variable in (x, s) has an associated *reduced gradient* d_j (also known as a *reduced cost*). The reduced gradients for the variables x are the quantities $g - A^{\mathrm{T}}\pi$, where g is the gradient of the QP objective function $q(x)$; the reduced gradients for the slack variables s are the dual variables π. The QP subproblem (5) 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 is found by slightly relaxing these conditions on d_j (see the description of the optional parameter Minor Optimality Tolerance).

After a QP subproblem has been solved, new estimates of the solution to (1) are computed using a linesearch on the augmented Lagrangian merit function

$$\mathcal{M}(x, s, \pi) = f(x) - \pi^{\mathrm{T}}(F(x) - s_N) + \tfrac{1}{2}(F(x) - s_N)^{\mathrm{T}} D (F(x) - s_N), \tag{6}$$

where D is a diagonal matrix of penalty parameters. If (x_k, s_k, π_k) denotes the current estimate of the solution and $(\hat{x}, \hat{s}, \hat{\pi})$ denotes the optimal QP solution, the linesearch determines a step α_k (where

$0 < \alpha_k \leq 1$) such that the new point

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

produces a *sufficient decrease* in the merit function (6). When necessary, the penalties in D are increased by the minimum-norm perturbation that ensures descent for $\mathcal{M}$ (see Gill *et al.* (1992)). As in E04WDF, s_N is adjusted to minimize the merit function as a function of s before the solution of the QP subproblem. Further details can be found in Eldersveld (1991) and Gill *et al.* (1986).

10.2 Treatment of Constraint Infeasibilities

E04UGF/E04UGA makes explicit allowance for infeasible constraints. Infeasible linear constraints are detected first by solving a problem of the form

$$\underset{x,v,w}{\text{minimize}}\, e^{\mathrm{T}}(v+w) \qquad \text{subject to} \qquad l \leq \begin{Bmatrix} x \\ Gx - v + w \end{Bmatrix} \leq u, \qquad v \geq 0, \qquad w \geq 0, \tag{7}$$

where $e = (1, 1, \ldots, 1)^{\mathrm{T}}$. This is equivalent to minimizing the sum of the general linear constraint violations subject to the simple bounds. (In the linear programming literature, the approach is often called *elastic programming*.)

If the linear constraints are infeasible (i.e., $v \neq 0$ or $w \neq 0$), the routine terminates without computing the nonlinear functions.

If the linear constraints are feasible, all subsequent iterates will satisfy the linear constraints. (Such a strategy allows linear constraints to be used to define a region in which $f(x)$ and $F(x)$ can be safely evaluated.) The routine then proceeds to solve (1) as given, using search directions obtained from a sequence of QP subproblems (5). Each QP subproblem minimizes a quadratic model of a certain Lagrangian function subject to linearized constraints. An augmented Lagrangian merit function (6) is reduced along each search direction to ensure convergence from any starting point.

The routine enters 'elastic' mode if the QP subproblem proves to be infeasible or unbounded (or if the dual variables π for the nonlinear constraints become 'large') by solving a problem of the form

$$\underset{x,v,w}{\text{minimize}}\, \bar{f}(x,v,w) \qquad \text{subject to} \qquad l \leq \begin{Bmatrix} x \\ F(x) - v + w \\ Gx \end{Bmatrix} \leq u, \qquad v \geq 0, \qquad w \geq 0, \tag{8}$$

where

$$\bar{f}(x,v,w) = f(x) + \gamma e^{\mathrm{T}}(v+w) \tag{9}$$

is called a *composite objective* and γ is a non-negative parameter (the *elastic weight*). If γ is sufficiently large, this is equivalent to minimizing the sum of the nonlinear constraint violations subject to the linear constraints and bounds. A similar l_1 formulation of (1) is fundamental to the Sl_1QP algorithm of Fletcher (1984). See also Conn (1973).

11 Optional Parameters

Several optional parameters in E04UGF/E04UGA define choices in the problem specification or the algorithm logic. In order to reduce the number of formal parameters of E04UGF/E04UGA 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. A complete list of optional parameters and their default values is given in Section 11.1.

Optional parameters may be specified by calling one, or both, of the routines E04UHF/E04UHA and E04UJF/E04UJA before a call to E04UGF/E04UGA.

E04UHF/E04UHA 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 E04UHF (IOPTNS, INFORM)
```

can then be used to read the file on unit IOPTNS. INFORM will be zero on successful exit. E04UHF/E04UHA should be consulted for a full description of this method of supplying optional parameters.

E04UJF/E04UJA can be called to supply options directly, one call being necessary for each optional parameter. For example,

```
CALL E04UJF ('Print Level = 5')
```

E04UJF/E04UJA 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 E04UGF/E04UGA (unless they define invalid values) and so remain in effect for subsequent calls to E04UGF/E04UGA from the calling program (unless altered by you).

11.1 Optional Parameter Checklist and Default Values

The following list gives the valid options. For each option, we give the keyword, any essential optional qualifiers and the default value. A definition for each option can be found in Section 11.2, where the minimum abbreviation of each keyword is underlined (if no characters of an optional qualifier are underlined, the qualifier may be omitted), the letters i and r denote INTEGER and ***double precision*** values required with certain options, and the default value of an option is used whenever the condition $|i| \geq 100000000$ is satisfied.

Optional Parameter	**Default Value**
Central Difference Interval	Default $= \sqrt[3]{\text{Function Precision}}$
Check Frequency	Default $= 60$
Crash Option	Default $= 0$ or 3
Crash Tolerance	Default $= 0.1$
Defaults	
Derivative Level	Default $= 3$
Derivative Linesearch	Default
Elastic Weight	Default $= 1.0$ or 100.0
Expand Frequency	Default $= 10000$
Factorization Frequency	Default $= 50$ or 100
Feasibility Tolerance	See Minor Feasibility Tolerance
Feasible Exit	See Infeasible Exit
Feasible Point	See Minimize
Forward Difference Interval	Default $= \sqrt{\text{Function Precision}}$
Function Precision	Default $= \epsilon^{0.8}$
Hessian Frequency	Default $= 99999999$
Hessian Full Memory	Default when $\bar{n} < 75$
Hessian Limited Memory	Default when $\bar{n} \geq 75$. See Hessian Full Memory.
Hessian Updates	Default $= 20$ or 99999999
Infeasible Exit	Default
Infinite Bound Size	Default $= 10^{20}$
Iteration Limit	Default $= 10000$
Linesearch Tolerance	Default $= 0.9$
List	Default for E04UGF $=$ List
LU Density Tolerance	Default $= 0.6$
LU Factor Tolerance	Default $= 5.0$ or 100.0
LU Singularity Tolerance	Default $= \epsilon^{0.67}$. See LU Density Tolerance.

LU Update Tolerance	Default $= 5.0$ or 10.0. See LU Factor Tolerance.
Major Feasibility Tolerance	Default $= \sqrt{\epsilon}$
Major Iteration Limit	Default $= 1000$
Major Optimality Tolerance	Default $= \sqrt{\epsilon}$
Major Print Level	Default for E04UGF $= 10$ Default for E04UGA $= 0$
Major Step Limit	Default $= 2.0$
Maximize	See Minimize
Minimize	Default
Minor Feasibility Tolerance	Default $= \sqrt{\epsilon}$
Minor Iteration Limit	Default $= 500$
Minor Optimality Tolerance	Default $= \sqrt{\epsilon}$
Minor Print Level	Default $= 0$
Monitoring File	Default $= -1$
Nolist	Default for E04UGA = Nolist. See List.
Nonderivative Linesearch	See Derivative Linesearch
Optimality Tolerance	See Major Optimality Tolerance
Partial Price	Default $= 1$ or 10
Pivot Tolerance	Default $= \epsilon^{0.67}$
Print Level	See Major Print Level
Scale Option	Default $= 1$ or 2
Scale Tolerance	Default $= 0.9$
Start Constraint Check At Column	Default $= 1$. See Start Objective Check At Column.
Start Objective Check At Column	Default $= 1$
Stop Constraint Check At Column	Default $= n_1''$. See Start Objective Check At Column.
Stop Objective Check At Column	Default $= n_1'$. See Start Objective Check At Column.
Superbasics Limit	Default $= \min(500, \bar{n} + 1)$
Unbounded Objective	Default $= 10^{15}$
Unbounded Step Size	Default $= \max(bigbnd, 10^{20})$. See Unbounded Objective.
Verify Level	Default $= 0$
Violation Limit	Default $= 10.0$

11.2 Description of the Optional Parameters

Central Difference Interval r Default $= \sqrt[3]{\text{Function Precision}}$

Note that this option does not apply when Derivative Level $= 3$.

The value of r is used near an optimal solution in order to obtain more accurate (but more expensive) estimates of gradients. This requires twice as many function evaluations as compared to using forward differences (see optional parameter Forward Difference Interval). The interval used for the jth variable is $h_j = r\left(1 + \left|x_j\right|\right)$. The resulting gradient estimates should be accurate to $O\left(r^2\right)$, unless the functions are badly scaled. The switch to central differences is indicated by `c` at the end of each line of intermediate printout produced by the major iterations (see Section 8.1). See Gill *et al.* (1981) for a discussion of the accuracy in finite difference approximations.

If $r \leq 0$, the default value is used.

Check Frequency i Default $= 60$

Every ith minor iteration after the most recent basis factorization, a numerical test is made to see if the current solution (x, s) satisfies the general linear constraints (including any linearized nonlinear constraints). The constraints are of the form $Ax - s = b$, where s is the set of slack variables. If the largest element of the residual vector $r = b - Ax + s$ is judged to be too large, the current basis is refactorized and the basic variables recomputed to satisfy the general constraints more accurately.

If $i < 0$, the default value is used. If $i = 0$, the value $i = 99999999$ is used and effectively no checks are made.

Crash Option i Default $= 0$ or 3

The default value of i is 0 if there are any nonlinear constraints and 3 otherwise. Note that this option does not apply when START = 'W' (see Section 5).

If START = 'C', an internal Crash procedure is used to select an initial basis from various rows and columns of the constraint matrix $(A \quad -I)$. The value of i determines which rows and columns of A are initially eligible for the basis and how many times the Crash procedure is called. Columns of $-I$ are used to pad the basis where necessary. The possible choices for i are the following.

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 any 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 (whilst retaining the current basis for linear rows).
3	The Crash procedure is called up to three times (if there are any nonlinear constraints). The first two calls treat linear *equality* constraints and linear *inequality* constraints separately. 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 (whilst retaining the current basis for linear rows).

If $i < 0$ or $i > 3$, the default value is used.

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.

Crash Tolerance r Default $= 0.1$

The value r $(0 \leq r < 1)$ allows the Crash procedure to ignore certain 'small' nonzero elements in the columns of A while searching for a triangular basis. If $a_{\max}$ is the largest element in the jth column, other nonzeros a_{ij} in the column are ignored if $|a_{ij}| \leq a_{\max} \times r$.

When $r > 0$, 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 earlier on some problems.

If $r < 0$ or $r \geq 1$, the default value is used.

Defaults

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

Derivative Level i Default $= 3$

This parameter indicates which nonlinear function gradients are provided in user-supplied subroutines OBJFUN and CONFUN. The possible choices for i are the following.

i	**Meaning**
3	All elements of the objective gradient and the constraint Jacobian are provided.
2	All elements of the constraint Jacobian are provided, but some (or all) elements of the objective gradient are not specified.
1	All elements of the objective gradient are provided, but some (or all) elements of the constraint Jacobian are not specified.
0	Some (or all) elements of both the objective gradient and the constraint Jacobian are not specified.

The default 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, E04UGF/E04UGA will *estimate* the unspecified elements of the objective gradient, using finite differences. This may simplify the coding of OBJFUN. However, the computation of finite difference approximations usually increases the total run-time substantially (since a call to OBJFUN is required for each unspecified element) and there is less assurance that an acceptable solution will be found.

If $i = 0$ or 1, E04UGF/E04UGA will approximate unspecified elements of the constraint Jacobian. For each column of the Jacobian, one call to CONFUN is needed to estimate all unspecified elements in that column (if any). For example, if the sparsity pattern of the Jacobian has the form

$$\begin{pmatrix} * & * & & * \\ & ? & ? & \\ * & & ? & \\ & * & & * \end{pmatrix}$$

where '$*$' indicates an element provided by you and '?' indicates an unspecified element, E04UGF/E04UGA will call CONFUN twice: once to estimate the missing element in column 2 and again to estimate the two missing elements in column 3. (Since columns 1 and 4 are known, they require no calls to CONFUN.)

At times, central differences are used rather than forward differences, in which case twice as many calls to OBJFUN and CONFUN are needed. (The switch to central differences is not under your control.)

If $i < 0$ or $i > 3$, the default value is used.

Derivative Linesearch Default
Nonderivative Linesearch

At each major iteration, a linesearch is used to improve the value of the Lagrangian merit function (6). The default linesearch uses safeguarded cubic interpolation and requires both function and gradient values in order to compute estimates of the step α_k. If some analytic derivatives are not provided or optional parameter Nonderivative Linesearch is specified, a linesearch based upon safeguarded quadratic interpolation (which does not require the evaluation or approximation of any gradients) is used instead.

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

If optional parameter Nonderivative Linesearch is selected, E04UGF/E04UGA signals the evaluation of the linesearch by calling OBJFUN and CONFUN with MODE $= 0$. Once the linesearch is complete, the nonlinear functions are re-evaluated with MODE $= 2$. If the potential savings offered by a nonderivative linesearch are to be fully realized, it is essential that OBJFUN and CONFUN be coded so that no derivatives are computed when MODE $= 0$.

Elastic Weight r Default $= 1.0$ or 100.0

The default value of r is 100.0 if there are any nonlinear constraints and 1.0 otherwise.

This option defines the initial weight γ associated with problem (8).

At any given major iteration k, elastic mode is entered if the QP subproblem is infeasible or the QP dual variables (Lagrange multipliers) are larger in magnitude than $r \times \left(1 + \|g(x_k)\|_2\right)$, where g is the objective gradient. In either case, the QP subproblem is resolved in elastic mode with $\gamma = r \times \left(1 + \|g(x_k)\|_2\right)$.

Thereafter, γ is increased (subject to a maximum allowable value) at any point that is optimal for problem (8), but not feasible for problem (1). After the pth increase, $\gamma = r \times 10^p \times \left(1 + \left\|g\left(x_{k_1}\right)\right\|_2\right)$, where x_{k_1} is the iterate at which γ was first needed.

If $r < 0$, the default value is used.

Expand Frequency i Default $= 10000$

This option is part of the EXPAND anti-cycling procedure due to Gill *et al.* (1989), which is 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 constraints by a small amount. Suppose that the value of optional parameter Minor Feasibility Tolerance is δ. Over a period of i iterations, the feasibility tolerance actually used by E04UGF/E04UGA (i.e., the *working* feasibility tolerance) 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 only occur 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 the value of i helps reduce the number of slightly infeasible nonbasic basic variables (most of which are eliminated during the resetting procedure). However, it also diminishes the freedom to choose a large pivot element (see optional parameter Pivot Tolerance).

If $i < 0$, the default value is used. If $i = 0$, the value $i = 99999999$ is used and effectively no anti-cycling procedure is invoked.

Factorization Frequency i Default $= 50$ or 100

The default value of i is 50 if there are any nonlinear constraints and 100 otherwise.

If $i > 0$, at most i basis changes will occur between factorizations of the basis matrix.

For linear problems, the basis factors are usually updated at every iteration. The default value $i = 100$ is reasonable for typical problems, particularly those that are extremely sparse and well-scaled.

When the objective function is nonlinear, fewer basis updates will occur as the solution is approached. The number of iterations between basis factorizations will therefore increase. During these iterations a test is made regularly according to the value of 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.

If $i \leq 0$, the default value is used.

Infeasible Exit Default
Feasible Exit

Note that this option is ignored if the value of optional parameter Major Iteration Limit is exceeded, or the linear constraints are infeasible.

If termination is about to occur at a point that does not satisfy the nonlinear constraints and optional parameter Feasible Exit is selected, this option requests that additional iterations be performed in order to find a feasible point (if any) for the nonlinear constraints. This involves solving a feasible point problem in which the objective function is omitted.

Otherwise, this option requests no additional iterations be performed.

Minimize Default
Maximize
Feasible Point

If optional parameter Feasible Point is selected, this option attempts to find a feasible point (if any) for the nonlinear constraints by omitting the objective function. It can also be used to check whether the nonlinear constraints are feasible.

Otherwise, this option specifies the required direction of the optimization. It applies to both linear and nonlinear terms (if any) in the objective function. Note that if two problems are the same except that one minimizes $f(x)$ and the other maximizes $-f(x)$, their solutions will be the same but the signs of the dual variables π_i and the reduced gradients d_j will be reversed.

Forward Difference Interval r Default $= \sqrt{\text{Function Precision}}$

This option defines an interval used to estimate derivatives by forward differences in the following circumstances:

(a) For verifying the objective and/or constraint gradients (see the description of the optional parameter Verify Level).

(b) For estimating unspecified elements of the objective gradient and/or the constraint Jacobian.

A derivative with respect to x_j is estimated by perturbing that element of x to the value $x_j + r(1 + |x_j|)$ and then evaluating $f(x)$ and/or $F(x)$ (as appropriate) 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. See Gill *et al.* (1981) for a discussion of the accuracy in finite difference approximations.

If $r \leq 0$, the default value is used.

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

This parameter defines the *relative function precision* ϵ_r, which is intended to be a measure of the relative accuracy with which the nonlinear functions can be computed. For example, if $f(x)$ (or $F_i(x)$) is computed as 1000.56789 for some relevant x and the first 6 significant digits are known to be correct then the appropriate value for ϵ_r would be 10^{-6}.

Ideally the functions $f(x)$ or $F_i(x)$ 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 $f(x)$ (or $F_i(x)$) is computed as $1.23456789 \times 10^{-4}$ for some relevant x and the first 6 significant digits are known to be correct then the appropriate value for ϵ_r would be 10^{-10}.

The choice of ϵ_r can be quite complicated for badly scaled problems; see Chapter 8 of Gill *et al.* (1981) for a discussion of scaling techniques. The default value is appropriate for most simple functions that are computed with full accuracy.

In some cases the function values will be the result of extensive computation, possibly involving an iterative procedure that can provide few digits of precision at reasonable cost. Specifying an appropriate value of r may therefore 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.

If $r < \epsilon$ or $r \geq 1$, the default value is used.

Hessian Frequency i Default $= 99999999$

This option forces the approximate Hessian formed from i BFGS updates to be reset to the identity matrix upon completion of a major iteration. It is intended to be used in conjunction with optional parameter Hessian Full Memory.

If $i \leq 0$, the default value is used and effectively no resets occur.

Hessian Full Memory Default when $\bar{n} < 75$
Hessian Limited Memory Default when $\bar{n} \geq 75$

These options specify the method for storing and updating the quasi-Newton approximation to the Hessian of the Lagrangian function.

If Hessian Full Memory is specified, the approximate Hessian is treated as a dense matrix and BFGS quasi-Newton updates are applied explicitly. This is most efficient when the total number of nonlinear variables is not too large (say, $\bar{n} < 75$). In this case, the storage requirement is fixed and you can expect 1-step Q-superlinear convergence to the solution.

Hessian Limited Memory should only be specified when $\bar{n}$ 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.)

Note that if Hessian Frequency $= 20$ is used in conjunction with Hessian Full Memory, the effect will be similar to using Hessian Limited Memory in conjunction with Hessian Updates $= 20$, except that the latter will retain the current diagonal during resets.

Hessian Updates i Default $= 20$ or 99999999

The default value of i is 20 when Hessian Limited Memory is in effect and 99999999 when Hessian Full Memory is in effect, in which case no updates are performed.

If Hessian Limited Memory is selected, this option defines the maximum number of pairs of Hessian update vectors that are to be used to define the quasi-Newton approximate Hessian. Once the limit of i updates is reached, all but the diagonal elements of the accumulated updates are discarded and the process starts again. Broadly speaking, the more updates that are stored, the better the quality of the approximate Hessian. On the other hand, the more vectors that are stored, the greater the cost of each QP iteration.

The default value of i is likely to give a robust algorithm without significant expense, but faster convergence may be obtained with far fewer updates (e.g., $i = 5$).

If $i < 0$, the default value is used.

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 $r \leq 0$, the default value is used.

Iteration Limit i Default $= 10000$

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 parameters Major Iteration Limit and Minor Iteration Limit.)

If $i < 0$, the default value is used.

Linesearch Tolerance r Default $= 0.9$

This option controls the accuracy with which a step length will be located along the direction of search at each iteration. At the start of each linesearch a target directional derivative for the Lagrangian merit function is identified. The value of r therefore determines the accuracy to which this target value is approximated.

The default value $r = 0.9$ requests an inaccurate search and is appropriate for most problems, particularly those with any nonlinear constraints.

If the nonlinear functions are cheap to evaluate, a more accurate search may be appropriate; try $r = 0.1, 0.01$ or 0.001. The number of major iterations required to solve the problem might decrease.

If the nonlinear functions are expensive to evaluate, a less accurate search may be appropriate. If Derivative Level $= 3$, try $r = 0.99$. (The number of major iterations required to solve the problem might increase, but the total number of function evaluations may decrease enough to compensate.)

If Derivative Level < 3, a moderately accurate search may be appropriate; try $r = 0.5$. Each search will (typically) require only $1 - 5$ function values, but many function calls will then be needed to estimate the missing gradients for the next iteration.

If $r < 0$ or $r \geq 1$, the default value is used.

List Default for E04UGF $=$ List
Nolist Default for E04UGA $=$ Nolist

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 restore printing.

LU Density Tolerance r_1 Default $= 0.6$
LU Singularity Tolerance r_2 Default $= \epsilon^{0.67}$

If $r_1 > 0$, r_1 defines the density tolerance used during the LU factorization of the basis matrix. 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. The remaining matrix is then factorized using a dense LU procedure. Increasing the value of r_1 towards unity may give slightly sparser LU factors, with a slight increase in factorization time. If $r_1 \leq 0$, the default value is used.

If $r_2 > 0$, r_2 defines the singularity tolerance used to guard against ill-conditioned basis matrices. Whenever the basis is refactorized, the diagonal elements of U are tested as follows. If $\left|u_{jj}\right| \leq r_2$ or $\left|u_{jj}\right| < r_2 \times \max_i \left|u_{ij}\right|$, the jth column of the basis is replaced by the corresponding slack variable. This is most likely to occur when START = 'W' (see Section 5), or at the start of a major iteration. If $r_2 \leq 0$, the default value is used.

In some cases, the Jacobian matrix may converge to values that make the basis exactly singular (e.g., a whole row of the Jacobian matrix could be zero at an optimal solution). Before exact singularity occurs, the basis could become very ill-conditioned and the optimization could progress very slowly (if at all). Setting $r_2 = 0.00001$ (say) may therefore help cause a judicious change of basis in such situations.

LU Factor Tolerance r_1 Default $= 5.0$ or 100.0
LU Update Tolerance r_2 Default $= 5.0$ or 10.0

The default value of r_1 is 5.0 if there are any nonlinear constraints and 100.0 otherwise. The default value of r_2 is 5.0 if there are any nonlinear constraints and 10.0 otherwise.

If $r_1 \geq 1$ and $r_2 \geq 1$, the values of r_1 and r_2 affect the stability and sparsity of the basis factorization $B = LU$, during refactorization and updating, 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 satisfy $|\mu| \leq r_i$. Smaller values of r_i favour stability, while larger values favour sparsity. The default values of r_1 and r_2 usually strike a good compromise. For large and relatively dense problems, setting $r_1 = 10.0$ or 5.0 (say) may give a marked improvement in sparsity without impairing stability to a serious degree. Note that for problems involving band matrices, it may be necessary to reduce r_1 and/or r_2 in order to achieve stability.

If $r_1 < 1$ or $r_2 < 1$, the appropriate default value is used.

Major Feasibility Tolerance r Default $= \sqrt{\epsilon}$

This option specifies how accurately the nonlinear constraints should be satisfied. The default value is appropriate when the linear and nonlinear constraints contain data to approximately that accuracy. A larger value may be appropriate if some of the problem functions are known to be of low accuracy.

Let *rowerr* be defined as the maximum nonlinear constraint violation normalized by the size of the solution. It is required to satisfy

$$rowerr = \max_i \frac{viol_i}{\|(x, s)\|} \leq r,$$

where $viol_i$ is the violation of the ith nonlinear constraint.

If $r \leq \epsilon$, the default value is used.

Major Iteration Limit i Default $= 1000$

The value of i specifies the maximum number of major iterations allowed before termination. It is intended to guard against an excessive number of linearizations of the nonlinear constraints. Setting $i = 0$ and Major Print Level > 0 means that the objective and constraint gradients will be checked if Verify Level > 0 and the workspace needed to start solving the problem will be computed and printed, but no iterations will be performed.

If $i < 0$, the default value is used.

Major Optimality Tolerance r Default $= \sqrt{\epsilon}$
Optimality Tolerance r

This option specifies the final accuracy of the dual variables. If E04UGF/E04UGA terminates with IFAIL $= 0$, a primal and dual solution (x, s, π) will have been computed such that

$$maxgap = \max_j \frac{gap_j}{\|\pi\|} \leq r,$$

where gap_j is an estimate of the complementarity gap for the jth variable and $\|\pi\|$ is a measure of the size of the QP dual variables (or Lagrange multipliers) given by

$$\|\pi\| = \max\left(\frac{\sigma}{\sqrt{m}}, 1\right), \qquad \text{where} \qquad \sigma = \sum_{i=1}^{m} |\pi_i|.$$

It is included to make the tests independent of a scale factor on the objective function. Specifically, gap_j is computed from the final QP solution using the reduced gradients $d_j = g_j - \pi^{\mathrm{T}} a_j$, where g_j is the jth element of the objective gradient and a_j is the associated column of the constraint matrix $(A \quad -I)$:

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

If $r \leq 0$, the default value is used.

Major Print Level i Default for E04UGF $= 10$
Default for E04UGA $= 0$
Print Level

The value of i controls the amount of printout produced by the major iterations of E04UGF/E04UGA, as indicated below. A detailed description of the printed output is given in Section 8.1 (summary output at each major iteration and the final solution) and Section 12 (monitoring information at each major iteration). (See also the description of Minor Print Level.)

The following printout is sent to the current advisory message unit (as defined by X04ABF):

i	Output
0	No output.
1	The final solution only.
5	One line of summary output (< 80 characters; see Section 8.1) for each major iteration (no printout of the final solution).
≥ 10	The final solution and one line of summary output for each major iteration.

The following printout is sent to the logical unit number defined by the optional parameter Monitoring File:

i	Output
0	No output.
1	The final solution only.
5	One long line of output (< 120 characters; see Section 12) for each major iteration (no printout of the final solution).
≥ 10	The final solution and one long line of output for each major iteration.
≥ 20	The final solution, one long line of output for each major iteration, matrix statistics (initial status of rows and columns, number of elements, density, biggest and smallest elements, etc.), details of the scale factors resulting from the scaling procedure (if Scale Option $= 1$ or 2), basis factorization statistics and details of the initial basis resulting from the Crash procedure (if START $=$ 'C'; see Section 5).

If Major Print Level ≥ 5 and the unit number defined by the optional parameter Monitoring File is the same as that defined by X04ABF then the summary output for each major iteration is suppressed.

Major Step Limit r Default $= 2.0$

If $r > 0, r$ 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 in the interval $0 < \alpha \leq \beta$, where $\beta = 1$ if there are any nonlinear constraints, or the step to the nearest upper or lower bound on x if all the constraints are linear. Normally, the first step attempted is $\alpha_1 = \min(1, \beta)$.

In some cases, such as $f(x) = ae^{bx}$ or $f(x) = ax^b$, even a moderate change in the elements of x can lead to floating-point overflow. The parameter r is therefore used to define a step limit $\bar{\beta}$ given by

$$\bar{\beta} = \frac{r\big(1 + \|x\|_2\big)}{\|p\|_2},$$

where p is the search direction and the first evaluation of $f(x)$ is made at the (potentially) smaller step length $\alpha_1 = \min\big(1, \bar{\beta}, \beta\big)$.

Wherever possible, upper and lower bounds on x should be used to prevent evaluation of nonlinear functions at meaningless points. The default value $r = 2.0$ should not affect progress on well-behaved functions, but values such as $r = 0.1$ or 0.01 may be helpful when rapidly varying functions are present. If a small value of r is selected, a 'good' starting point may be required. An important application is to the class of nonlinear least-squares problems.

If $r \leq 0$, the default value is used.

Minor Feasibility Tolerance r Default $= \sqrt{\epsilon}$
Feasibility Tolerance r

This option attempts to ensure that all variables eventually satisfy their upper and lower bounds to within the tolerance r. Since this includes slack variables, general linear constraints should also be satisfied to within r. Note that feasibility with respect to nonlinear constraints is judged by the value of optional parameter Major Feasibility Tolerance and not by r.

If the bounds and linear constraints cannot be satisfied to within r, the problem is declared *infeasible*. Let `Sinf` be the corresponding sum of infeasibilities. If `Sinf` 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.

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

Nonlinear functions will only be evaluated at points that satisfy the bounds and linear constraints. If there are regions where a function is undefined, every effort should be made to eliminate these regions from the problem. For example, if $f(x_1, x_2) = \sqrt{x_1} + \log(x_2)$, it is essential to place lower bounds on both x_1 and x_2. If the value $r = 10^{-6}$ is used, the bounds $x_1 \geq 10^{-5}$ and $x_2 \geq 10^{-4}$ might be appropriate. (The log singularity is more serious; in general, you should attempt to keep x as far away from singularities as possible.)

In reality, r is used 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 and the routine is then in *elastic mode* thereafter (with only the linearized nonlinear constraints defined to be elastic). (See also the description of Elastic Weight.)

If $r \leq \epsilon$, the default value is used.

Minor Iteration Limit i Default $= 500$

The value of i specifies the maximum number of iterations allowed between successive linearizations of the nonlinear constraints. A value in the range $10 \leq i \leq 50$ prevents excessive effort being expended on early major iterations, but allows later QP subproblems to be solved to completion. Note that an extra m minor iterations are allowed if the first QP subproblem to be solved starts with the all-slack basis $B = I$. (See the description of the optional parameter Crash Option.)

In general, it is unsafe to specify values as small as $i = 1$ or 2 (because even when an optimal solution has been reached, a few minor iterations may be needed for the corresponding QP subproblem to be recognized as optimal).

If $i \leq 0$, the default value is used.

Minor Optimality Tolerance r Default $= \sqrt{\epsilon}$

This option is used to judge optimality for each QP subproblem. Let the QP reduced gradients be $d_j = g_j - \pi^{\mathrm{T}} a_j$, where g_j is the jth element of the QP gradient, a_j is the associated column of the QP constraint matrix and π is the set of QP dual variables.

By construction, the reduced gradients for basic variables are always zero. The QP subproblem will be declared optimal if the reduced gradients for nonbasic variables at their upper or lower bounds satisfy

$$\frac{d_j}{\|\pi\|} \geq -r \qquad \text{or} \qquad \frac{d_j}{\|\pi\|} \leq r$$

respectively, and if $\dfrac{|d_j|}{\|\pi\|} \leq r$ for superbasic variables.

Note that $\|\pi\|$ is a measure of the size of the dual variables. It is included to make the tests independent of a scale factor on the objective function. (The value of $\|\pi\|$ actually used is defined in the description for optional parameter Major Optimality Tolerance.)

If the objective is scaled down to be very *small*, the optimality test reduces to comparing d_j against r.

If $r \leq 0$, the default value is used.

Minor Print Level i Default $= 0$

The value of i controls the amount of printout produced by the minor iterations of E04UGF/E04UGA (i.e., the iterations of the quadratic programming algorithm), as indicated below. A detailed description of the printed output is given in Section 8.2 (summary output at each minor iteration) and Section 12 (monitoring information at each minor iteration). (See also the description of the optional parameter Major Print Level.)

The following printout is sent to the current advisory message unit (as defined by X04ABF):

i	**Output**
0	No output.
≥ 1	One line of summary output (< 80 characters; see Section 8.2) for each minor iteration.

The following printout is sent to the logical unit number defined by the optional parameter Monitoring File:

i	**Output**
0	No output.
≥ 1	One long line of output (< 120 characters; see Section 12) for each minor iteration.

If Minor Print Level ≥ 1 and the unit number defined by the optional parameter Monitoring File is the same as that defined by X04ABF then the summary output for each minor iteration is suppressed.

Monitoring File i Default $= -1$

If $i \geq 0$ and Major Print Level ≥ 5 or $i \geq 0$ and Minor Print Level ≥ 1 then monitoring information is produced by E04UGF/E04UGA at every iteration is sent to a file with logical unit number i. If $i < 0$ and/or Major Print Level < 5 and Minor Print Level < 1 then no monitoring information is produced.

Partial Price i Default $= 1$ or 10

The default value of i is 1 if there are any nonlinear constraints and 10 otherwise.

This option is recommended for large problems that have significantly more variables than constraints (i.e., $n \gg m$). It reduces the work required for each 'pricing' operation (i.e., when a nonbasic variable is selected to become superbasic). The possible choices for i are the following.

i	**Meaning**
1	All columns of the constraint matrix $(A \quad -I)$ are searched.
≥ 2	Both A and I are partitioned to give i roughly equal segments A_j, I_j, for $j = 1, 2, \ldots, p$ (modulo p). If the previous pricing search was successful on A_j, I_j, the next search begins on the segments A_{j+1}, I_{j+1}. 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 enter the basis. If nothing is found, the search continues on the next segments A_{j+2}, I_{j+2} and so on.

If $i \leq 0$, the default value is used.

Pivot Tolerance r Default $= \epsilon^{0.67}$

If $r > 0$, r is used during the solution of QP subproblems 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 specified search direction p, a 'ratio test' is used to determine which element 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 r.

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

If $r \leq 0$, the default value is used.

Scale Option i Default $= 1$ or 2

The default value of i is 1 if there are any nonlinear constraints and 2 otherwise.

This option enables you to scale the variables and constraints using an iterative procedure due to Fourer (1982), which attempts to compute row scales r_i and column scales c_j such that the scaled matrix coefficients $\bar{a}_{ij} = a_{ij} \times \left(c_j / r_i\right)$ are as close as possible to unity. (The lower and upper bounds on the variables and slacks for the scaled problem are redefined as $\bar{l}_j = l_j / c_j$ and $\bar{u}_j = u_j / c_j$ respectively, where $c_j \equiv r_{j-n}$ if $j > n$.) The possible choices for i are the following.

i	**Meaning**
0	No scaling is performed. This is recommended if it is known that the elements of x and the constraint matrix A (along with its Jacobian) never become large (say, > 1000).
1	All linear constraints and variables are scaled. This may improve the overall efficiency of the routine on some problems.
2	All constraints and variables are scaled. Also, an additional scaling is performed that takes into account columns of $(A \quad -I)$ that are fixed or have positive lower bounds or negative upper bounds.

If there are any nonlinear constraints present, the scale factors depend on the Jacobian at the first point that satisfies the linear constraints and the upper and lower bounds. The setting $i = 2$ should therefore be used only if a 'good' starting point is available and the problem is not highly nonlinear.

If $i < 0$ or $i > 2$, the default value is used.

Scale Tolerance r Default $= 0.9$

Note that this option does not apply when Scale Option $= 0$.

The value r $(0 < r < 1)$ is used to control the number of scaling passes to be made through the constraint matrix A. At least 3 (and at most 10) passes will be made. More precisely, let a_p denote the largest

column ratio (i.e., $\frac{\text{'biggest' element}}{\text{'smallest' element}}$ in some sense) after the pth scaling pass through A. The scaling procedure is terminated if $a_p \geq a_{p-1} \times r$ for some $p \geq 3$. Thus, increasing the value of r from 0.9 to 0.99 (say) will probably increase the number of passes through A.

If $r \leq 0$ or $r \geq 1$, the default value is used.

Start Objective Check At Column	i_1	Default $= 1$
Stop Objective Check At Column	i_2	Default $= n_1'$
Start Constraint Check At Column	i_3	Default $= 1$
Stop Constraint Check At Column		Default $= n_1''$

These keywords take effect only if Verify Level > 0. They may be used to control the verification of gradient elements computed by OBJFUN and/or Jacobian elements computed by 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 then it is reasonable to specify Start Objective Check At Column $= 31$. Similarly for columns of the Jacobian. If the first 30 variables occur nonlinearly in the constraints but the remaining variables are nonlinear only in the objective, then OBJFUN must set the first 30 elements of the array OBJGRD to zero, but these hardly need to be verified. Again it is reasonable to specify Start Objective Check At Column $= 31$.

If $i_2 \leq 0$ or $i_2 > n_1'$, the default value is used.

If $i_1 \leq 0$ or $i_1 > \min(n_1', i_2)$, the default value is used.

If $i_4 \leq 0$ or $i_4 > n_1''$, the default value is used.

If $i_3 \leq 0$ or $i_3 > \min(n_1'', i_4)$, the default value is used.

Superbasics Limit i Default $= \min(500, \bar{n} + 1)$

Note that this option does not apply to linear problems.

It places a limit on the storage allocated for superbasic variables. Ideally, the value of i should be set slightly larger than the 'number of degrees of freedom' expected at the solution.

For nonlinear problems, the number of degrees of freedom is often called the 'number of independent variables'. Normally, the value of i need not be greater than $\bar{n} + 1$, but for many problems it may be considerably smaller. (This will save storage if $\bar{n}$ is very large.)

If $i \leq 0$, the default value is used.

Unbounded Objective	r_1	Default $= 10^{15}$
Unbounded Step Size	r_2	Default $= \max(bigbnd, 10^{20})$

These options are intended to detect unboundedness in nonlinear problems. During the linesearch, the objective function 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, the iterations are terminated and the routine returns with IFAIL $= 3$.

If singularities are present, unboundedness in $f(x)$ may manifest itself 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.

If $r_1 \leq 0$ or $r_2 \leq 0$, the appropriate default value is used.

Verify Level i Default $= 0$

This option refers to finite difference checks on the gradient elements computed by OBJFUN and CONFUN. Gradients are verified at the first point that satisfies the linear constraints and the upper and lower bounds. Unspecified gradient elements are not checked and hence they result in no overhead. The possible choices for i are the following.

i	**Meaning**
-1	No checks are performed.

0 Only a 'cheap' test will be performed, requiring three calls to OBJFUN and two calls to CONFUN. Note that no checks are carried out if every column of the constraint gradients (Jacobian) contains a missing element.

1 Individual objective gradient elements will be checked using a reliable (but more expensive) test. If Major Print Level > 0, a key of the form `OK` or `BAD?` indicates whether or not each element appears to be correct.

2 Individual columns of the constraint gradients (Jacobian) will be checked using a reliable (but more expensive) test. If Major Print Level > 0, a key of the form `OK` or `BAD?` indicates whether or not each element appears to be correct.

3 Check both constraint and objective gradients (in that order) as described above for $i = 2$ and $i = 1$ respectively.

The value $i = 3$ should be used whenever a new function routine is being developed. The Start Objective Check At Column and Stop Objective Check At Column keywords may be used to limit the number of nonlinear variables to be checked.

If $i < -1$ or $i > 3$, the default value is used.

Violation Limit r Default $= 10.0$

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

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

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

The effect of the 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 function is expected to be well-defined and bounded below (or above in the case of maximization). If the objective function is bounded below (or above in the case of maximization) for all values of the variables, then r may be any large positive value.

If $r \leq 0$, the default value is used.

12 Description of Monitoring Information

This section describes the intermediate printout and final printout which constitutes the monitoring information produced by E04UGF/E04UGA. (See also the description of the optional parameters Monitoring File, Major Print Level and Minor Print Level.) You can control the level of printed output.

When Major Print Level ≥ 20 and Monitoring File ≥ 0, the following line of intermediate printout (< 120 characters) is produced at every major iteration on the unit number specified by optional parameter Monitoring File. Unless stated otherwise, the values of the quantities printed are those in effect *on completion* of the given iteration.

`Major` is the major iteration count.

`Minor` is the number of minor iterations required by the feasibility and optimality phases of the QP subproblem. Generally, `Minor` will be 1 in the later iterations, since theoretical analysis predicts that the correct active set will be identified near the solution (see Section 10).

`Step` is the step α_k taken along the computed search direction. On reasonably well-behaved problems, the unit step (i.e., $\alpha_k = 1$) will be taken as the solution is approached.

`nObj` is the number of times OBJFUN has been called to evaluate the nonlinear part of the objective function. Evaluations needed for the estimation of the gradients by finite differences are not included. `nObj` is printed as a guide to the amount of work required for the linesearch.

`nCon` is the number of times CONFUN has been called to evaluate the nonlinear constraint functions (not printed if NCNLN is zero).

`Merit` is the value of the augmented Lagrangian merit function (6) at the current iterate. This function will decrease at each iteration unless it was necessary to increase the penalty parameters (see Section 8.1). As the solution is approached, `Merit` will converge to the value of the objective function at the solution.

In elastic mode (see Section 10.2), the merit function is a composite function involving the constraint violations weighted by the value of the optional parameter Elastic Weight.

If there are no nonlinear constraints present, this entry contains `Objective`, the value of the objective function $f(x)$. In this case, $f(x)$ will decrease monotonically to its optimal value.

`Feasibl` is the value of *rowerr*, the largest element of the scaled nonlinear constraint residual vector defined in the description of the optional parameter Major Feasibility Tolerance. The solution is regarded as 'feasible' if `Feasibl` is less than (or equal to) the optional parameter Major Feasibility Tolerance. `Feasibl` will be approximately zero in the neighbourhood of a solution.

If there are no nonlinear constraints present, all iterates are feasible and this entry is not printed.

`Optimal` is the value of *maxgap*, the largest element of the maximum complementarity gap vector defined in the description of the optional parameter Major Optimality Tolerance. The Lagrange multipliers are regarded as 'optimal' if `Optimal` is less than (or equal to) the optional parameter Major Optimality Tolerance. `Optimal` will be approximately zero in the neighbourhood of a solution.

`nS` is the current number of superbasic variables.

`Penalty` is the Euclidean norm of the vector of penalty parameters used in the augmented Lagrangian merit function (not printed if NCNLN is zero).

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

If there are nonlinear constraints present, the basis factorization $B = LU$ is computed at the start of the first minor iteration. At this stage, `LU` $=$ `lenL` $+$ `lenU`, where `lenL` is the number of subdiagonal elements in the columns of a lower triangular matrix and `lenU` 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, the value of `LU` may fluctuate up or down (but in general will tend to increase). As the solution is approached and the number of minor iterations required to solve each QP subproblem decreases towards zero, `LU` will reflect the number of nonzeros in the LU factors at the start of each QP subproblem.

If there are no nonlinear constraints present, refactorization is subject only to the value of the optional parameter Factorization Frequency and hence `LU` will tend to increase between factorizations.

`Swp` is the number of columns of the basis matrix B that were swapped with columns of S in order to improve the condition number of B (not printed if NCNLN is zero). The swaps are determined by an LU factorization of the rectangular matrix $B_S = (B \quad S)^{\mathrm{T}}$, with stability being favoured more than sparsity.

`Cond Hz` is an estimate of the condition number of the reduced Hessian of the Lagrangian (not printed if NCNLN and NONLN are both zero). It is the square of the ratio between the largest and smallest diagonal elements of the upper triangular matrix R. This constitutes a lower bound on the condition number of the matrix $R^{\mathrm{T}}R$ that approximates the reduced Hessian. The larger this number, the more difficult the problem.

`PD` is a two-letter indication of the status of the convergence tests involving the feasibility and optimality of the iterates defined in the descriptions of the optional parameters Major Feasibility Tolerance and Major Optimality Tolerance. Each letter is `T` if the test is satisfied and `F` otherwise. The tests indicate whether the values of `Feasibl` and `Optimal` are sufficiently small. For example, `TF` or `TT` is printed if there are no nonlinear constraints present (since all iterates are feasible). If either indicator is `F` when E04UGF/E04UGA terminates with IFAIL $= 0$, you should check the solution carefully.

`M` is printed if an extra evaluation of user-supplied subroutines OBJFUN and CONFUN was needed in order to define an acceptable positive-definite quasi-Newton update to the Hessian of the Lagrangian. This modification is only performed when there are nonlinear constraints present.

`m` is printed if, in addition, it was also necessary to modify the update to include an augmented Lagrangian term.

`s` is printed if a self-scaled BFGS (Broyden–Fletcher–Goldfarb–Shanno) update was performed. This update is always used when the Hessian approximation is diagonal and hence always follows a Hessian reset.

`S` is printed if, in addition, it was also necessary to modify the self-scaled update in order to maintain positive-definiteness.

`n` is printed if no positive-definite BFGS update could be found, in which case the approximate Hessian is unchanged from the previous iteration.

`r` is printed if the approximate Hessian was reset after 10 consecutive major iterations in which no BFGS update could be made. The diagonal elements of the approximate Hessian are retained if at least one update has been performed since the last reset. Otherwise, the approximate Hessian is reset to the identity matrix.

`R` is printed if the approximate Hessian has been reset by discarding all but its diagonal elements. This reset will be forced periodically by the values of the optional parameters Hessian Frequency and Hessian Updates. However, it may also be necessary to reset an ill-conditioned Hessian from time to time.

`l` is printed if the change in the norm of the variables was greater than the value defined by the optional parameter Major Step Limit. If this output occurs frequently during later iterations, it may be worthwhile increasing the value of Major Step Limit.

`c` is printed if central differences have been used to compute the unknown elements 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.

`u` is printed if the QP subproblem was unbounded.

`t` is printed if the minor iterations were terminated after the number of iterations specified by the value of the optional parameter Minor Iteration Limit was reached.

`i` is printed if the QP subproblem was infeasible when the routine was not in elastic mode. 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 (8) (see Section 10.2). It is also printed if the minimizer of the elastic subproblem does not satisfy the linearized constraints when the routine is already in elastic mode. (In this case, a feasible point for the usual QP subproblem may or may not exist.)

`w` is printed if a weak solution of the QP subproblem was found.

When Minor Print Level ≥ 1 and Monitoring File ≥ 0, the following line of intermediate printout (< 120 characters) is produced at every minor iteration on the unit number specified by optional parameter

Monitoring File. Unless stated otherwise, the values of the quantities printed are those in effect *on completion* of the given iteration.

In the description below, a 'pricing' operation is defined to be the process by which a nonbasic variable is selected to become superbasic (in addition to those already in the superbasic set). If the problem is purely linear, the variable selected will usually become basic immediately (unless it happens to reach its opposite bound and return to the nonbasic set).

`Itn` is the iteration count.

`pp` is the partial price indicator. The variable selected by the last pricing operation came from the `pp`th partition of A and $-I$. Note that `pp` is reset to zero whenever the basis is refactorized.

`dj` is the value of the reduced gradient (or reduced cost) for the variable selected by the pricing operation at the start of the current iteration.

`+SBS` is the variable selected by the pricing operation to be added to the superbasic set.

`-SBS` is the variable chosen to leave the superbasic set. It has become basic if the entry under `-B` is nonzero; otherwise it has become nonbasic.

`-BS` is the variable removed from the basis (if any) to become nonbasic.

`-B` is the variable removed from the basis (if any) to swap with a slack variable made superbasic by the latest pricing operation. The swap is done to ensure that there are no superbasic slacks.

`Step` is the value of the step length α taken along the current search direction p. The variables x have just been changed to $x + \alpha p$. If a variable is made superbasic during the current iteration (i.e., `+SBS` is positive), `Step` will be the step to the nearest bound. During the optimality phase, the step can be greater than unity only if the reduced Hessian is not positive-definite.

`Pivot` is the rth element of a vector y satisfying $By = a_q$ whenever a_q (the qth column of the constraint matrix $(A \quad -I)$) replaces the rth column of the basis matrix B. Wherever possible, `Step` is chosen so as to avoid extremely small values of `Pivot` (since they may cause the basis to be nearly singular). In extreme cases, it may be necessary to increase the value of the optional parameter Pivot Tolerance to exclude very small elements of y from consideration during the computation of `Step`.

`Ninf` is the number of infeasibilities. This will not increase unless the iterations are in elastic mode. `Ninf` will be zero during the optimality phase.

`Sinf/Objective` is the value of the current objective function. If x is infeasible, `Sinf` gives the value of the sum of infeasibilities at the start of the current iteration. It will usually decrease at each nonzero value of `Step`, but may occasionally increase if the value of `Ninf` decreases by a factor of 2 or more. However, in elastic mode this entry gives the value of the composite objective function (9), which will decrease monotonically at each iteration. If x is feasible, `Objective` is the value of the current QP objective function.

`L` is the number of nonzeros in the basis factor L. Immediately after a basis factorization $B = LU$, this entry contains `lenL`. Further nonzeros are added to `L` when various columns of B are later replaced. (Thus, `L` increases monotonically.)

`U` is the number of nonzeros in the basis factor U. Immediately after a basis factorization $B = LU$, this entry contains `lenU`. As columns of B are replaced, the matrix U is maintained explicitly (in sparse form). The value of `U` may fluctuate up or down; in general, it will tend to increase.

`Ncp` is the number of compressions required to recover workspace in the data structure for U. This includes the number of compressions needed during the previous basis factorization. Normally, `Ncp` should increase very slowly. If it does not, increase LENIZ and LENZ by at least `L` + `U` and rerun E04UGF/E04UGA (possibly using START = 'W'; see Section 5).

The following items are printed only if the problem is nonlinear or the superbasic set is non-empty (i.e., if the current solution is nonbasic).

`Norm rg` is the Euclidean norm of the reduced gradient of the QP objective function. During the optimality phase, this norm will be approximately zero after a unit step.

`nS` is the current number of superbasic variables.

`Cond Hz` is an estimate of the condition number of the reduced Hessian of the Lagrangian (not printed if NCNLN and NONLN are both zero). It is the square of the ratio between the largest and smallest diagonal elements of the upper triangular matrix R. This constitutes a lower bound on the condition number of the matrix $R^{\mathrm{T}}R$ that approximates the reduced Hessian. The larger this number, the more difficult the problem.

When Major Print Level ≥ 20 and Monitoring File ≥ 0, the following lines of intermediate printout (< 120 characters) are produced on the unit number specified by optional parameter Monitoring File whenever the matrix B or $B_S = (B \quad S)^{\mathrm{T}}$ is factorized before solving the next QP subproblem. Gaussian elimination is used to compute a sparse LU factorization of B or B_S, where PLP^{T} is a lower triangular matrix and PUQ is an upper triangular matrix for some permutation matrices P and Q. The factorization is stabilized in the manner described under the optional parameter LU Factor Tolerance (default value = 5.0 or 100.0).

Note that B_S may be factorized at the beginning of just some of the major iterations. It is immediately followed by a factorization of B itself.

`Factorize` is the factorization count.

`Iteration` is the iteration count.

`Nonlinear` is the number of nonlinear variables in the current basis B (not printed if B_S is factorized).

`Linear` is the number of linear variables in B (not printed if B_S is factorized).

`Slacks` is the number of slack variables in B (not printed if B_S is factorized).

`Elems` is the number of nonzeros in B (not printed if B_S is factorized).

`Density` is the percentage nonzero density of B (not printed if B_S is factorized). More precisely, $\texttt{Density} = 100 \times \texttt{Elems}/(\texttt{Nonlinear} + \texttt{Linear} + \texttt{Slacks})^2$.

`Compressns` is the number of times the data structure holding the partially factorized matrix needed to be compressed, in order to recover unused workspace. Ideally, it should be zero. If it is more than 3 or 4, increase LENIZ and LENZ and rerun E04UGF/E04UGA (possibly using START = 'W'; see Section 5).

`Merit` 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 `m` such quantities. It gives an indication of how much work was required to preserve sparsity during the factorization.

`lenL` is the number of nonzeros in L.

`lenU` is the number of nonzeros in U.

`Increase` is the percentage increase in the number of nonzeros in L and U relative to the number of nonzeros in B. More precisely, $\texttt{Increase} = 100 \times (\texttt{lenL} + \texttt{lenU} - \texttt{Elems})/\texttt{Elems}$.

`m` is the number of rows in the problem. Note that $\texttt{m} = \texttt{Ut} + \texttt{Lt} + \texttt{bp}$.

`Ut` is the number of triangular rows of B at the top of U.

`d1` is the number of columns remaining when the density of the basis matrix being factorized reached 0.3.

`Lmax` is the maximum subdiagonal element in the columns of L. This will not exceed the value of the optional parameter LU Factor Tolerance.

`Bmax` is the maximum nonzero element in B (not printed if B_S is factorized).

`BSmax` is the maximum nonzero element in B_S (not printed if B is factorized).

`Umax` is the maximum nonzero element in U, excluding elements of B that remain in U unchanged. (For example, if a slack variable is in the basis, the corresponding row of B will become a row of U without modification. Elements in such rows will not contribute to `Umax`. If the basis is strictly triangular then *none* of the elements of B will contribute and `Umax` will be zero.)

Ideally, `Umax` should not be significantly larger than `Bmax`. If it is several orders of magnitude larger, it may be advisable to reset the optional parameter LU Factor Tolerance to some value nearer unity.

`Umax` is not printed if B_S is factorized.

`Umin` is the magnitude of the smallest diagonal element of PUQ.

`Growth` is the value of the ratio `Umax`/`Bmax`, which should not be too large.

Providing `Lmax` is not large (say, < 10.0), the ratio $\max(\texttt{Bmax}, \texttt{Umax})/\texttt{Umin}$ is an estimate of the condition number of B. If this number is extremely large, the basis is nearly singular and some numerical difficulties might occur. (However, an effort is made to avoid near-singularity by using slacks to replace columns of B that would have made `Umin` extremely small and the modified basis is refactorized.)

`Lt` is the number of triangular columns of B at the left of L.

`bp` is the size of the 'bump' or block to be factorized nontrivially after the triangular rows and columns of B have been removed.

`d2` is the number of columns remaining when the density of the basis matrix being factorized has reached 0.6.

When Major Print Level ≥ 20, Monitoring File ≥ 0 and Crash Option > 0 (default value $= 0$ or 3), the following lines of intermediate printout (< 80 characters) are produced on the unit number specified by optional parameter Monitoring File whenever START = 'C' (see Section 5). They refer to the number of columns selected by the Crash procedure during each of several passes through A while searching for a triangular basis matrix.

`Slacks` is the number of slacks selected initially.

`Free cols` is the number of free columns in the basis, including those whose bounds are rather far apart.

`Preferred` is the number of 'preferred' columns in the basis (i.e., ISTATE$(j) = 3$ for some $j \leq n$). It will be a subset of the columns for which ISTATE$(j) = 3$ was specified.

`Unit` is the number of unit columns in the basis.

`Double` is the number of columns in the basis containing two nonzeros.

`Triangle` is the number of triangular columns in the basis with three (or more) nonzeros.

`Pad` is the number of slacks used to pad the basis (to make it a nonsingular triangle).

When Major Print Level $= 1$ or ≥ 10 and Monitoring File ≥ 0, the following lines of final printout (< 120 characters) are produced on the unit number specified by optional parameter Monitoring File.

Let x_j denote the jth 'column variable', for $j = 1, 2, \ldots, n$. We assume that a typical variable x_j has bounds $\alpha \leq x_j \leq \beta$.

The following describes the printout for each column (or variable). A full stop (.) is printed for any numerical value that is zero.

`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` gives the state of x_j relative to the bounds α and β.

The various possible states 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 $\alpha < x_j < \beta$.

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` *Alternative optimum possible*. 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 current value, there would be no change in the value of the objective function. The values of the basic and superbasic variables *might* change, giving a genuine alternative solution. The values of the Lagrange multipliers *might* also change.

`D` *Degenerate*. The variable is basic, but it is equal to (or very close to) one of its bounds.

`I` *Infeasible*. The variable is basic and is currently violating one of its bounds by more than the value of the optional parameter Minor Feasibility Tolerance.

`N` *Not precisely optimal*. The variable is nonbasic. Its reduced gradient is larger than the value of the optional parameter Major Feasibility Tolerance.

`Activity` is the value of x_j at the final iterate.

`Obj Gradient` is the value of g_j at the final iterate. (If any x_j is infeasible, g_j is the gradient of the sum of infeasibilities.)

`Lower Bound` is the lower bound specified for the variable. `None` indicates that $\mathrm{BL}(j) \leq -bigbnd$.

`Upper Bound` is the upper bound specified for the variable. `None` indicates that $\mathrm{BU}(j) \geq bigbnd$.

`Reduced Gradnt` is the value of d_j at the final iterate.

`m + j` is the value of $m + j$.

General linear constraints take the form $l \leq Ax \leq u$. The ith constraint is therefore of the form $\alpha \leq a_i^{\mathrm{T}} x \leq \beta$ and the value of $a_i^{\mathrm{T}} x$ is called the *row activity*. Internally, the linear constraints take the form $Ax - s = 0$, where the slack variables s should satisfy the bounds $l \leq s \leq u$. For the ith 'row', it is the slack variable s_i that is directly available and it is sometimes convenient to refer to its state. Slacks may be basic or nonbasic (but not superbasic).

Nonlinear constraints $\alpha \leq F_i(x) + a_i^{\mathrm{T}} x \leq \beta$ are treated similarly, except that the row activity and degree of infeasibility are computed directly from $F_i(x) + a_i^{\mathrm{T}} x$ rather than from s_i.

The following describes the printout for each row (or constraint). A full stop (.) is printed for any numerical value that is zero.

`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 ith row.

`State` gives the state of the ith row relative to the bounds α and β.

The various possible states are as follows:

`LL` The row is at its lower limit, α.

`UL` The row is at its upper limit, β.

`EQ` The limits are the same ($\alpha = \beta$).

`BS` The constraint is not binding. s_i is basic.

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` *Alternative optimum possible*. 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 current value, there would be no change in the value of the objective function. The values of the basic and superbasic variables *might* change, giving a genuine alternative solution. The values of the Lagrange multipliers *might* also change.

`D` *Degenerate*. The variable is basic, but it is equal to (or very close to) one of its bounds.

`I` *Infeasible*. The variable is basic and is currently violating one of its bounds by more than the value of the optional parameter Minor Feasibility Tolerance.

`N` *Not precisely optimal*. The variable is nonbasic. Its reduced gradient is larger than the value of the optional parameter Major Feasibility Tolerance.

`Activity` is the value of $a_i^{\mathrm{T}}x$ (or $F_i(x) + a_i^{\mathrm{T}}x$ 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 Bound` is α, the lower bound specified for the ith row. `None` indicates that $\mathrm{BL}(n+i) \leq -bigbnd$.

`Upper Bound` is β, the upper bound specified for the ith row. `None` indicates that $\mathrm{BU}(n+i) \geq bigbnd$.

`Dual Activity` is the value of the dual variable π_i.

`i` gives the index i of the ith row.

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