JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 125, No. 3, pp. 497–521, June 2005 (© 2005)
DOI: 10.1007/s10957-005-2086-2

Globally Convergent Interior-Point Algorithm for Nonlinear Programming[1]

I. Akrotirianakis[2] and B. Rustem[3]

Communicated by L. C. W. Dixon

Abstract. This paper presents a primal-dual interior-point algorithm for solving general constrained nonlinear programming problems. The inequality constraints are incorporated into the objective function by means of a logarithmic barrier function. Also, satisfaction of the equality constraints is enforced through the use of an adaptive quadratic penalty function. The penalty parameter is determined using a strategy that ensures a descent property for a merit function. Global convergence of the algorithm is achieved through the monotonic decrease of a merit function. Finally, extensive computational results show that the algorithm can solve large and difficult problems in an efficient and robust way.

Key Words. Primal-dual interior-point algorithms, merit functions, convergence theory.

1. Introduction

In this paper, we discuss a primal-dual interior-point algorithm for solving general (nonconvex) nonlinear programming problems. The algorithm is based on two different approaches. The first is the augmented Lagrangian SQP framework for general constrained optimization problems, discussed in Rustem (Ref. 1); the second is the primal-dual

[1]The research reported in this paper was done while the first author was at Imperial College. The authors gratefully acknowledge constructive comments from Professor L. C. W. Dixon and an anonymous referee. They are also grateful to Dr. Stanislav Zakovic for helpful suggestions and comments. Financial support was provided by EPSRC Grants M16016 and GR/G51377/01.
[2]Postdoctoral Research Associate, Department of Chemical Engineering, Princeton University, Princeton, New Jersey.
[3]Professor, Department of Computing, Imperial College, London, UK.

interior-point method, where a barrier function and a damped Newton framework are used in order to solve NLP problems. The algorithm is motivated by the fact that the solution of the first-order optimality conditions of any NLP problem, which is the core of primal-dual interior-point algorithms, is not sufficient to guarantee the convergence to a local minimum, unless the problem is convex. An algorithm that merely solves the first-order optimality conditions may converge to a saddle point or even to a local maximum, since these conditions are also satisfied at those points. To avoid such undesirable behavior, we use a merit function, whose aim is to guide the iterates of the algorithm toward a local minimum. The merit function incorporates the inequality constraints by means of the logarithmic barrier function and the equality constraints by means of the quadratic penalty function. Furthermore, the subproblem that is used to compute the search direction involves the augmented Lagrangian of the equality-constrained barrier problem. The search direction is shown to be of descent for the merit function. It is also shown that the penalty parameter in the merit function does not increase indefinitely, if the iterates of the algorithm are not near a feasible point of the barrier problem. This is a particularly important point as it involves the use of the equality-constrained problem and the corresponding augmented Lagrangian to establish the finiteness of the penalty parameter. If the iterates of the algorithm are near a feasible point of the barrier problem, then a switch in the merit function is activated. In this case, we use the Euclidean norm of the first-order optimality conditions as the merit function. The second merit function, on its own, ensures only convergence to a point satisfying the first-order perturbed optimality conditions, without distinguishing between a minimum or maximum. However, the second merit function is expected to be activated when the iterates have been placed, by the primary penalty-barrier merit function, within a neighborhood of a local minimum. As will be discussed later, it is this switch of the merit functions that enables the global convergence of the algorithm to a local minimum.

Although our algorithm is related to the approaches proposed by El-Bakry et al. (Ref. 2) and Yamashita (Ref. 3), it differs in significant aspects, such as the merit functions, the adaptive penalty selection rule, the stepsize rules, and the technique that switches between the different merit functions. Other algorithms which use an adaptive penalty have been developed recently by Gay et al. (Ref. 4) and Vanderbei and Shanno (Ref. 5).

Recently, general (nonconvex) NLP problems have been the subject of intensive research in the optimization community and several primal-dual interior-point algorithms have emerged (e.g. Refs. 2–4 and 6–7).

The common feature of these algorithms is that they use a merit function within a line-search or trust-region framework to achieve global convergence.

This paper is organized as follows. Section 2 describes the primal-dual interior-point algorithm. In Section 3, we establish the global convergence of the algorithm. In Section 4, we report our numerical experience. We provide also an example where we demonstrate how the mechanism that switches between the two merit functions enables the algorithm to converge to a local minimum and avoid saddle points or local maxima.

2. Description of the Algorithm

In this paper, we consider the following constrained optimization problem:

$$\min \quad f(x), \tag{1a}$$

$$\text{s.t.} \quad g(x)=0, \quad x \geq 0, \tag{1b}$$

where $x \in \Re^n$, $f : \Re^n \to \Re$, and $g(x) : \Re^n \to \Re^q$. The formulation in (1) is quite general because every equality-constrained and inequality-constrained optimization problem can be reduced to that form, by adding for example slack variables to the constraints.

The following assumptions are used throughout the paper:

(A1) The second-order derivatives of f and g are continuous.

(A2) The columns of the matrix $[\nabla g(x), e_i : i \in I_x^0\}$ are linear independent, where $I_x^0 = \{i : \liminf_{k\to\infty} x_k^i = 0, i = 1, 2, \ldots, n\}$ and e_i represents the ith column of the $n \times n$ identity matrix. Also, the sequence $\{x_k\}$ is bounded.

(A3) Strict complementarity of the solution $w_* = (x_*, y_*, z_*)$ is satisfied.

(A4) The second-order sufficiency condition for optimality is satisfied at the solution point; i.e., for all vectors $0 \neq v \in \Re^n$ such that $\nabla g^i(x_*)^T v = 0, i = 1, 2, \ldots, q$, and such that $e_i^T v = 0$, for $i \in I_x^0$, $v^T \nabla_{xx}^2 L(x, y, z) v > 0$.

The original equality and inequality constrained optimization problem (1) is approximated by

$$\min \quad f(x) + (c/2)\|g(x)\|_2^2 - \mu \sum_{i=1}^{n} \log(x^i), \tag{2a}$$

$$\text{s.t.} \quad g(x) = 0, \tag{2b}$$

for $c, \mu \geq 0$. The objective in (1) is augmented by the penalty and logarithmic barrier functions. The penalty is used to enforce satisfaction of the equality constraints by adding a high cost to the objective function for infeasible points. The barrier is needed to introduce an interior-point method to solve the initial problem (1), since it creates a positive singularity at the boundary of the feasible region. Thus, strict feasibility is enforced, while approaching the optimum solution.

The Lagrangian associated with the optimization problem (2) is given by

$$L(x, y; c, \mu) = f(x) + (c/2)\|g(x)\|_2^2 - \mu \sum_{i=1}^{n} \log\left(x^i\right) - g(x)^T y,$$

and the perturbed optimality conditions are

$$F(x, y, z; c, \mu) = \begin{pmatrix} \nabla f(x) - z + c\nabla g(x)^T g(x) - \nabla g(x)^T y \\ g(x) \\ XZe - \mu e \end{pmatrix} = 0, \tag{3}$$

where

$$z = \mu X^{-1} e, \quad X = \mathrm{diag}\{x_1, \ldots, x_n\}, \quad Z = \mathrm{diag}\{z_1, \ldots, z_n\}.$$

For μ fixed, the system (3) is solved by using the quasi-Newton method. At the kth iteration, the Newton system is

$$\begin{pmatrix} H_k & -\nabla g_k^T & -I \\ \nabla g_k & 0 & 0 \\ Z_k & 0 & X_k \end{pmatrix} \begin{pmatrix} \Delta x_k \\ \Delta y_k \\ \Delta z_k \end{pmatrix} = - \begin{pmatrix} \nabla f_k - z_k + c_k \nabla g_k^T g_k - \nabla g_k^T y_k \\ g_k \\ X_k Z_k e - \mu e, \end{pmatrix}, \tag{4}$$

where H_k is a positive-definite approximation of the Hessian of the augmented Lagrangian. In matrix-vector form, equation (4) can be written as

$$F(w_k; c_k)\Delta w_k = -F(w_k; c_k, \mu_k), \tag{5}$$

where $F(w_k; c_k)$ is the Jacobian matrix of the vector function $F(w_k; c_k, \mu_k)$.

To initiate the algorithm, a strictly-interior starting point is needed, that is, a point

$$w^0 = \left(x^0, y^0, z^0\right), \quad \text{with } x^0, z^0 > 0.$$

By controlling the steplength α_{xk} of the primal variables x_k and the steplength α_{zk} of the dual variables y_k and z_k, the algorithm ensures that the generated iterates remain strictly in the interior of the feasible region. Moreover, the algorithm moves from one inner iteration to another inner iteration (i.e., with μ fixed) by seeking to minimize the merit function

$$\Phi(x;c,\mu)=f(x)+(c/2)\|g(x)\|_2^2-\mu\sum_{i=1}^{n}\log(x^i), \tag{6}$$

which is basically the objective function of the barrier problem (2). This is achieved by properly selecting the values of the penalty parameter c at each inner iteration. In order to avoid situations where the penalty parameter may grow to large values, we introduce a second merit function, defined by the ℓ_2 norm of the KKT residuals of the barrier problem (2). The potential difficulties of the penalty parameter in the merit function $\Phi(x;c,\mu)$ have been considered by other authors. For example, Bartholomew-Biggs (Ref. 8) avoids this undesirable situation by constructing search directions based directly on the augmented Lagrangian of the barrier problem. As shown later, the monotonic decrease of both merit functions and the rules for determining the primal and dual stepsizes guarantee that the inner iterates converge to the solution of (2) for a fixed value of μ. Subsequently, by reducing μ, such that $\{\mu\}\to 0$, the optimum of the initial problem (1) is reached.

A detailed description of the algorithm follows.

Algorithm 2.1.

Step 0. Initialization. Choose $\tilde{x}^0,\tilde{z}^0\in\Re^n$, and $\tilde{y}^0\in\Re^q$, such that $\tilde{x}^0,\tilde{z}^0>0$, penalty and barrier parameters $c_0>0,\mu^0>0$, and parameters $\beta,\gamma,\epsilon_0,\eta,\rho\in(0,1),\delta,m,M>0$. Set $l=0$ and $k=0$

Step 1. Test for Convergence of Outer-Iterations. If

$$\|F(\tilde{x}^l,\tilde{y}^l,\tilde{z}^l;c_*,\mu^l)\|_2/(1+\|(\tilde{x}^l,\tilde{y}^l,\tilde{z}^l)\|_2)\le\epsilon_0,$$

then stop.

Step 2. Start of Inner Iterations (μ is fixed to μ^l throughout this step). Set

$$(x_k,y_k,z_k)=(\tilde{x}^l,\tilde{y}^l,\tilde{z}^l).$$

Step 2.1. Test for Convergence of Inner Iterations. If

$\|F(x_k,y_k,z_k;c_k,\mu^l)\|\le\eta\mu^l$ and $\|g(x)\|^2\le\epsilon_g$, then set $(\tilde{x}^{l+1},\tilde{y}^{l+1},\tilde{z}^{l+1})=(x_k,y_k,z_k)$ and go to Step 3.

Step 2.2. Solve the Newton system (4) to obtain $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k)$.

Step 2.3. Penalty Parameter Selection. If

$\Delta x_k^T \nabla f_k - c_k \|g_k\|_2^2 - \mu^l \Delta x_k^T X_k^{-1} e + \|\Delta x_k\|_{H_k}^2 > 0$ and $\|g_k\|^2 > \epsilon_g$, then set

$$c_{k+1} = \max\left\{[\Delta x_k^T \nabla f_k - \mu^l \Delta x_k^T X_k^{-1} e + \|\Delta x_k\|_{H_k}^2]/\|g_k\|_2^2,\ c_k + \delta\right\}.$$

Otherwise, set $c_{k+1} = c_k$.

Step 2.4. Steplength Selection Rules.

Set $\alpha_{xk}^{\max} = \min_{1 \le i \le n} \left\{ -x_k^i / \Delta x_k^i : \Delta x_k^i < 0 \right\}$.

If $\Delta x_k^T \nabla f_k - c_k \|g_k\|_2^2 - \mu^l \Delta x_k^T X_k^{-1} e + \|\Delta x_k\|_{H_k}^2 > 0$ and $0 < \|g_k\|^2 \le \epsilon_g$, then set

$\alpha_{zk}^{\max} = \min_{1 \le i \le n} \left\{ \left(-z_k^i / \Delta z_k^i \right) | \Delta z_k^i < 0 \right\}$,

$\hat{\alpha}_k = \min\left\{1, \gamma \alpha_{xk}^{\max}, \gamma \alpha_{zk}^{\max}\right\}$.

Set $\alpha_k = \beta^\theta \hat{\alpha}_k$, where θ is the smallest nonnegative integer such that

$$\|F(w_k + \alpha_k \Delta w_k)\|^2 - \|F(w_k)\|^2 \le \rho \alpha_k (\nabla F^t(w_k) F(w_k), \Delta w_k).$$

Set $w_{k+1} = w_k + \alpha_k \Delta w_k$.

Otherwise, set $\hat{\alpha}_{xk} = \min\{\gamma \alpha_{xk}^{\max}, 1\}$ and $\alpha_{xk} = \beta^\theta \hat{\alpha}_{xk}$, where θ is the smallest nonnegative integer such that

$$\Phi\left(x_{k+1}; c_{k+1}, \mu^l\right) - \Phi\left(x_k; c_{k+1}, \mu^l\right) \le \rho \alpha_{xk} \nabla \Phi\left(x_k; c_{k+1}, \mu^l\right)^T \Delta x_k.$$

Set $LB_k^i = \min\left\{(1/2) m\mu, x_{k+1}^i z_k^i\right\}$ and $UB_k^i = \max\left\{2M\mu, x_{k+1}^i z_k^i\right\}$.

For $i = 1, 2, \ldots, n$, find

$\alpha_{zk}^i = \max\{\alpha^i : LB_k^i \le x_{k+1}^i \left(z_k^i + \alpha^i \Delta z_k^i\right) \le UB_k^i\}$.

Set $\alpha_{zk} = \min\left\{1, \min_{1 \le i \le n} \{\alpha_{zk}^i\}\right\}$.

Set $x_{k+1} = x_k + \alpha_{xk} \Delta x_k, \quad y_{k+1} = y_k + \alpha_{zk} \Delta y_k,$
$z_{k+1} = z_k + \alpha_{zk} \Delta z_k$.

Step 2.5. Set $k=k+1$ and go to Step 2.1.

Step 3. Reduce the barrier parameter as described in Section 2.4.
Step 4. Set $l=l+1$ and go to Step 1.

2.1. Penalty Parameter Selection Rule. At every iteration, the value of the penalty parameter is determined such that a descent property is ensured for the merit function $\Phi(x; c, \mu)$. The direction Δx_k is a descent for Φ at the current point x_k if

$$\Delta x_k^T \nabla \Phi(x_k; c_k, \mu) \leq 0. \tag{7}$$

By considering the second equation of the Newton system (4), we have

$$\Delta x_k^T \nabla \Phi(x_k; c_k, \mu) = \Delta x_k^T \nabla f(x_k) - c_k \|g_k\|^2 - \mu \Delta x_k^T X_k^{-1} e, \tag{8}$$

where c_k is the value of the penalty parameter at the beginning of the kth iteration. Since the barrier parameter μ is fixed, we can deduce from (8) that, if c_k is not large enough, then the descent property (7) may not be satisfied. Thus, a new value $c_{k+1} > c_k$ must be determined to guarantee that (7) holds. The next lemmas show that Algorithm 2.1 chooses the value of the penalty parameter in such a way that Δx_k is a descent direction for the merit function.

In Lemmas 2.1 and 2.4, we show that the descent is guaranteed always if $\|g(x)\|^2 > \epsilon_g$ or $g(x)=0$, where ϵ_g denote a finite precision. As a result, the penalty parameter $c_k = c_k(\epsilon_g)$ remains finite. On the other hand, at some inner iteration k, if we have $0 < \|g(x_k)\|^2 \leq \epsilon_g$ and the descent condition (7) is not satisfied, then a switch to the following merit function:

$$\|F(x, y, z; c, \mu)\|^2 \tag{9}$$

is performed for all consecutive inner iterations. Once the convergence of the inner iteration is achieved, the algorithm returns to minimizing the merit function (6). This is a variation of the so called watchdog technique, which was suggested first by Chamberlain et al. Ref. 9. In the context of interior-point methods, it was also used by Gay et al. in Ref. 4. The convergence criteria for (9) have been well established (Refs. 2, 10).

Lemma 2.1. Let f and g be differentiable functions and let there exist a small $\epsilon_g > 0$ such that $\|g_k\|^2 > \epsilon_g$. If Δx_k is calculated by solving the Newton system (4) and c_{k+1} is chosen as in Step 2.3 of Algorithm 2.1, then Δx_k is a descent direction for the merit function Φ at the current point x_k. Furthermore,

$$\Delta x_k^T \nabla \Phi(x_k; c_{k+1}, \mu) \leq -\|\Delta x_k\|_{H_k}^2 \leq 0. \tag{10}$$

Proof. In Step 2.3, Algorithm 2.1 initially checks the inequality

$$\Delta x_k^T \nabla f_k - c_k \|g_k\|^2 - \mu \Delta x_k^T X_k^{-1} e + \|\Delta x_k\|_{H_k}^2 \leq 0. \tag{11}$$

If (11) is satisfied, then by setting $c_{k+1} = c_k$ and rearranging (11), we obtain (10). On the other hand, if (11) is not satisfied, by setting

$$c_{k+1} = \max\left\{\left[\Delta x_k^T \nabla f_k - \mu \Delta x_k^T X_k^{-1} e + \|\Delta x_k\|_{H_k}^2\right]/\|g_k\|^2, c_k + \delta\right\}, \quad \delta > 0,$$

and substituting it into (8), it can be verified that (10) holds also. □

Remark 2.1. The role of the parameter δ in the definition of c_{k+1} is to guarantee that the penalty parameter increases by at least a certain amount every time it needs to be updated.

In the previous lemma, it is assumed that $\|g_k\|^2 > \epsilon_g$. The next lemma demonstrates that Δx_k remains a descent direction for the merit function Φ when $g_k = 0$.

Lemma 2.2. Let f and g be differentiable functions and let $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k)$ be the Newton direction taken by solving the system (4). If $g_k = 0$, for some or all iterations k, then the descent property (10) is satisfied for any choice of the penalty parameter $c_k \in [0, \infty)$.

Proof. If $g_k = 0$, then (8) yields

$$\Delta x_k^T \nabla \Phi(x_k; c_k, \mu) = \Delta x_k^T \nabla f_k - \mu \Delta x_k^T X_k^{-1} e \tag{12}$$

and the second equation of the Newton system (4) becomes

$$\nabla g_k \Delta x_k = 0. \tag{13}$$

Furthermore, solving the third equation of (4) for Δz_k, we have

$$\Delta z_k = -X_k^{-1} Z_k \Delta x_k - z_k + \mu X_k^{-1} e. \tag{14}$$

Substituting Δz_k into first equation of (4) yields

$$\nabla f_k + c_k \nabla g_k^T g_k - \mu X_k^{-1} e = -\left(H_k + X_k^{-1} Z_k\right) \Delta x_k + \nabla g_k^T (y_k + \Delta y_k). \tag{15}$$

Premultiplying (15) by Δx_k^T and using (12) and (13) yields

$$\begin{aligned}\Delta x_k^T \nabla \Phi(x_k; c_k, \mu) &= -\Delta x_k^T \left(H_k + X_k^{-1} Z_k\right) \Delta x_k \\ &< -\Delta x_k^T H_k \Delta x_k. \end{aligned} \tag{16}$$

From (16), it is derived that (10) holds for every $c_k \in [0, \infty)$. □

Lemma 2.3. Let the assumptions of the previous lemma hold and let $g_k = 0$ for some k. Then, the algorithm chooses $c_{k+1} = c_k$ in Step 2.3. Also, Δx_k is still a descent direction for the merit function Φ at x_k.

Proof. In the previous lemma, it was proved that the descent property (10) is satisfied for $g_k = 0$. Basically, this means that the condition in Step 2.3 of Algorithm 2.1 is satisfied always. Consequently, the algorithm does not need to increase the value of the penalty parameter and simply sets $c_{k+1} = c_k$. For this choice of the penalty parameter, it can be verified that the descent property (10) still holds. □

Corollary 2.1. If $\|\Delta x_k\| = 0$, then the algorithm chooses $c_{k+1} = c_k$.

Lemma 2.4. Let f and g be continuously differentiable functions and

$$\Delta x_k^T \nabla f_k - \mu \Delta x_k^T X_k^{-1} e + \|\Delta x_k\|_{H_k}^2 \leq \mathcal{M}^* < \infty.$$

Then, for μ fixed, the following properties hold:

(i) There exists always a constant $c_{k+1} \geq 0$ satisfying Step 2.3 of Algorithm 2.1.

(ii) Assuming that the sequence $\{x_k\}$ is bounded, c_k is increased finitely often; that is, there exists an integer $k_* \geq 0$ such that, for all $k \geq k_*$, we have $c_* \in [0, \infty)$.

Proof. Part (i) is a direct consequence of Lemmas 2.1–2.3 and Corollary 2.1, since a finite value c_{k+1} is always generated in Step 2.3. Part (ii) will be shown by contradiction. Assume that $c_k \to \infty$ as $k \to \infty$. From the way c_{k+1} is defined in Step 2.3, we can deduce that, if $c_k \to \infty$, then $\|g_k\|^2 \to 0$. Hence, there exists an integer k_1 such that, for all $k \geq k_1$, we have

$$0 < \|g_k\|^2 \leq \epsilon_g.$$

As can be seen in Step 2.4, however, in the case where $0 < \|g_k\|^2 \leq \epsilon_g$, the algorithm stops increasing the penalty parameter, since it switches to the second merit function. Therefore, the maximum value that c_k can take is

$$c_* = c_{k_1} = \mathcal{M}^*/\epsilon_g,$$

where $\mathcal{M}^*$ and c_* are finite values. Hence, we have $c_* < \infty$. This contradicts our assumption that $c_k \to \infty$ as $k \to \infty$. Hence, the penalty parameter does not increase indefinitely; that is, there exists an integer $k_* \geq 0$ such that, for all $k \geq k_*$, we have $c_k < \infty$. □

2.2. Primal Stepsize Rule. In Step 2.4 of the algorithm, we adopt the Armijo rule to determine the new iterate x_{k+1}. The maximum allowable stepsize is determined by the boundary of the feasible region and is given by

$$\alpha_{xk}^{\max} = \min_{1 \leq i \leq n} \{x_k^i / (-\Delta x_k^i) : \Delta x_k^i < 0\}.$$

We take as initial step $\hat{\alpha}_{xk}$ a number very close to $\alpha_{xk}^{\max}$ and we ensure that it is never greater than one, i.e.,

$$\hat{\alpha}_{xk} = \min\{\gamma \alpha_{xk}^{\max}, 1\}, \qquad \text{with } \gamma \in (0, 1).$$

The final step is

$$\alpha_{xk} = \beta^\theta \hat{\alpha}_{xk},$$

where θ is the first nonnegative integer for which the Armijo rule is satisfied and the factor β is usually chosen in the interval [0.1, 0.5], depending on the confidence we have on the initial step $\hat{\alpha}_{xk}$.

2.3. Dual Stepsize Rule. In this section, we discuss the calculation of the stepsize of the dual variables z. The strategy uses the information provided by the new primal iterate x_{k+1} in order to find the new iterate z_{k+1}. It is a modification of the strategy suggested by Yamashita (Ref. 3) and Yamashita and Yabe (Ref. 11).

While the barrier parameter μ is fixed, we determine a step α_{zk}^i along the direction Δz_k^i, for each dual variable z_k^i, $i = 1, 2, \ldots n$, such that the box constraints are satisfied,

$$\alpha_{zk}^i = \max\{\alpha > 0 : LB_k^i \leq (x_k^i + \alpha_{xk} \Delta x_k^i)(z_k^i + \alpha \Delta z_k^i) \leq UB_k^i\}.$$

The lower bounds LB_k^i and upper bounds $UB_k^i, i=1,2,\dots,n$, are defined as

$$LB_k^i = \min\{(1/2)m\mu,\, (x_k^i + \alpha_{xk}\Delta x_k^i)z_k^i\},$$
$$UB_k^i = \max\{2M\mu,\, (x_k^i + \alpha_{xk}\Delta x_k^i)z_k^i\},$$

where the parameters m and M are chosen such that

$$0 < m \le \min\{1, [(1-\gamma)(1-\gamma/(M_0)^{\mu})\min_i\{x_k^i z_k^i\}]/\mu\},$$
$$M \ge \max\{1, \max_i\{x_k^i z_k^i\}/\mu\} > 0,$$

with $\gamma \in (0,1)$ and M_0 a positive large number. The common dual steplength α_{zk} is defined as

$$\alpha_{zk} = \min\{1, \min_{1\le i\le n}\{\alpha_{zk}^i\}\}.$$

Also, the stepsize for the dual variables y can be either $\alpha_{yk} = 1$ or $\alpha_{yk} = \alpha_{zk}$.

2.4. Barrier Parameter Selection Rule. The basic characteristic of our barrier reduction strategy is that it determines the new value of μ by taking into consideration the distance of the current point (x_k, y_k, z_k) from the central path and he optimum solution of the initial problem. The barrier reduction strategy is defined as follows:

$$\mu^{l+1} = \min\{0.95\mu^l, 0.01(0.95)^k \|F(x_k, y_k, z_k)\|_2\}.$$

If $\|F(x_k, y_k, z_k; \mu^l)\|_2 \le 0.1\eta\mu^l$, then:

if $\mu^l < 10^{-4}$, then $\mu^{l+1} = \min\{0.85\mu^l, 0.01(0.85)^{k+2\sigma}\|F(x_k, y_k, z_k)\|_2\}$;
else, $\mu^{l+1} = \min\{0.85\mu^l, 0.01(0.85)^{k+\sigma}\|F(x_k, y_k, z_k)\|_2\}$.

3. Global Convergence

In this section, we show that the algorithm is globally convergent, because it guarantees always progress toward a solution from any starting point.

We show that, while the barrier parameter is fixed to a value μ^l, the algorithm produces iterates

$$w_k(\mu^l) = (x_k(\mu^l), y_k(\mu^l), z_k(\mu^l)), \qquad \text{for } k \ge 0,$$

which are bounded and converge to a point

$$w_*(\mu^l) = (x_*(\mu^l), y_*(\mu^l), z_*(\mu^l))$$

such that

$$\|F(x_*(\mu^l), y_*(\mu^l), z_*(\mu^l); c_*, \mu^l)\| = 0,$$

where $F(x, y, z; c, \mu)$ is the vector of the perturbed KKT conditions, defined in (3). In other words, we show that the inner steps 2.1–2.5 of Algorithm 2.1 converge to a solution of the perturbed KKT conditions. For simplicity, we suppress the index l, and we use w_k instead of $w_k(\mu^l)$ to denote the iterates produced while $\mu = \mu^l$.

The basic result of Lemmas 2.1 to 2.4 is that the direction Δx_k, taken from the solution of the Newton system (4), is a descent direction for the merit function Φ at the current point x_k; that is the inequality (10) holds. In the next theorem, we show that the sequence $\{\Phi(x_k; c_*, \mu)\}$ is monotonically decreasing if the barrier parameter μ is fixed. We show also that the step α_{xk} chosen in Step 2.4 is always positive.

Theorem 3.1. Assume that the following conditions hold:

(i) The objective function f and the constraints g are twice continuously differentiable.

(ii) For every iteration k and every vector $v \in \Re^n$, there exist constants $M' > m' > 0$, such that $m'\|v\|_2^2 \leq v^T H_k v \leq M'\|v\|_2^2$.

(iii) For every k, there exists a vector $(\Delta x_k, \Delta y_k, \Delta z_k)$ as a solution of (4).

(iv) There exists an iteration k_*, small $\epsilon_g > 0$, $\|g_k\|^2 \notin (0, \epsilon_g)$, and a scalar $c_* \geq 0$, with $c_* = c_*(\epsilon_g)$, such that the penalty parameter restriction in Step 2.3,

$$\Delta x_k^T \nabla f_k - c_k(\epsilon_g)\|g_k\|_2^2 - \mu \Delta x_k^T X_k^{-1} e + \|\Delta x_k\|_{H_k}^2 \leq 0,$$

is satisfied for all $k \geq k_*$, with

$$c_{k+1}(\epsilon_g) = c_k(\epsilon_g) = c_*(\epsilon_g).$$

Then, the stepsize computed in Step 2.4 is such that $\alpha_{xk} \in (0, 1]$, hence, the sequence $\{\Phi(x_k; c_*, \mu)\}$ is monotonically decreasing for $k \geq k_*$ and μ fixed.

Proof. Consider the case $\|g_k\|^2 \notin (0, \epsilon_g)$ and the first-order approximation with remainder of the function $\Phi(x; c_*, \mu)$ around the point $x_{k+1} = x_k + \alpha_{xk} \Delta x_k$,

$$\begin{aligned} &\Phi(x_{k+1}; c_*, \mu) - \Phi(x_k; c_*, \mu) \\ &\leq \alpha_{xk} \Delta x_k^T \nabla \Phi(x_k; c_*, \mu) + (\alpha_{xk}^2/2) \Delta x_k^T H_k \Delta x_k + \alpha_{xk}^2 \psi_k \|\Delta x_k\|_2^2, \end{aligned} \tag{17}$$

where

$$\psi_k = \int_0^1 (1-t) \|\nabla_x^2 \Phi(x_k + t\alpha_{xk} \Delta x_k; c_*, \mu) - H_k\|_2 \, dt.$$

Furthermore, from Assumption (A2) and Lemmas 2.1 and 2.2, we have

$$\begin{aligned} \|\Delta x_k\|_2^2 &\leq (1/m) \Delta x_k^T H_k \Delta x_k \\ &\leq -\Delta x_k^T \nabla \Phi(x_k; c_*, \mu). \end{aligned} \tag{18}$$

Using (18) in (17), we obtain

$$\begin{aligned} &\Phi(x_{k+1}; c_*, \mu) - \Phi(x_k; c_*, \mu) \\ &\leq \alpha_{xk} \Delta x_k^T \nabla \Phi(x_k; c_*, \mu)(1 - \alpha_{xk}(1/2 + \psi_k/m)). \end{aligned} \tag{19}$$

The scalar ρ in Step 2.4 determines a steplength α_{xk} such that

$$\rho \leq 1 - \alpha_{xk}(1/2 + \psi_k/m) \leq 1/2.$$

From Lemmas 2.1 to 2.4, since we have always

$$\Delta x_k^T \nabla \Phi(x_k; c_*, \mu) \leq 0,$$

there must exist $\alpha_{xk} \in (0, 1]$, to ensure (19) and the Armijo rule in Step 2.4. Assume that α^0 is the largest step in the interval (0, 1] satisfying both (19) and the Armijo rule. Consequently, for every $\alpha \leq \alpha^0$, inequality (19) and the Armijo rule are also satisfied. Hence, the strategy in Step 2.4 selects always a steplength $\alpha_{xk} \in [\beta\alpha^0, \alpha^0]$, where $0 < \beta \leq 1$. From the above analysis, it follows that the sequence $\{\Phi(x_k; c_*, \mu)\}$ is monotonically decreasing. □

Remark 3.1. The results of the above theorem can be proved to hold before the penalty parameter c_k achieves a constant value c_*. This can be done by considering the difference $\Phi(x_{k+1}; c_{k+1}, \mu) - \Phi(x_k; c_{k+1}, \mu)$ and the Taylor expansion of the function $\Phi(x; c_{k+1}, \mu)$ instead of $\Phi(x; c_*, \mu)$. In the above theorem, we choose to prove the case where $c_k = c_*$ has been achieved, in order to show that, asymptotically, Φ is monotonically decreasing and the strategy in Step 2.4 selects a steplength $\alpha_{xk} \in (0, 1]$.

Corollary 3.1. The sequence $\{x_k\}$ of primal variables generated by Algorithm 2.1, with μ fixed, is bounded away from zero.

The following lemma, proved by Yamashita in Ref. 3, shows that the dual stepsize rule, used by Algorithm 2.1, generates iterates z_k which are also bounded above and away from zero.

Lemma 3.1. While μ is fixed, the lower bounds LB_k^i and the upper bounds $UB_k^i, i=1,2,\ldots,n$, of the box constraints in the dual stepsize rule are bounded away form zero and bounded from above, respectively, if the corresponding components x_k^i of the iterates x_k are also bounded above and away from zero.

Proof. The proof can be found in Ref. 3. □

Having established that the sequences of iterates $\{x_k\}$ and $\{z_k\}$ are bounded above and away from zero, we show that the iterates $\{y_k\}, k \geq 0$, are also bounded. In particular, Lemma 3.3 shows that, if at each iteration of the algorithm we take a unit step along the direction Δy_k, then the resulting sequence $\{y_k + \Delta y_k\}$ is bounded. In addition to this, Lemma 3.3 shows also that the Newton direction $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k)$ is bounded, while μ is fixed. We first establish the following technical result.

Lemma 3.2. Let w_k be a sequence of vectors generated by Algorithm 2.1 for μ fixed Then, the matrix sequence $\{\Theta_k^{-1}\}$ is bounded, where

$$\Theta_k = \begin{pmatrix} 0 & \nabla g_k \\ -\nabla g_k^T & H_k + X_k^{-1} Z_k \end{pmatrix}.$$

Proof. The proof can be found in Ref. 12. □

Lemma 3.3. Let w_k is a sequence of vectors generated by Algorithm 2.1 for μ fixed. Then, the sequence of vectors $\{(\Delta x_k, y_k + \Delta y_k, \Delta z_k)\}$ is bounded.

Proof. The proof can be found in Ref. 12. □

Lemmas 3.4 and 3.5 provide the necessary results by Theorem 3.2, which shows that the sequence of $\{w_k\}$ converges to a point $w_* = (x_*, y_*, z_*)$ satisfying the KKT conditions of problem (2).

Lemma 3.4. Let the assumptions of Theorem 3.1 be satisfied and let the barrier parameter μ be fixed. Suppose that also, for some iteration $k_0 \geq 0$, the level set

$$S_1 = \{x \in \Re^n_+ : \Phi(x; c_*, \mu) \leq \Phi(x_{k_0}; c_*, \mu)\} \tag{20}$$

is compact. Then, for all $k \geq k_0$, we have

$$\lim_{k\to\infty} \Delta x_k^T \nabla\Phi(x_k; c_*, \mu) = 0. \tag{21}$$

Proof. The scalar $\rho \in (0, 1/2)$ in the Armijo stepsize strategy at Step 2.4 determines a stepsize α_{xk} such that

$$\rho \leq 1 - \alpha_{xk}(1/2 + \psi_k/m') \leq 1/2.$$

By solving for α_{xk}, we obtain

$$(1/2)/(1/2 + \psi_k/m') \leq \alpha_{xk} \leq (1-\rho)/(1/2 + \psi_k/m').$$

Hence, the largest value that α_{xk} can take and still satisfy the Armijo rule in Step 2.4 is

$$\alpha^0_{xk} = \min\{1, (1-\delta)/(1/2 + \psi_k/m')\}.$$

Recall that the steplength α_{xk} is chosen by reducing the maximum allowable steplength $\hat{\alpha}_{xk}$ until the Armijo rule is satisfied. Therefore, $\alpha_{xk} \in [\beta\alpha^0_{xk}, \alpha^0_{xk}]$ and thereby satisfies also the Armijo rule.

As the merit function Φ is twice continuously differentiable and the level set S_1 is bounded, there exists a scalar $\bar{M} < \infty$ such that

$$\psi_k = \int_0^1 (1-t)\|\nabla_x^2\Phi(x_k + t\alpha_{xk}\Delta x_k; c_*, \mu) - H_k\|_2 dt \leq \bar{M} < \infty.$$

Thus, we have always

$$\alpha_{xk} \geq \bar{\alpha}_{xk} > 0,$$

where

$$\bar{\alpha}_{xk} = \min\{1, (1-\rho)/(1/2 + M/m')\}.$$

Hence, the stepsize α_{xk} is always bounded away from zero. Furthermore, from the Armijo rule and Lemmas 2.1 and 2.2, we have

$$\Phi(x_{k+1}; c_*, \mu) - \Phi(x_k; c_*, \mu) \leq \rho\alpha_{xk}\nabla\Phi(x_k; c_*, \mu)^T \Delta x_k < 0. \tag{22}$$

From our assumption that the level set S_1 is bounded, it can be deduced that the sequence

$$\{|\Phi(x_{k+1}; c_*, \mu) - \Phi(x_k; c_*, \mu)|\} \to 0.$$

Consequently, from (22), we have

$$\lim_{k\to\infty} (\rho\alpha_{xk} \nabla\Phi(x_k; c_*, \mu)^T \Delta x_k) = 0.$$

Finally, since $\rho, \alpha_{xk} > 0$, it can be deduced that (21) holds. □

Lemma 3.5. Let the assumptions of the previous lemma hold. Then,

$$\lim_{k\to\infty} \|\Delta x_k\|^2_{H_k} = 0. \tag{23}$$

Proof. From (10) we have

$$-\nabla\Phi(x_k; c_*, \mu)^T \Delta x_k \geq \|\Delta x_k\|^2_{H_k}.$$

Using (21), we have that (23) holds. □

Theorem 3.2. Let the assumptions of the previous lemma hold and let ϵ_g be a sufficiently small positive scalar[4]. Then, the algorithm converges in the limit to a point satisfying $F(x, y, z; c, \mu) = 0$ for μ fixed.

Proof. Consider the case where $\|g_k\|^2 \notin (0, \epsilon_g)$ and let $x_*(\mu), z_*(\mu) \in \Re^n$, $y_*(\mu) \in \Re^q$ be such that

$$\lim_{k\to\infty} x_k = x_*(\mu), \quad \lim_{k\to\infty} z_k = z_*(\mu), \quad \lim_{k\to\infty} y_k = y_*(\mu),$$
$$\forall k \geq k_*, \quad k \in K \subseteq \{1, 2, \ldots\}.$$

The existence of such points is ensured since by Assumption A2 and Lemmas 3.1 and 3.3, the sequence $\{(x_k(\mu), y_k(\mu), z_k(\mu))\}$ is bounded for μ fixed and by Theorem 3.1 the algorithm decreases always the merit function Φ sufficiently at each iteration, thereby ensuring that $x_k \in S_1$, with S_1 compact.

We first prove that, for k sufficiently large, the dual step α_{zk} becomes the unit one, by showing that

$$\lim_{k\to\infty} \|z_k + \Delta z_k - \mu X^{-1}_{k+1} e\| = 0. \tag{24}$$

[4]In our numerical experiments, we used $\epsilon_g = 10^{-8}$.

Adding $-\mu X_{k+1}^{-1} e$ to both sides of (14), we have

$$\|z_k + \Delta z_k - \mu X_{k+1}^{-1} e\| \leq \| - X_k^{-1} Z_k\| \|\Delta x_k\| + \mu \|X_k^{-1} - X_{k+1}^{-1}\| \|e\|. \quad (25)$$

Moreover,

$$\|X_k^{-1} - X_{k+1}^{-1}\|^2 \leq n \max_{1\leq i\leq n} \{(\alpha_{xk} \Delta x_k^i)^2)/(x_k^i x_{k+1}^i)^2\}.$$

Since we have always

$$\alpha_{xk} \in (0, 1], \quad (\Delta x_k^i)^2 \leq \|\Delta x_k\|^2,$$

and since the sequence $\{x_k\}$ bounded away from zero, from the above inequality and (23), we can derive that

$$\lim_{k\to\infty} \|X_k^{-1} - X_{k+1}^{-1}\|^2 \leq n \lim_{k\to\infty} \max_{1\leq i\leq n} \{\|\Delta x_k\|^2/(x_k^i x_{k+1}^i)^2\} = 0. \quad (26)$$

Hence, letting $k \to \infty$ in (25) and using (23) and (26), we can deduce that (24) holds. Consequently,

$$z_{k+1} = z_k + \Delta z_k,$$

for k sufficiently large.

Furthermore, using (14) for k sufficiently large, the complementarity condition becomes

$$\begin{aligned} X_{k+1} z_{k+1} &= X_{k+1}(z_k + \Delta z_k) \\ &= X_{k+1} X_k^{-1}(-Z_k \Delta x_k + \mu e). \end{aligned} \quad (27)$$

From (23) and the fact that the elements of the diagonal matrix $X_{k+1} X_k^{-1}$ can be written as

$$x_{k+1}^i / x_k^i = 1 + \alpha_{xk} \Delta x_k^i / x_k^i, \quad \text{for all } i = 1, 2, \ldots, n,$$

we can derive that

$$\lim_{k\to\infty} X_{k+1} X_k^{-1} = I_n, \quad (28)$$

where I_n is the $n \times n$ identity matrix. Letting $k \to \infty$ in (27) and using (23) and (28) yields

$$\lim_{k\to\infty} X_{k+1} z_{k+1} = X_*(\mu) z_*(\mu) = \mu e. \quad (29)$$

Also, for $k\to\infty$, the second equation of the Newton system (4) and (23) yield

$$\lim_{k\to\infty}(\nabla g_k \Delta x_k)=g(x_*(\mu))=0. \tag{30}$$

The first equation of the Newton system (4) can be written as

$$\nabla f_k-\nabla g_k^T y_{k+1}+c_*\nabla g_k^T g_k-\mu X_k^{-1}e=-(H_k+X_k^{-1}Z_k)\Delta x_k,$$

where

$$y_{k+1}=y_k+\Delta y_k.$$

Letting $k\to\infty$ and using (23), the above equation yields

$$\lim_{k\to\infty}\|\nabla f_k-\nabla g_k^T y_{k+1}+c_*\nabla g_k^T g_k-\mu X_k^{-1}e\|=0. \tag{31}$$

From the assumptions that the functions f and g have continuous gradients and ∇g_k^T has full column rank and using (26), equation (31) yields

$$\lim_{k\to\infty}\|\nabla f_{k+1}-\nabla g_{k+1}^T y_{k+1}+c_*\nabla g_{k+1}^T g_{k+1}-\mu X_{k+1}^{-1}e\|=0,$$

or equivalently,

$$\nabla f(x_*(\mu))-\nabla g(x_*(\mu))^T y_*(\mu)+c_*\nabla g(x_*(\mu))^T g(x_*(\mu))-\mu X_*(\mu)^{-1}e=0. \tag{32}$$

From (32), (30), and (29), we can conclude that the vector $(x_*(\mu),y_*(\mu),z_*(\mu))$ is a solution of the perturbed KKT conditions (3).

The convergence result for $\|g_k\|^2\in(0,\epsilon_g)$ is a consequence of El-Bakry et al. (Ref. 2) and Zakovic et al. (Ref. 10). □

An immediate consequence of Theorem 3.2 is that, for any convergent subsequence produced by the algorithm for $\mu=\mu^l$, there is an iteration $\tilde{k}$, such that

$$\|F(x_{\tilde{k}},y_{\tilde{k}},z_{\tilde{k}};c_*,\mu)\|\le\eta\mu, \tag{33}$$

for all $k\ge\tilde{k}$, where $\eta\ge 0$ and $F(x,y,z;c,\mu)$ is given by (3). At this point, we record the value of the current iterate

$$(\tilde{x}^l,\tilde{y}^l,\tilde{z}^l)=(x_{\tilde{k}},y_{\tilde{k}},z_{\tilde{k}}),$$

and set μ to a smaller value $\mu^{l+1}<\mu^l$. Therefore, a sequence of approximate central points $\{(\tilde{x}^l,\tilde{y}^l,\tilde{z}^l)\}$ is generated.

In the remaining part of this section, we show that the sequence of approximation central points converges to a KKT point $\{\tilde{x}^*, \tilde{y}^*, \tilde{z}^*\}$ of the initial constrained optimization problem (1).

For a given $\epsilon \geq 0$ sufficiently small, consider the set of all the approximate central points generated by Algorithm 2.1,

$$S_2(\epsilon)=\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l): \epsilon \leq \|F(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l; c_*, \mu^l)\| \\ \leq \|F(\tilde{x}^0, \tilde{y}^0, \tilde{z}^0; c_*, \mu^0)\|, \forall \mu^l < \mu^0\}.$$

If $\epsilon > 0$, then the stepsize rules described in Section 2 guarantee that $\tilde{x}^l, \tilde{z}^l \in S_2(\epsilon)$ are bounded away from zero for $l \geq 0$. Consequently, $(\tilde{x}^l)^T \tilde{z}^l$ is also bounded away from zero in $S_2(\epsilon)$. The following lemma shows that the sequence $\{\tilde{y}^l\}$ is bounded if the sequence $\{\tilde{z}^l\}$ is also bounded.

Lemma 3.6. Assuming that the columns of $\nabla g(\tilde{x}^l)$ are linearly independent and the iterates $\tilde{x}^l$ are in a compact set for $l \geq 0$, then there exists a constant $M_1 > 0$ such that

$$\|\tilde{y}^l\| \leq M_1(1+\|\tilde{z}^l\|).$$

Proof. The proof can be found in Ref. 12. □

Lemma 3.7. If $(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) \in S_2(\epsilon)$ for all $l \geq 0$, then the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ is bounded above.

Proof. The proof can be found in Ref. 12. □

The following theorem shows that the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ converges to $\{(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)\}$ which is a KKT point of the initial constrained problem (1).

Theorem 3.3. Let $\{\mu^l\}$ be a positive monotonically decreasing sequence of barrier parameters with $\{\mu^l\} \to 0$ and let $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ be a sequence of approximate central points satisfying (33) for $\mu = \mu^l, l \geq 0$. Then, the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ is bounded its limit point $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ satisfies $F(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*; c, \mu)=0$, for $\mu=0$.

Proof. From Lemma 3.6, the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ is bounded and remains in the compact set $S_2(\epsilon)$. Thus, it has a limit point in $S_2(\epsilon)$, denoted as $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$. From (33) and the fact that $\mu^l \to 0$, we obtain easily that $\lim_{l\to\infty} \|F(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\| = 0$. Therefore,

$$\nabla f(\tilde{x}^*) - \tilde{z}^* + c\nabla g(x)^T g(x) - \nabla g(\tilde{x}^*)^T \tilde{y}^* = 0,$$
$$g(\tilde{x}^*) = 0,$$
$$\tilde{X}^* \tilde{Z}^* e = 0.$$

Clearly, from the above equations, we may derive that $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ is a KKT point of the initial constrained optimization problem (1). □

4. Numerical Results

The test problems solved by the algorithm fall into two categories. The first category consists of small-size problems drawn mainly from the Hock and Schittkowski collection (Ref. 13). The second category consists of real -world problems of larger size, which were taken from the Vanderbei website (Ref. 14). The implementation of the algorithm has been done using standard C and is interfaced with the powerful mathematical programming language.

The various parameters used in the algorithm are as follows. In Step 1, the accuracy of the stopping criterion is $\epsilon_0 = 10^{-8}$. In Step 2.3, $\delta = 10$ and $\epsilon_g = 10^{-8}$. In Step 2.4, we set $\gamma = 0.995, \beta = 0.5, \rho = 10^{-4}, m = 1$, and $M = 10$. In the barrier reduction rule, described in Section 2.4, we set $\sigma = 6$.

Tables 1 and **??** summarize the numerical results for the test-problems taken from the Hock and Schittkowski collection. The following terminology is used:

Problem: problem number given in the Hock and Schittkowski collection (Ref. 13);

Iterations: total number of inner iterations required to find the optimum solution of the original problem (1);

c_*: final value of the penalty parameter.

All the numerical results have been obtained by using the exact Hessian, provided by AMPL, except those marked with $*$ (i.e., tests 74, 75, 93 and 97), which were solved using the BFGS updating formula (Ref. 15). In those iterations where the exact Hessian H_k is not positive semidefinite, it is replaced it by

$$\hat{H}_k = H_k + E_k,$$

where E_k is a nonnegative diagonal matrix that is zero if H_k is positive definite. The matrix $\hat{H}_k$ is generated by the modified Cholesky factorization as described in Gill et al. (Ref. 16). This technique has

Table 1. Numerical results for the Hock–Schitkowski problems.

Problem	Iterations	c_*	Problem	Iterations	c_*
1	27	0	29	47	5.8×10^{11}
2	15	0	30	10	1.6×10^{11}
3	10	0	31	9	1.7×10^{11}
4	10	0	32	45	0
5	10	0	33	10	0
6	6	0	34	19	3.4×10^{11}
7	9	2.3×10^{6}	35	12	0
8	5	0	36	20	0
9	5	0	37	57	175.4
10	14	1.2×10^{12}	38	25	0
11	11	1.9×10^{12}	39	8	5.3×10^{6}
12	15	3.8×10^{11}	40	25	5.8×10^{5}
14	11	2.3×10^{11}	41	36	1.2×10^{11}
15	22	3.6×10^{6}	42	13	2.1×10^{7}
16	13	6.6×10^{12}	43	13	0
17	55	2.3×10^{12}	44	23	0
18	12	4.4×10^{12}	45	22	0
19	21	8.1×10^{12}	46	26	33.69
20	16	6.7×10^{11}	47	20	71.8
21	10	0	48	5	0
22	12	5.3×10^{12}	49	19	0
23	20	3.6×10^{12}	50	9	0
24	19	9.7×10^{10}	51	5	0
25	31	0	52	4	0
26	21	10	53	9	0
27	9	7065	54	13	0
28	4	0	55	11	4.1×10^{12}

worked very well in practice enabling the algorithm to solve large problems.

The algorithm solved all the problems to the given accuracy. For all of the problems, the initial value c_0 of the penalty parameter c is set to zero. Its final value c is usually kept at low levels. However, for some tests, its final value needed to become large in order to achieve convergence, which was achieved in all tests. We have observed also that usually the penalty parameter becomes large for the problems whose starting point, provided by Hock and Schittkowski (Ref. 13), is close to the boundary of the feasible region. In such cases, Vanderbei and Shanno (Ref. 5) suggest that the starting point should be set to a 90%-10% mixture of the two bounds of the box constraints, with the higher value placed on the nearer bound. If the algorithm starts from such a point, the final value of the penalty parameter can be kept low, with similar convergence.

Table 2. Numerical results for the Hock–Schitkowski problems (continued).

Problem	Iterations	c_*	Problem	Iterations	c_*
56	11	9.4×10^8	86	16	7.6×10^{12}
57	22	18.9	87	13	1.7×10^{12}
59	12	1.7×10^{11}	88	16	2.1×10^9
60	11	0	89	16	1.5×10^{11}
61	20	9.5×10^5	90	20	1.1×10^{11}
62	12	0	91	16	1.0×10^{11}
63	20	2.3×10^6	92	18	7.1×10^{11}
64	23	4.1×10^{11}	93*	15	2.3×10^{12}
65	12	1.2×10^{11}	95	31	1.0×10^{12}
66	20	9.8×10^{12}	96	31	4.3×10^{12}
67	45	3.8×10^{11}	97*	22	1.9×10^8
68	11	2.3×10^7	98*	19	3.1×10^5
69	43	4.2×10^5	99	35	1.2×10^{10}
70	14	3.0×10^{12}	100	14	6.1×10^{12}
71	25	2.8×10^{11}	102	69	7.8×10^{12}
72	18	6.1×10^{10}	103	66	1.3×10^{12}
73	12	4.1×10^{11}	104	59	6.6×10^{12}
74	8*	6.0×10^9	105	12	0
75	9*	104	107	15	7.1×10^{12}
76	12	0	108	21	6.8×10^{12}
77	12	3243	110	11	0
78	50	1.3×10^4	111	13	13957
79	7	10	112	26	0
80	14	2.2×10^6	113	15	4.8×10^{12}
81	56	465	114	48	2.7×10^{12}
83	16	2.8×10^{10}	117	18	1.1×10^{12}
84	35	3.6×10^5	119	19	1157.6

Moreover, the algorithm was tested on many large optimization problem. All of these problems were limited up to 300 variables and constraints, since we used the student version of AMPL. In order to overcome this limitation and be able to test the algorithm on a diverse collection of problems, in some problems, we have reduced the number of variables or constraints to 300, by modifying some parameters in their definition. The numerical results are summarized in Table 3.

At this points, we present an example in order to show the importance of the mechanism that switches between the two merit functions. Consider the following box-constrained optimization problem:

$$\begin{aligned} \min \quad & f(x)=(x_1-1)(x_1-2)(x_1-3)+(x_1-2)(x_1-3)(x_2-1) \\ & \qquad -(x_1-3)(x_2-1)(x_2-2)-(x_2-1)(x_2-2)(x_2-3), && \text{(34a)} \\ \text{s.t.} \quad & -5\le x_i\le 5, \quad i=1,2. && \text{(34b)} \end{aligned}$$

Table 3. Numerical results for the large problems.

Problem	Iterations	c_*	Problem	Iterations	c_*
Antenna	79	1.3×10^{10}	dea2	12	0
Catenary	23	1.5×10^{8}	dea_lp	11	1481
markowita100	31	25.1	dea_lp2	11	1481
1_lp_150	18	9.1×10^{11}	dea_frac_lin	14	0
nnls	32	0	fir_linear	20	5.0×10^{12}
nnls2	17	0	fir_convex	21	2.4×10^{11}
oetl_148	16	0	fir_socp	20	2.3×10^{11}
oet3_148	18	0	fir_exp	20	4.1×10^{10}
minsurf	9	0	svanberg299	26	1.7×10^{10}
obstclal	24	0	vanderm1	74	3.9×10^{11}
dea	25	0	vanderm3	78	1.1×10^{10}

Problem (34) has three local minima,

$$x^1_{\min} = (-5, -0.698),$$

$$x^2_{\min} = (3.395, 5),$$

$$x^3_{\min} = (2.5, 1.5),$$

and two saddle points,

$$x^1_{\text{sad}} = (1.293, 1.293),$$

$$x^2_{\text{sad}} = (2.707, 2.707).$$

We solved (34) with the original Algorithm 2.1 and with a variant that uses only the norm of the KKT conditions as the merit function in all iterations. We observed that, if we use the KKT residual norm as the only merit function, there is no guarantee that Algorithm 2.1 will converge to a local minimum. In fact the algorithm tends to converge to saddle points. On the other hand, the strategy of switching merit functions ensures that the Algorithm 2.1 converges to local minima.

Finally, we mention that performance of the algorithm is very good and comparable with that of other primal-dual interior-point algorithms. For example, our algorithm seems to have superior performance, in terms of the number of iterations, than the interior points algorithms described in Ref. 5 and Ref. 4.

References

1. Rustem, B., *Equality and Inequality Constrained Optimization Algorithms with Convergent Stepsizes*, Journal of Optimization Theory and Applications, Vol. 76, pp. 429–453, 1993.
2. El-Bakry, A. S., Tapia, R. A., Tsuchiya, T., and Zhang, Y., *On the Formulation and Theory of the Newton Interior-Point Method for Nonlinear Programming*, Journal of Optimization Theory and Applications, Vol. 89, pp. 507–541, 1996.
3. Yamashita, H., *Globally Convergent Primal-Dual Interior-Point Method for Constrained Optimization*, Technical Report, Mathematical Systems Institute, Shinjuku, Shinjuku-Ku, Tokyo, Japan, 1995.
4. Gay, D. M., Overton, M. L., and Wright, M. H., *A Primal-Dual Interior Method for Nonconvex Nonlinear Programming*, Technical Report 97-4-08, Bell Laboratories, Murray Hill, New Jersey, 1997.
5. Vanderbei, R. J., and Shanno, D. F., *An Interior-Point Algorithm for Nonconvex Nonlinear Programming*, Technical Report SOR-97-21, Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey, 1997.
6. Akrotirianakis, I., and Rustem, B., *A Primal-Dual Interior-Point Algorithm with an Exact and Differentiable Merit Function for Nonlinear Programming*, Optimization Methods and Software, Vol. 14, pp. 1–36, 2000.
7. Forsgren, A., and Gill, P. E., *Primal-Dual Interior-Point Methods for Nonconvex Nonlinear Programming*, SIAM Journal on Optimization, Vol. 8, pp. 1132–1152, 1998.
8. Bartholomew-Biggs, M. C., *IP from an SQP Point of View*, Optimization Methods and Software, Vol, 16, pp. 69–84, 2001.
9. Chamberlain, R. M., Lemarechal, C., Pedersen, H. C., and Powell, M. J. D., *The Watchdog Technique for Forcing Convergence in Algorithms for Constrained Optimization*, Mathematical Programming Study, Vol. 16, pp. 1–17, 1982.
10. Zakovic, S., Rustem, B., and Pantelides, C., *An Interior-Point Algorithm for Computing Saddle Points of Constrained Continuous Minimax*, Annals of Optimization Research, Vol. 99, pp. 59–78, 2000.
11. Yamashita H., and Yabe, H., *Superlinear and Quadratic Convergence of Some Primal-Dual Interior Point Methods for Constrained Optimization*, Mathematical Programming, Vol. 75, pp. 377–397, 1996.
12. Akrotirianakis, I., and Rustem, B., *A Globally Convergent Interior-Point Algorithm for Nonlinear Programming Problems*, Technical Report 97-14, Department of Computing, Imperial College, London, UK, 1997.
13. Hock, W., and Schittkowski, K., *Test Examples for Nonlinear Programming Codes*, Springer Verlag, New York, NY, 1981.
14. Vanderbei, R. J., *Large-Scale Nonlinear AMPL Models*, Princeton University, Princeton, New Jersey 1997; see http://www.princeton.edu/rvdb/ampl/nlmodels.
15. Powell, M. J. D., *A Fast Algorithm for Nonlinearly Constrained Optimization Calculation*, Numerical Analysis Proceedings, Biennial Conference, Dundee,

Scotland, Edited by G. A. Watson, Lecture Notes in Mathematics, Springer Verlag, Berlin, Germany, Vol. 630, pp. 144–157, 1978.

16. Gill, P. E., Murray, W., and Wright, M. H., *Practical Optimization*, Academic Press, New York, NY, 1981.
17. Rustem, B., *Convergent Stepsizes for Constrained Optimization Algorithms*, Journal of Optimization Theory and Applications, Vol. 49, pp. 135–160, 1986.
18. Rustem, B., *Algorithms for Nonlinear Programming and Multiple-Objective Decisions*, J. Wiley, New York, NY, 1998.