The convergence of a modified barrier method for convex programming

by D. L. Jensen
R. A. Polyak

We show, using elementary considerations, that a modified barrier function method for the solution of convex programming problems converges for any fixed positive setting of the barrier parameter. With mild conditions on the primal and dual feasible regions, we show how to use the modified barrier function method to obtain primal and dual optimal solutions, even in the presence of degeneracy. We illustrate the argument for convergence in the case of linear programming, and then generalize it to the convex programming case.

Introduction

In this paper we discuss the convergence of the modified barrier method for solving linear and convex programming problems. A linear program is an optimization problem that entails finding the best solution satisfying the specified constraints. A linear measure of the quality of the possible solutions, called the objective function, is given, as well as a set of linear equalities and inequalities constraining variables. A convex program is an optimization problem in which both the constraints and the objective function are represented by convex functions. There are many algorithms for solving linear and convex programming problems. Some algorithms relax the constraints by permitting, but penalizing, solutions that come close to violating the constraints. These penalties are applied by means of rapidly growing "barrier" functions. By iteratively decreasing the penalty given to such violations, the algorithms generate a sequence of solutions that converge to a solution of the original problem. These algorithms are referred to as barrier methods. Fiacco and McCormick initiated the study of these methods in [1] and examined logarithmic barrier functions. These barrier functions are also referred to as classical barrier functions.

Barrier methods to solve linear and nonlinear programming problems have been used for many years and have recently experienced a resurgence because of the success of interior-point algorithms for linear and nonlinear programming problems. This interest was rekindled by Karmarkar's result [2] that these methods are theoretically efficient and his claims that they are practical. The methods that are most widely used in practice employ a barrier parameter, which tends to zero during the solution process in order to weaken the effect of the barriers and to produce a sequence of solutions converging to an optimal solution. In this paper, we study a modified barrier function and a related method in which the barrier parameter is initialized to any positive number and remains fixed throughout the procedure. Instead of the barrier parameter being updated, individual weights on the barrier functions are updated, and these weights cluster at optimal dual solutions.

The work in the present paper includes transforming a classical barrier method into a Lagrangian with weights and shifts and is philosophically similar to augmented Lagrangian methods. Powell's 1969 derivation of augmented Lagrangian methods involved the inclusion of weights and shifts for each constraint in the classical quadratic penalty function, in order to avoid the need to drive the penalty parameter to infinity [3]. It is interesting indeed (as Osborne and Jittorntrum pointed out nearly twenty years ago [4, 5]) that a similar idea can be applied to the classical logarithmic barrier function. Numerous authors, including [6, 7], have studied weighted and shifted barrier functions. Besides the derivation and analysis of another weighted and shifted barrier function method in Conn et al. [8], the introduction of that paper provides a clear and concise comparison of several of these methods, including the work of Jittorntrum and Osborne [4, 5]. For further references, the reader can consult a bibliography on interior-point methods compiled by Kranich [9].

The modified barrier functions and corresponding methods were introduced by Polyak in [10] and analyzed under standard second-order optimality conditions. In the case of linear programming problems, the assumption underlying the analysis was that both primal and dual problems are nondegenerate [11]. In other words, the theory of modified barrier functions and the modified barrier function method for constrained optimization problems were developed for problems that have unique primal and dual solutions.

Under the nondegeneracy assumption, there exists a constant $k_0 > 0$, which is a function of the problem data, such that for any fixed barrier parameter $k \geq k_0$, the modified barrier function method converges to the unique primal and dual solutions with a linear rate of convergence. Moreover, for any given constant specifying the rate of reduction in error at each step, one can find a fixed barrier parameter $k \geq k_0$ that guarantees that the modified barrier function method will generate a pair of sequences that converge to the primal and dual solutions linearly at this specified rate. Consequently, by increasing the barrier parameter from step to step, one can obtain a superlinear rate of convergence.

In practice, the performance of the modified barrier function method and its variations does not seem to depend on the nondegeneracy assumption, for either linear or nonlinear programming [12, 13]. It is important to establish a theoretical foundation for the behavior of the method in the presence of degeneracy and in the case in which the barrier parameter remains fixed throughout the procedure. Recently, M. Powell [14] has undertaken a similar analysis for linear programming problems.

In this paper, we consider the behavior of the modified barrier function method for convex programming problems. We do so by replacing the nondegeneracy assumption with mild conditions on the primal and dual feasible regions. We prove convergence for any fixed barrier parameter $k > 0$ and fixed positive constraint shifts. We first demonstrate convergence of the method for linear programming problems, both to illustrate the basic ideas and to set the framework of the more general argument. We show that the sequence of primal (dual) iterates has these properties: The associated primal (dual) objective function converges to the optimal value; the dual iterates have dual optimal solutions as their limit points; and the primal iterates can be averaged to provide asymptotic primal solutions. We then generalize the argument to the case of convex programming problems, taking Wolfe's dual [15] for our dual program.

The motivation for fixing k at an arbitrary value is to begin to explain behavior that we have observed in practice. It is in no way a proposal to fix the parameter at some arbitrary level. Indeed, we consider an example in the next section in which the rate of convergence of the multipliers becomes unacceptably slow even for moderate values of k. It does, however, argue that in an algorithm where k is increased as the algorithm proceeds, k can be fixed after any finite number of increases without jeopardizing convergence. From a practical viewpoint, the algorithm as stated is not implementable, since it requires exact optimization of the modified barrier function, which is not possible on any finite-precision machine.

We conclude the paper by showing the equivalence of the modified barrier function method for linear programming discussed here and a method for linear programming using shifted "entropy functions." There, the linear objective function is shifted but the constraints are not. Entropy functions and their use in optimization have been studied by many authors and have a rich literature; the interested reader can refer to [16, 17].

Convergence for linear programming problems

We consider a self-dual formulation of linear programs. To motivate this formulation, we first consider one of the several equivalent canonical representations of a dual pair of linear programming problems,

$$\begin{array}{llll} & \min\ \mathbf{u}^{0\mathrm{T}}\mathbf{x} & & \max\ \mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{x}^{0} \\ (\mathrm{P}) & \mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}^{0} & (\mathrm{D}) & \mathbf{u} = \mathbf{u}^{0} - \mathbf{A}^{\mathrm{T}}\mathbf{y} \\ & \mathbf{x} \geq 0 & & \mathbf{u} \geq 0 \end{array}$$

where $\mathbf{A} \in \Re^{m \times n}$, and $\mathbf{u}^0$ and $\mathbf{Ax}^0$ determine the primal and dual objective functions, respectively. Assuming that both programs have feasible solutions, we may take $\mathbf{x}^0$ and $\mathbf{u}^0$ to be fixed, nonnegative, and feasible. If we define the vector subspace $V^\perp$ of $\Re^n$ as $V^\perp \equiv \{\mathbf{w} \mid A\mathbf{w} = 0\}$ and its orthogonal complement $V \equiv \{\mathbf{v} \mid \mathbf{v} - \mathbf{A}^\top \mathbf{y} = 0\}$, this dual pair of programs can be rewritten as

$$\min \mathbf{u}^{0\top}\mathbf{x}$$

$$\text{(P)} \qquad \mathbf{x} \in \{\mathbf{x}^0 + \mathbf{w} \mid \mathbf{w} \in V^\perp\} \cap \Re^n_+$$

$$\max \mathbf{x}^{0\top}(\mathbf{u}^0 - \mathbf{u}) = \mathbf{x}^{0\top}\mathbf{u}^0 - \min \mathbf{x}^{0\top}\mathbf{u}$$

$$\text{(D)} \qquad \mathbf{u} \in \{\mathbf{u}^0 + \mathbf{v} \mid \mathbf{v} \in V\} \cap \Re^n_+ .$$

Any dual pair of feasible linear programming problems can be cast in the framework illustrated above, which we now formalize. We let V be a Euclidean vector space in $\Re^n$ and $V^\perp$ be the orthogonal complement of V. The reader will note that any V that is a Euclidean vector space in $\Re^n$ can be represented in terms of a matrix $\mathbf{A}$, as above. We strengthen the assumption that both (P) and (D) are feasible by assuming that $\mathbf{u}^0$ is strictly positive and $\mathbf{x}^0$ is nonnegative. We define the sets X_+ and U_+ as

$$U = \{\mathbf{u} \mid \mathbf{u} = \mathbf{u}^0 + \mathbf{v}, \mathbf{v} \in V\}, \qquad U_+ = U \cap \Re^n_+ ,$$

$$X = \{\mathbf{x} \mid \mathbf{x} = \mathbf{x}^0 + \mathbf{w}, \mathbf{w} \in V^\perp\}, \qquad X_+ = X \cap \Re^n_+ .$$

Thus, a canonical form for a dual pair of programs is

$$\text{(P)} \min_{\mathbf{x} \in X_+} \mathbf{u}^{0\top}\mathbf{x} \quad \text{and} \quad \text{(D)} \min_{\mathbf{u} \in U_+} \mathbf{x}^{0\top}\mathbf{u}. \tag{1}$$

By construction, both linear programs are feasible, since $\mathbf{x}^0 \in X_+$ and $\mathbf{u}^0 \in U_+$. Furthermore, since the objectives of each are nonnegative, they both are bounded below in value by zero. The formulation is nonstandard, in that both programs are minimization problems and the programs do not have the same value. However, the sum of the optimal values of the programs is $\mathbf{x}^{0\top}\mathbf{u}^0$, so we could rewrite the first program as

$$\max_{\mathbf{x} \in X_+} \{\mathbf{u}^{0\top}\mathbf{x}^0 - \mathbf{u}^{0\top}\mathbf{x}\}$$

or the second program as

$$\max_{\mathbf{u} \in U_+} \{\mathbf{x}^{0\top}\mathbf{u}^0 - \mathbf{x}^{0\top}\mathbf{u}\}$$

to obtain a pair of programs with the same value and with opposite directions of optimization.

A consequence of this formulation is that any pair of optimal solutions is characterized by a complementarity condition. Suppose $\mathbf{x}^*$ and $\mathbf{u}^*$ are optimal for (P) and (D), respectively. Since $0 = (\mathbf{x}^0 - \mathbf{x}^*)^\top(\mathbf{u}^0 - \mathbf{u}^*)$ and $\mathbf{u}^{0\top}\mathbf{x}^0 = \mathbf{u}^{0\top}\mathbf{x}^* + \mathbf{x}^{0\top}\mathbf{u}^*$, it follows that $0 = \mathbf{u}^{*\top}\mathbf{x}^*$, which implies that $\mathbf{x}^*$ and $\mathbf{u}^*$ are complementary.

We now discuss the modified barrier function method for solving linear programming problems in this framework, in order to underscore the independence of our reasoning from the representation of the linear programming problem. Since the framework is symmetric, the algorithm presented can equally well be applied to (P) or (D) of (1).

The modified barrier function method is developed in great detail in [10, 11]. The method is motivated by a reformulation of the problem (P) of (1) as

$$\min_{\mathbf{x} \in X} \mathbf{u}^{0\top}\mathbf{x}$$

$$\frac{1}{k} \ln (k\mathbf{x}_i + 1) \geq 0, \qquad i = 1, 2, \cdots, n. \tag{2}$$

In this paper we consider the scalar $k > 0$ to be a fixed parameter of the problem. In order to solve the original problem, the modified barrier function method finds a saddle-point for the classical Lagrangian for problem (2),

$$\max_{\mathbf{u} \in U_+} \min_{\mathbf{x} \in X} F(\mathbf{x}, \mathbf{u}; k)^{a},$$

where

$$F(\mathbf{x}, \mathbf{u}; k) \equiv \begin{cases} \mathbf{u}^{0\top}\mathbf{x} - \dfrac{1}{k}\displaystyle\sum_{i=1}^{n} \mathbf{u}_i \ln (k\mathbf{x}_i + 1) & \\ \qquad \text{for } \left\{\mathbf{x} \mid \mathbf{x} \in X, \mathbf{x}_i \geq -\dfrac{1}{k}, i = 1, 2, \cdots, n\right\}, \\ \infty \qquad \text{otherwise.} \end{cases}$$

The classical Lagrangian for problem (2), defined by $F(\mathbf{x}, \mathbf{u}; k)$, is the modified barrier function.

It is easy to verify that the basic properties of the modified barrier function, enumerated below, are satisfied for every $k > 0$, where $\mathbf{x}^*$ and $\mathbf{u}^*$ are an optimal, hence complementary, pair for Equation (2):

1. $F(\mathbf{x}^*, \mathbf{u}^*; k) = \mathbf{u}^{0\top}\mathbf{x}^*$.

2. $\nabla_{\mathbf{x}} F(\mathbf{x}^*, \mathbf{u}^*; k) = \left(\mathbf{u}^0 - \dfrac{\mathbf{u}^*}{k\mathbf{x}^* + 1}\right) \in V.$[b]

3. $\nabla^2_{\mathbf{xx}} F(\mathbf{x}^*, \mathbf{u}^*; k) = \operatorname{diag}\left[\dfrac{\mathbf{u}^*_i}{(k\mathbf{x}^*_i + 1)^2}\right]_{i=1}^{n} = U^*,$

   where $U^* = \operatorname{diag}[\mathbf{u}^*_i]^n_{i=1}$,[c] is positive semidefinite.

[a] We use a semicolon to offset k in $F(\cdot, \cdot; k)$, to emphasize the fact that $k > 0$ is a fixed parameter.
[b] $\mathbf{u}^*/(kX^* + 1)$ denotes the vector in $\Re^n$ whose ith component is $\mathbf{u}^*_i/(kX^*_i + 1)$.
[c] $\operatorname{diag}[a_i]^n_{i=1}$ denotes the diagonal matrix with ith diagonal entry a_i.

Thus, given $\mathbf{u}^*$, we may optimize $F(\mathbf{x}, \mathbf{u}^*; k)$ to obtain $\mathbf{x}^*$, if the linear program (D) is nondegenerate. [If (D) is degenerate, we are guaranteed only that optimizing F will produce a complementary, but not necessarily feasible, $\mathbf{x}^*$.] Furthermore, if $\tilde{\mathbf{u}} > 0$ and is close to a dual optimal solution, the minimizer of $F(\mathbf{x}, \tilde{\mathbf{u}}; k)$ will be close to a primal optimal solution. We show that by successively fixing $\mathbf{u}$ and minimizing the modified barrier function F with respect to $\mathbf{x}$, one can produce an update to $\mathbf{u}$ as a byproduct of the minimization. The sequence of dual solutions $\mathbf{u}$ produced by this mechanism is strictly increasing in dual objective function value. In this way, the modified barrier function method solves the saddle-point problem by maximizing over all feasible $\mathbf{u}$ the minimum over $\mathbf{x}$ of the classical Lagrangian for the equivalent problem of (2).

These considerations lead to the modified barrier function method, which is described in Algorithm 1.

Algorithm 1 The modified barrier function method for linear programming is as follows:

1. Initialization:

$$s = 0; \qquad 0 < \mathbf{u}^0 \in U_+;$$

2. Primal update:

$$\mathbf{x}^{s+1} = \operatorname*{argmin}_{\mathbf{x}\in X_k} F(\mathbf{x}, \mathbf{u}^s; k)$$

$$= \operatorname*{argmin}_{\mathbf{x}\in X_k} \left[\mathbf{u}^{0^\top}\mathbf{x} - \frac{1}{k}\sum_{i=1}^{n} \mathbf{u}_i^s \ln(k\mathbf{x}_i + 1) \right],$$

$$\text{where } X_k \equiv \left\{ \mathbf{x} \in X \mid \mathbf{x}_i \geq -\frac{1}{k}, \quad i = 1, 2, \cdots, n \right\};$$

3. Dual update:

$$\mathbf{u}_i^{s+1} = \frac{\mathbf{u}_i^s}{k\mathbf{x}_i^{s+1} + 1}, \qquad i = 1, 2, \cdots, n;$$

4. Iteration:

$$s = s + 1;$$

Goto 2.

The algorithm is well defined as long as the minimum in step 2 is attained. A necessary and sufficient condition for the minimum to exist is that the set

$$X^0 = \{\mathbf{x} \in X_+ \mid \mathbf{u}^{0^\top}\mathbf{x} \leq \mathbf{u}^{0^\top}\mathbf{x}^0\}$$

be bounded. This, in turn, is guaranteed if and only if the dual feasible region U_+ has a strictly interior point. We also require that the set of dual feasible solutions of dual objective function value better than our initial solution

$$U^0 = \{\mathbf{u} \in U_+ \mid \mathbf{x}^{0^\top}\mathbf{u} \leq \mathbf{x}^{0^\top}\mathbf{u}^0\}$$

be bounded. This will happen if and only if the primal feasible region X_+ has a strictly interior point. Under these conditions, the algorithm is well defined, since the strict convexity of F implies that the minimum in step 2 is attained at a unique point $\mathbf{x}^{s+1}$.

Before establishing the convergence of this algorithm, we consider its behavior on a simple example. The primal program is given by

$$\min \tfrac{1}{3}\mathbf{x}_1 + 2\mathbf{x}_2 + \tfrac{1}{3}\mathbf{x}_3 + \tfrac{1}{3}\mathbf{x}_4 + \tfrac{1}{3}\mathbf{x}_5,$$

$$\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 - \mathbf{x}_4 = 1,$$

$$\mathbf{x}_1 - \mathbf{x}_2 + \mathbf{x}_3 + \mathbf{x}_5 = 1,$$

$$\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5 \geq 0,$$

and a dual program is given by

$$\min \tfrac{1}{2}\mathbf{u}_1 + \tfrac{1}{2}\mathbf{u}_2 + \tfrac{1}{2}\mathbf{u}_3 + \tfrac{1}{2}\mathbf{u}_4 + \tfrac{1}{2}\mathbf{u}_5,$$

$$\tfrac{1}{3} - \mathbf{y}_1 - \mathbf{y}_2 = \mathbf{u}_1,$$

$$2 - \mathbf{y}_1 + \mathbf{y}_2 = \mathbf{u}_2,$$

$$\tfrac{1}{3} - \mathbf{y}_1 - \mathbf{y}_2 = \mathbf{u}_3,$$

$$\tfrac{1}{3} + \mathbf{y}_1 = \mathbf{u}_4,$$

$$\tfrac{1}{3} - \mathbf{y}_2 = \mathbf{u}_5,$$

$$\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5 \geq 0.$$

Note that these programs are formulated as in the discussion at the beginning of this section. Also note that any point on the line segment joining (1, 0, 0, 0, 0) and (0, 0, 1, 0, 0) is primal optimal, and every point on the segment has primal objective value 1/3. Any point on the line segment joining (0, 7/3, 0, 1/3, 0) and (0, 5/3, 0, 2/3, 1/3) is dual optimal and has dual objective value 4/3. Note that the inner product of the objective functions is 5/3, which is the sum of the optimal values. Also note that the problem is both primal and dual degenerate. That is, both primal and dual optimal faces are overspecified and are line segments. Additionally, each of the feasible regions is unbounded, but their optimal faces are bounded.

Pictorially, the primal feasible region is given in **Figure 1**, and the face of optimal solutions is the bold line segment in the $\mathbf{x}_2 = 0$ plane.

We consider starting the algorithm with a variety of initial settings for $\mathbf{u}$ and k. First, suppose we use the objective function of the primal problem as our initial dual feasible $\mathbf{u}^0$ and take $k = 1$. These results are given in **Table 1**. Algorithm 1 converges slowly to an optimal pair of solutions.

Second, suppose we use the objective function of the primal problem as our initial dual feasible $\mathbf{u}^0$, as above, but take $k = 100$. The results are given in **Table 2**. The convergence is much faster, as indicated by the relative quality of the tabulated solutions in Tables 1 and 2.

Finally, suppose we use a different dual feasible solution as $\mathbf{u}^0$ and take $k = 100$, as indicated in **Table 3**. Tables 2 and 3 indicate that the dual optimal solution obtained is dependent on the initial dual feasible solution, but that the primal optimal solution is not. Powell has shown that, for linear programming, this algorithm converges to a unique primal solution [14]. This example supports that conclusion, and also shows that the dual solution is a function of the initial barrier multipliers.

We now prove the following theorem.

Theorem 1 If X_+ is nonempty, U_+ has an interior point, and both U^0 and X^0 are bounded, Algorithm 1 is well defined, and the sequences $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$ generated by Algorithm 1 satisfy the following properties:

1. The sequence $\{\mathbf{u}^s\}_{s=0}^{\infty}$ is a sequence of points in U_+ whose dual objective values monotonically converge to the optimal value

   $$\mathbf{u}^{*\mathrm{T}}\mathbf{x}^0 \le \cdots \le \mathbf{u}^{s+1^{\mathrm{T}}}\mathbf{x}^0 \le \mathbf{u}^{s^{\mathrm{T}}}\mathbf{x}^0 \le \cdots \mathbf{u}^{0^{\mathrm{T}}}\mathbf{x}^0.$$

   If any of the inequalities hold at equality, an optimal solution to the program, which is a fixed point of the iterates, has been found.
2. The sequence $\{\mathbf{x}^s\}_{s=0}^{\infty}$ is a sequence of points in X_k whose primal objective values converge to the optimal value

   $$\lim_{s\to\infty} \mathbf{u}^{0^{\mathrm{T}}}\mathbf{x}^s = \mathbf{u}^{0^{\mathrm{T}}}\mathbf{x}^*.$$
3. There is a method of averaging the elements of $\{\mathbf{x}^s\}_{s=0}^{\infty}$ to obtain a sequence $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ of primal solutions arbitrarily close to X_+ and converging in primal objective value to the optimum.
4. The value gap[d] $\Sigma_{i=1}^{m} \mathbf{u}_i^s \mathbf{x}_i^s$ between primal and dual iterates converges to zero.

We prove this theorem by establishing four propositions. The first proposition is that the solutions $\mathbf{u}^{s+1}$ generated in step 3 of Algorithm 1 are dual feasible.

Proposition 1 [dual feasibility] The vectors in the sequence $\{\mathbf{u}^s\}_{s=0}^{\infty}$ are contained in U_+.

Proof This is a direct consequence of the optimality conditions for the minimization performed in Step 2 of the algorithm:

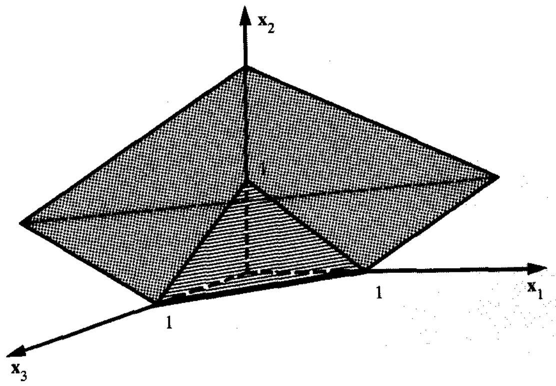

Figure 1

Primal feasible region of the example. The bold line is the set of primal optimal solutions.

$$\nabla_{\mathbf{x}} F(\mathbf{x}, \mathbf{u}; k) = \left(\mathbf{u}^0 - \frac{\mathbf{u}}{k\mathbf{x} + 1}\right) \in V. \qquad \blacksquare$$

The second proposition is that the sequence of dual values $\{\mathbf{x}^{0^{\mathrm{T}}}\mathbf{u}^s\}_{s=1}^{\infty}$ is monotonically decreasing.

Proposition 2 [dual-value monotonicity]

$$\mathbf{u}^{s+1^{\mathrm{T}}}\mathbf{x}^{s+1} = \mathbf{u}^{s^{\mathrm{T}}}\mathbf{x}^{s+1} - \sum_{i=1}^{n} \frac{k\mathbf{u}_i^s(\mathbf{x}_i^{s+1})^2}{k\mathbf{x}_i^{s+1} + 1}.$$

Proof This follows from the definition of $\mathbf{u}^{s+1}$. Since

$$\mathbf{u}_i^{s+1} = \frac{\mathbf{u}_i^s}{k\mathbf{x}_i^{s+1} + 1}, \qquad i = 1, 2, \cdots, n,$$

we have

$$\mathbf{u}^{s+1^{\mathrm{T}}}\mathbf{x}^{s+1} = \sum_{i=1}^{n} \frac{\mathbf{u}_i^s\mathbf{x}_i^{s+1}}{k\mathbf{x}_i^{s+1} + 1}$$

$$= \mathbf{u}_i^{s^{\mathrm{T}}}\mathbf{x}_i^{s+1} - \sum_{i=1}^{n} \frac{k\mathbf{u}_i^s(\mathbf{x}_i^{s+1})^2}{k\mathbf{x}_i^{s+1} + 1}. \qquad \blacksquare$$

Remark 1 Note that the difference $\mathbf{u}^{s+1^{\mathrm{T}}}\mathbf{x}^{s+1} - \mathbf{u}^{s^{\mathrm{T}}}\mathbf{x}^{s+1}$ is the difference between the objective function values at $\mathbf{u}^{s+1}$ and $\mathbf{u}^s$. This follows from the orthogonality of V and $V^{\perp}$; from the fact that $\mathbf{u}^s \in U$, $\mathbf{u}^{s+1} \in U$, $\mathbf{x}^{s+1} \in X$, and $\mathbf{x}^0 \in X$; and from consideration of the relation

$$(\mathbf{u}^{s+1} - \mathbf{u}^s)^{\mathrm{T}}(\mathbf{x}^{s+1} - \mathbf{x}^0) = 0.$$

[d] We use the term "value" for a vector $\mathbf{x}$ or $\mathbf{u}$ to mean the objective function value in the associated primal or dual program, respectively.

Table 1 Algorithm 1 applied with initial dual feasible $\mathbf{u}^0 = (1/3, 2, 1/3, 1/3, 1/3)$ and modified barrier parameter $k = 1$.

$\mathbf{u}^0$	$\mathbf{u}^1$	$\mathbf{u}^2$	$\mathbf{u}^3$	$\mathbf{u}^4$	$\mathbf{u}^5$	$\mathbf{u}^6$
3.3333D − 01	2.3374D − 01	1.6154D − 01	1.1044D − 01	7.4916D − 02	5.0539D − 02	3.3964D − 02
2.0000D + 00	1.9890D + 00	1.9755D + 00	1.9633D + 00	1.9535D + 00	1.9462D + 00	1.9410D + 00
3.3333D − 01	2.3374D − 01	1.6154D − 01	1.1044D − 01	7.4916D − 02	5.0539D − 02	3.3964D − 02
3.3333D − 01	3.8862D − 01	4.3147D − 01	4.6315D − 01	4.8580D − 01	5.0163D − 01	5.1252D − 01
3.3333D − 01	2.8902D − 01	2.5968D − 01	2.4026D − 01	2.2738D − 01	2.1883D − 01	2.1315D − 01
$\mathbf{x}^1$	$\mathbf{x}^2$	$\mathbf{x}^3$	$\mathbf{x}^4$	$\mathbf{x}^5$	$\mathbf{x}^6$	
4.2611D − 01	4.4692D − 01	4.6269D − 01	4.7419D − 01	4.8235D − 01	4.8802D − 01	
5.5196D − 03	6.8369D − 03	6.2363D − 03	5.0041D − 03	3.7433D − 03	2.6910D − 03	
4.2611D − 01	4.4602D − 01	4.6269D − 01	4.7419D − 01	4.8235D − 01	4.8802D − 01	
−1.4227D − 01	−9.9319D − 02	−6.8387D − 02	−4.6620D − 02	−3.1562D − 02	−2.1260D − 02	
1.5331D − 01	1.1299D − 01	8.0860D − 02	5.6629D − 02	3.9048D − 02	2.6642D − 02	

Table 2 Algorithm 1 applied with initial dual feasible $\mathbf{u}^0 = (1/3, 2, 1/3, 1/3, 1/3)$ and modified barrier parameter $k = 100$.

$\mathbf{u}^0$	$\mathbf{u}^1$	$\mathbf{u}^2$	$\mathbf{u}^3$	$\mathbf{u}^4$
3.3333D − 01	6.6627D − 04	1.3064D − 05	2.5616D − 07	5.0228D − 09
2.0000D + 00	1.9295D + 00	1.9278D + 00	1.9292D + 00	1.9281D + 00
3.3333D − 01	6.6627D − 04	1.3064D − 05	2.5616D − 07	5.0228D − 09
3.3333D − 01	5.3493D − 01	5.3518D − 01	5.3538D − 01	5.3529D − 01
3.3333D − 01	2.0227D − 01	2.0182D − 01	2.0205D − 01	2.0185D − 01
$\mathbf{x}^1$	$\mathbf{x}^2$	$\mathbf{x}^3$	$\mathbf{x}^4$	
4.9976D − 01	4.9999D − 01	5.0000D − 01	5.0000D − 01	
5.9709D − 05	8.8031D − 06	−7.6048D − 06	5.7177D − 06	
4.9976D − 01	4.9999D − 01	5.0000D − 01	5.0000D − 01	
−4.1887D − 04	−4.6940D − 06	−3.7242D − 06	1.6350D − 06	
5.3829D − 04	2.2300D − 05	−1.1485D − 05	9.8004D − 06	

Also, note that the only circumstance in which the sequence in dual values is not strictly monotone is if $\mathbf{x}^s = 0 \in X$ for some s, as can be seen from Proposition 2. In that case, all points in U_+ are dual optimal.

Remark 2 Since the dual objective function value is monotone decreasing, it follows that $\mathbf{u}^{t^{\mathrm{T}}}\mathbf{x}^0 - \mathbf{u}^{r^{\mathrm{T}}}\mathbf{x}^0 \le 0$ whenever $t \ge r$. The orthogonality of V and $V^{\perp}$ implies that $\mathbf{u}^{t^{\mathrm{T}}}\mathbf{x} - \mathbf{u}^{r^{\mathrm{T}}}\mathbf{x} \le 0$ whenever $t \ge r$ and $\mathbf{x} \in X$.

We now have that the sequence $\{\mathbf{u}^s\}_{s=0}^{\infty}$ monotonically converges in value. To demonstrate that the limiting value is the optimal value, it is sufficient to produce a complementary primal feasible point.

Although we show below that the sequence $\{\mathbf{x}^s\}_{s=0}^{\infty}$ converges in value to the optimum, we have been unable to show that the sequence $\{\mathbf{x}^s\}_{s=0}^{\infty}$ converges to an optimal solution. We can show, however, that if we average the elements of the sequence over appropriately chosen subsequences of consecutive elements, the average solutions converge in value to the optimal value and the average solutions come arbitrarily close to the nonnegative orthant. To define this sequence of average solutions, we focus on the dual solutions. We first choose a subsequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ of $\{\mathbf{u}^s\}_{s=0}^{\infty}$ satisfying the following properties.

Property 1 The subsequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ converges to a point $\bar{\mathbf{u}}$.

Such a subsequence can be chosen since, by assumption, U^0 is compact.

Property 2 Let $I^0 = \{i \mid \bar{\mathbf{u}}_i = 0\}$ and let $I^+ = \{i \mid \bar{\mathbf{u}}_i > 0\}$. We require that the subsequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ satisfy

$$\min\{\mathbf{u}_i^{s_l} \mid i \in I^0\} \ge 2 \times \max\{\mathbf{u}_i^{s_{l+1}} \mid i \in I^0\}, \qquad \text{for } l = 0, 1, \cdots, \infty.$$

We can select a subsequence satisfying Property 2, since $\{\mathbf{u}^s\}_{s=0}^{\infty}$ (and consequently $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$) is a sequence of strictly positive vectors. Given any subsequence satisfying Property 1, we can choose a subsequence of it satisfying Property 2.

We select this subsequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ of $\{\mathbf{u}^s\}_{s=0}^{\infty}$ because of the fundamental observation that a component i of $\mathbf{u}$ decreases as a result of a dual update only if $\mathbf{x}_i^s > 0$. This observation generalizes nicely. If a component i of $\mathbf{u}$

decreases over a consecutive sequence of dual updates, the average of the $\mathbf{x}_i^s$ over this sequence must be positive.

Our argument in proving Theorem 1 proceeds as follows. First we show that the sequence of average solutions

$$\bar{\mathbf{x}}^{l+1} \equiv \frac{1}{s_{l+1} - s_l} \sum_{s=s_l+1}^{s_{l+1}} \mathbf{x}^s$$

is arbitrarily close, componentwise, to the nonnegative orthant. Since each element of $\{\mathbf{x}^s\}_{s=0}^{\infty}$ is in X_k, so is each element of the sequence of averages $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$. Thus, although the sequence $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ may not converge, the distance between its elements and X_+ converges to zero.

Next we use the fact that $\mathbf{x}^s$ and $\mathbf{u}^s$ are orthogonal in the limit to show that the sequences $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ and $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ satisfy complementarity in the limit. Our assumption that X^0 and U^0 are compact allows us to conclude that the sequences $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ and $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ have cluster points that are primal and dual feasible, respectively, and that are complementary.

Proposition 3 [asymptotic primal feasibility] For every real number $\epsilon_0 > 0$, there exists an l_0 such that

$$\bar{\mathbf{x}}_i^l \geq -\frac{\epsilon_0}{k} \qquad \text{for } i = 1, 2, \cdots, n$$

for every $l \geq l_0$.

Proof To establish the fact that the components of $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ are nonnegative in the limit, we consider the index sets L^0 and I^+ separately. First, consider an index $i \in I^0$. Using the definition of $\{\mathbf{u}^s\}_{s=0}^{\infty}$ and Property 2 of the sequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$, we have

$$\mathbf{u}_i^{s_{l+1}} \prod_{s=s_l+1}^{s_{l+1}} (k\mathbf{x}_i^s + 1) = \mathbf{u}_i^{s_l} \geq 2 \times \mathbf{u}_i^{s_{l+1}}, \qquad \text{for } i = 1, 2, \cdots, n.$$

Using the arithmetic–geometric means inequality, we have

$$\frac{1}{s_{l+1} - s_l} \sum_{s=s_l+1}^{s_{l+1}} (k\mathbf{x}_i^s + 1) \geq \left(\prod_{s=s_l+1}^{s_{l+1}} (k\mathbf{x}_i^s + 1) \right)^{1/(s_{l+1}-s_l)}$$

$$\geq 2^{1/(s_{l+1}-s_l)} > 1,$$

which implies

$$\bar{\mathbf{x}}_i^{l+1} > 0.$$

Now, consider an index $i \in I^+$. Since $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ converges to $\bar{\mathbf{u}}$, for any $\epsilon_0 > 0$, there exists an $l_0 > 0$ such that

$$\mathbf{u}_i^{s_{l+1}} \prod_{s=s_l+1}^{s_{l+1}} (k\mathbf{x}_i^s + 1) = \mathbf{u}_i^{s_l} \geq (1 - \epsilon_0) \times \mathbf{u}_i^{s_{l+1}} \qquad (3)$$

for every $l \geq l_0$. [Indeed, if we choose l_0 so that $|\mathbf{u}_i^{s_l} - \bar{\mathbf{u}}_i| < (\epsilon_0/2)\bar{\mathbf{u}}_i$ for all $l \geq l_0$, then l_0 satisfies the condition in (3).] Again invoking the arithmetic–geometric means inequality, we can conclude that

$$\frac{1}{s_{l+1} - s_l} \sum_{s=s_l+1}^{s_{l+1}} (k\mathbf{x}_i^s + 1) \geq \left[\prod_{s=s_l+1}^{s_{l+1}} (k\mathbf{x}_i^s + 1) \right]^{1/(s_{l+1}-s_l)}$$

$$\geq (1 - \epsilon_0)^{1/(s_{l+1}-s_l)}.$$

Rearranging terms, we have

$$\bar{\mathbf{x}}_i^{l+1} \geq \frac{1}{k}[(1 - \epsilon_0)^{1/(s_{l+1}-s_l)} - 1] \geq -\frac{\epsilon_0}{k}. \qquad \blacksquare$$

To draw the conclusion that the points $\bar{\mathbf{x}}^l$ come arbitrarily close in value to the points $\mathbf{u}^s$, we prove the following consequence of Proposition 2.

Proposition 4 [complementarity] The sequences $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$ satisfy

$$\lim_{s\to\infty} \mathbf{u}^{s^{\mathrm{T}}} \mathbf{x}^s = 0.$$

Furthermore, with the subsequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ of $\{\mathbf{u}^s\}_{s=0}^{\infty}$ satisfying properties 1 and 2, and the average sequence $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$, defined as above, derived from $\{\mathbf{x}^s\}_{s=0}^{\infty}$, we have

$$\lim_{s\to\infty} \mathbf{u}^{s_l^{\mathrm{T}}} \bar{\mathbf{x}}^l = 0.$$

Proof Since the value of the dual solutions is bounded below by zero, we know that

$$\sum_{s=1}^{\infty} \sum_{i=1}^{n} \frac{k\mathbf{u}_i^s(\mathbf{x}_i^{s+1})^2}{k\mathbf{x}_i^{s+1} + 1} \leq \mathbf{u}^{0^{\mathrm{T}}} \mathbf{x}^0.$$

Thus,

Table 3 Algorithm 1 applied with initial dual feasible $\mathbf{u}^0 = (1/3, 2, 198/300, 1/3, 1/300, 1/300)$ and modified barrier parameter $k = 100$.

$\mathbf{u}^0$	$\mathbf{u}^1$	$\mathbf{u}^2$
3.3333D − 01	1.9160D − 11	3.7568D − 13
2.6600D + 00	2.3293D + 00	2.3304D + 00
3.3333D − 01	1.9160D − 11	3.7568D − 13
3.3333D − 03	3.3474D − 01	3.3477D − 01
3.3333D − 03	1.4322D − 03	1.4334D − 03
$\mathbf{x}^1$	$\mathbf{x}^2$	
5.0000D − 01	5.0000D − 01	
4.8350D − 06	−4.5746D − 06	
5.0000D − 01	5.0000D − 01	
7.7049D − 07	−7.6775D − 07	
8.8996D − 06	−8.3814D − 06	

$$\sum_{i=1}^{n} k\mathbf{u}_i^{s+1}(\mathbf{x}_i^{s+1})^2 = \sum_{i=1}^{n} \frac{k\mathbf{u}_i^{s}(\mathbf{x}_i^{s+1})^2}{k\mathbf{x}_i^{s+1}+1} < \alpha_s$$

for some sequence $\{\alpha^s\}_{s=1}^{\infty}$ with $\lim_{s\to\infty}\alpha_s = 0$. But this implies that

$$\alpha_s > \sum_{i=1}^{n} k\mathbf{u}_i^{s+1}(\mathbf{x}_i^{s+1})^2$$

$$= k\sum_{i=1}^{n} \frac{1}{\mathbf{u}_i^{s+1}}(\mathbf{u}_i^{s+1})^2(\mathbf{x}_i^{s+1})^2$$

$$\geq k \min_{1\leq i\leq n}\left\{\frac{1}{\mathbf{u}_i^{s+1}}\right\}\sum_{i=1}^{n}(\mathbf{u}_i^{s+1})^2(\mathbf{x}_i^{s+1})^2$$

$$\geq \frac{k}{n}\min_{1\leq i\leq n}\left\{\frac{1}{\mathbf{u}_i^{s+1}}\right\}\left(\sum_{i=1}^{n}\mathbf{u}_i^{s+1}|\mathbf{x}_i^{s+1}|\right)^2$$

$$\geq \frac{k}{n}\min_{1\leq i\leq n}\left\{\frac{1}{\mathbf{u}_i^{s+1}}\right\}\left(\sum_{i=1}^{n}\mathbf{u}_i^{s+1}\mathbf{x}_i^{s+1}\right)^2.$$

Rearranging terms, we have

$$\sqrt{\frac{n}{k}\alpha_s \max_{1\leq i\leq n}\{\mathbf{u}_i^{s+1}\}} > \sum_{i=1}^{n}\mathbf{u}_i^{s+1}\mathbf{x}_i^{s+1}. \tag{4}$$

Thus, $\mathbf{u}^{s^\top}\mathbf{x}^s$ converges to zero. Our assumption that U^0 is bounded allows us to conclude that

$$\lim_{s\to\infty}\sum_{i=1}^{n}\mathbf{u}_i^{s+1}\mathbf{x}_i^{s+1} = 0. \tag{5}$$

Now suppose we choose S large enough to guarantee that

$$\mathbf{u}^{s^\top}\mathbf{x}^s < \epsilon_0 \tag{6}$$

for all $s \geq S$. Consider $\mathbf{u}^{s_{l+1}\top}\bar{\mathbf{x}}^{l+1}$ when $s_l > S$. Then

$$\sum_{i=1}^{n}\mathbf{u}_i^{s_{l+1}}\bar{\mathbf{x}}_i^{l+1}$$

$$= \frac{1}{s_{l+1}-s_l}\sum_{i=1}^{n}\mathbf{u}_i^{s_{l+1}}\sum_{s=s_l+1}^{s_{l+1}}\mathbf{x}_i^s$$

$$= \frac{1}{s_{l+1}-s_l}\sum_{s=s_l+1}^{s_{l+1}}\sum_{i=1}^{n}\mathbf{u}_i^{s_{l+1}}\mathbf{x}_i^s.$$

We now invoke Remark 2 and inequality (6) to conclude that

$$\sum_{i=1}^{n}\mathbf{u}_i^{s_{l+1}}\bar{\mathbf{x}}_i^{l+1}$$

$$= \frac{1}{s_{l+1}-s_l}\sum_{s=s_l+1}^{s_{l+1}}\sum_{i=1}^{n}\mathbf{u}_i^{s_{l+1}}\mathbf{x}_i^s$$

$$\leq \frac{1}{s_{l+1}-s_l}\sum_{s=s_l+1}^{s_{l+1}}\sum_{i=1}^{n}\mathbf{u}_i^{s}\mathbf{x}_i^s$$

$$< \epsilon_0.$$ ■

Proof of Theorem 1 The inequalities of the first property of the theorem follow from Propositions 1 and 2. The convergence to the optimal value follows from the complementarity of $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ and $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$, which is a conclusion of Proposition 4. The second property of the theorem follows from the orthogonality of corresponding elements of the sequences $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$ in the limit. That is, using the orthogonality of the sequences $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$ in the limit and the fact that $\{\mathbf{u}^s\}_{s=0}^{\infty}$ is a sequence of dual feasible solutions converging to the optimal value, we have

$$\mathbf{u}^{0^\top}\mathbf{x}^* = \mathbf{u}^{0^\top}\mathbf{x}^* - \lim_{s\to\infty}\mathbf{u}^{s^\top}\mathbf{x}^* - \lim_{s\to\infty}\mathbf{u}^{0^\top}\mathbf{x}^s$$

$$+ \lim_{s\to\infty}\mathbf{u}^{s^\top}\mathbf{x}^s + \lim_{s\to\infty}\mathbf{u}^{0^\top}\mathbf{x}^s$$

$$= \lim_{s\to\infty}(\mathbf{u}^0-\mathbf{u}^s)^\top(\mathbf{x}^*-\mathbf{x}^s) + \lim_{s\to\infty}\mathbf{u}^{0^\top}\mathbf{x}^s$$

$$= \lim_{s\to\infty}\mathbf{u}^{0^\top}\mathbf{x}^s.$$

The third property is the content of Proposition 3. The fourth property is proved in Equation (4) of Proposition 4. ■

Convex programming problems

In the case of convex programming, we need to use a framework that is not self-dual. We define the primal convex programming problem to be

$$\min f_0(\mathbf{x})$$

$$f_i(\mathbf{x}) \geq 0, \qquad i = 1, 2, \cdots, m, \tag{7}$$

where $f_i : \Re^n \to \Re$ for $i = 0, 1, \cdots, m$, each of the functions f_i is differentiable, f_0 is convex, and each of the functions f_i for $i = 1, 2, \cdots, m$ is concave. The program that will serve as the dual to (7) is known as Wolfe's dual [15]:

$$\max\ L(\mathbf{x}, \mathbf{u}) \equiv f_0(\mathbf{x}) - \sum_{i=1}^{m} \mathbf{u}_i f_i(\mathbf{x})$$

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{u}) = \nabla f_0(\mathbf{x}) - \sum_{i=1}^{m} \mathbf{u}_i \nabla f_i(\mathbf{x}) = 0, \tag{8}$$

where $\mathbf{x}$ is unrestricted and $0 \le \mathbf{u} \in \Re^m$. It is well known that Wolfe's dual provides lower bounds for a convex programming problem. Furthermore, in the context of convex programming, a necessary and sufficient condition for this dual program to have an optimal value equal to the optimal value of the primal program is that the convex programming program be stable [18]. Our requirements that the functions $f_i(\mathbf{x})$ be real-valued and that the primal problem have an optimum are sufficient to guarantee that problem (7) is stable. Note that if $\mathbf{x}$ and $\mathbf{u}$ are feasible for (8), $\mathbf{x}$ minimizes the Lagrangian $L(\mathbf{x}, \mathbf{u})$ for fixed $\mathbf{u}$. In other words, program (8) is equivalent to

$$\max_{\mathbf{u} \ge 0} \min_{\mathbf{x} \in \Re^n} \left[f_0(\mathbf{x}) - \sum_{i=1}^{m} \mathbf{u}_i f_i(\mathbf{x}) \right].$$

Since problem (7) is a convex programming problem, given a fixed $\mathbf{u}$ the value of the dual program $L(\mathbf{x}, \mathbf{u})$ is constant for all $\mathbf{x}$ satisfying $\nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{u}) = 0$.

Given any fixed $k > 0$, we consider a primal feasible region of

$$X_+ = \{\mathbf{x} \in \Re^n \mid f_i(\mathbf{x}) \ge 0, i = 1, 2, \cdots, m\},$$

an extended primal feasible region over which the modified barrier function method is defined,

$$X_k = \left\{ \mathbf{x} \in \Re^n \mid f_i(\mathbf{x}) \ge -\frac{1}{k}, i = 1, 2, \cdots, m \right\},$$

a dual feasible region of

$$U_+ = \{\mathbf{u} \in \Re^m_+ \mid \nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{u}) = \nabla f_0(\mathbf{x}) - \sum_{i=1}^{m} \mathbf{u}_i \nabla f_i(\mathbf{x}) = 0 \quad \text{for some } \mathbf{x} \in X_+\},$$

and an extended dual feasible region

$$U_k = \{\mathbf{u} \in \Re^m_+ \mid \nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{u}) = \nabla f_0(\mathbf{x}) - \sum_{i=1}^{m} \mathbf{u}_i \nabla f_i(\mathbf{x}) = 0 \quad \text{for some } \mathbf{x} \in X_k\}.$$

We assume that there exists $\mathbf{x}^0$ in X_+ and that U_+ has a strictly interior point $\mathbf{u}^0 > 0$.

The modified barrier function for (7), the problem under consideration, is

$$F(\mathbf{x}, \mathbf{u}; k) = \begin{cases} f_0(\mathbf{x}) - \dfrac{1}{k} \displaystyle\sum_{i=1}^{m} \mathbf{u}_i \ln[k f_i(\mathbf{x}) + 1] & \mathbf{x} \in \operatorname{int} X_k, \\ \infty & \mathbf{x} \notin \operatorname{int} X_k. \end{cases}$$

We apply Algorithm 2, which follows, for convex programming problems.

Algorithm 2 The modified barrier function method for convex nonlinear programming is as follows:

1. Initialization:

   $s = 0; \quad \mathbf{u}^0 \in U_k;$

2. Primal update:

   $$\mathbf{x}^{s+1} \in \operatorname*{argmin}_{\mathbf{x} \in \Re^n} F(\mathbf{x}, \mathbf{u}^s; k) = \operatorname*{argmin}_{\mathbf{x} \in \Re^n} \left\{ f_0(\mathbf{x}) - \frac{1}{k} \sum_{i=1}^{m} \mathbf{u}_i^s \ln[k f_i(\mathbf{x}) + 1] \right\};$$

3. Dual update:

   $$\mathbf{u}_i^{s+1} = \frac{\mathbf{u}_i^s}{k f_i(\mathbf{x}^{s+1}) + 1}, \qquad i = 1, 2, \cdots, m;$$

4. Iteration:

   $s = s + 1;$

   Goto 2.

This algorithm is also well defined as long as the minimum in step 2 is attained. We assume that the image under the constraints of all points in X with lower objective values than $\mathbf{x}^0$,

$$V^0 \equiv \{(v_0, v_1, v_2, \cdots, v_m) \mid \mathbf{x} \in \Re^n, v_0 = f_0(\mathbf{x}) \le f_0(\mathbf{x}^0), v_i = f_i(\mathbf{x}) \ge 0, i = 1, 2, \cdots, m\},$$

is a compact subset of $\Re^m$. When the objective is convex and the constraints are concave, the minimum in step 2 exists if and only if V^0 is a compact subset of $\Re^m$. This, in turn, is guaranteed if and only if the dual feasible region U_+ has a strictly interior point. Our proof of convergence of the method also requires a compactness assumption on the dual feasible region U_k. We call $\psi(\mathbf{u})$ the value of the dual solution $\mathbf{u} \in U_k$ and define ψ by

$$\psi(\mathbf{u}) = \min_{\mathbf{x} \in \Re^n} \left[f_0(\mathbf{x}) - \sum_{i=1}^{m} \mathbf{u}_i f_i(\mathbf{x}) \right],$$

and require that

$$U^0 = \{\mathbf{u} \in U_k \mid \psi(\mathbf{u}) \ge \psi(\mathbf{u}^0)\},$$

the set of dual solutions with better value than $\mathbf{u}^0$, is compact. As in the case of linear programming, U^0 is compact if and only if V^0 contains a point that is positive on coordinates $1, 2, \cdots, m$.

In this section, we prove the convex programming analogue of Theorem 1. Given a fixed $\mathbf{u}^s$ and corresponding $\mathbf{x}^s$ determined in step 2 of Algorithm 2, we call $L(\mathbf{x}^s, \mathbf{u}^s) = \psi(\mathbf{u}^s)$ the value of $\mathbf{u}^s$. With this definition of the value of a dual solution and the conditions above on the functions $f_i(\mathbf{x})$, we can prove the following result.

Theorem 2 If X_+ is nonempty, U_+ has an interior point, and both U^0 and V^0 are bounded, then Algorithm 2 is well defined, and the sequences $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$ generated by Algorithm 2 satisfy the following properties:

1. $\{\mathbf{u}^s\}_{s=0}^{\infty}$ is a sequence of dual feasible solutions whose values monotonically converge to the optimal value

$$f_0(\mathbf{x}^*) = L(\mathbf{x}^*, \mathbf{u}^*) \geq \cdots \geq L(\mathbf{x}^{s+1}, \mathbf{u}^{s+1}) \geq L(\mathbf{x}^s, \mathbf{u}^s) \geq \cdots \geq L(\mathbf{x}^0, \mathbf{u}^0).$$

   If any of the inequalities hold at equality, an optimal solution to the program and a fixed point of the iterates has been found.
2. The elements of the sequence $\{\mathbf{x}^s\}_{s=0}^{\infty}$ are in X_k, and their values converge to the optimal value

$$\lim_{s\to\infty} f_0(\mathbf{x}^s) = f_0(\mathbf{x}^*).$$

3. There is a method of averaging the elements of $\{\mathbf{x}^s\}_{s=0}^{\infty}$ to obtain a sequence of primal feasible solutions that converge to the optimal value.
4. The value gap $\Sigma_{i=1}^m \mathbf{u}_i^s f_i(\mathbf{x}^s)$ between primal and dual iterates converges to zero.

We must show that each of the four propositions of the previous section remains valid in this broader context.

Proposition 5 [dual feasibility] For the sequence of vectors $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$, we have that $\mathbf{x}^s \in X_k$, $\mathbf{u}^s \in U_+$, and

$$\nabla_{\mathbf{x}} F(\mathbf{x}^{s+1}, \mathbf{u}^s; k) = \nabla_{\mathbf{x}} L(\mathbf{x}^{s+1}, \mathbf{u}^{s+1}) = \nabla f_0(\mathbf{x}^{s+1}) - \sum_{i=1}^{m} \mathbf{u}_i^{s+1} \nabla f_i(\mathbf{x}^{s+1}) = 0.$$

Proof The optimality conditions for the unconstrained optimization problem of step 2 of Algorithm 2 require that the gradient of the objective function vanish:

$$\nabla_{\mathbf{x}} F(\mathbf{x}^{s+1}, \mathbf{u}^s; k) = \nabla f_0(\mathbf{x}^{s+1}) - \sum_{i=1}^{m} \frac{\mathbf{u}_i^s}{[k f_i(\mathbf{x}^{s+1}) + 1]} \nabla f_i(\mathbf{x}^{s+1}) = 0.$$

Applying the dual update formula of step 3, we have that

$$\nabla_{\mathbf{x}} F(\mathbf{x}^{s+1}, \mathbf{u}^s; k) = \nabla_{\mathbf{x}} L(\mathbf{x}^{s+1}, \mathbf{u}^{s+1}) = 0,$$

and $\mathbf{u}^{s+1}$ is feasible for program (8). ■

The second proposition is that the value of the dual solutions $\{\mathbf{u}^s\}_{s=0}^{\infty}$ is monotonically increasing.

Proposition 6 [dual-value monotonicity]

$$L(\mathbf{x}^{s+1}, \mathbf{u}^{s+1}) \geq L(\mathbf{x}^s, \mathbf{u}^s) + \sum_{i=1}^{m} \frac{k \mathbf{u}_i^s [f_i(\mathbf{x}^{s+1})]^2}{k f_i(\mathbf{x}^{s+1}) + 1}.$$

Proof Since $\mathbf{x}^s$ minimizes the Lagrangian when $\mathbf{u}$ is set equal to $\mathbf{u}^s$,

$$L(\mathbf{x}^s, \mathbf{u}^s) = f_0(\mathbf{x}^s) - \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^s) \leq f_0(\mathbf{x}^{s+1}) - \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^{s+1}).$$

Applying the definition of the dual update, we have

$$f_0(\mathbf{x}^{s+1}) - \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^{s+1}) = f_0(\mathbf{x}^{s+1}) - \sum_{i=1}^{m} \mathbf{u}_i^{s+1} f_i(\mathbf{x}^{s+1}) + \sum_{i=1}^{m} [\mathbf{u}_i^{s+1} f_i(\mathbf{x}^{s+1}) - \mathbf{u}_i^s f_i(\mathbf{x}^{s+1})] = L(\mathbf{x}^{s+1}, \mathbf{u}^{s+1}) - \sum_{i=1}^{m} \frac{k \mathbf{u}_i^s [f_i(\mathbf{x}^{s+1})]^2}{k f_i(\mathbf{x}^{s+1}) + 1}.$$

Remark 3 We can actually assume that the strict inequality

$$L(\mathbf{x}^{s+1}, \mathbf{u}^{s+1}) > L(\mathbf{x}^s, \mathbf{u}^s)$$

holds. Otherwise, $f_i(\mathbf{x}^s) = 0$ for $i = 1, 2, \cdots, m$, and $\mathbf{x}^s$ and $\mathbf{u}^s$ are an optimal pair for the convex programming problem, since they are feasible for both programs (7) and (8) and satisfy complementarity.

As in the proof for the linear programming case, our assumption that U^0 is compact allows us to choose a sequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ satisfying properties 1 and 2. We again consider the associated sequence of averages $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$.

Proposition 7 [asymptotic primal feasibility] For every number $\epsilon_0 > 0$, there exists an l_0 such that

$$f_i(\bar{\mathbf{x}}^l) \geq -\frac{\epsilon_0}{k}, \qquad i = 1, 2, \cdots, m$$

for every $l \geq l_0$.

Proof The proof of Proposition 3 suffices to show that

$$\frac{1}{s_{l+1} - s_l} \sum_{s=s_l+1}^{s_{l+1}} f_i(\mathbf{x}^s) \geq -\frac{\epsilon_0}{k}$$

for every l exceeding a fixed l_0. Since each of the functions f_i is concave,

$$f_i(\bar{\mathbf{x}}^l) = f_i\left(\frac{1}{s_{l+1} - s_l} \sum_{s=s_l+1}^{s_{l+1}} \mathbf{x}^s\right) \geq \frac{1}{s_{l+1} - s_l} \sum_{s=s_l+1}^{s_{l+1}} f_i(\mathbf{x}^s),$$

and the result follows. ■

We strengthen Proposition 4 of the previous section, establishing complementarity, with the aid of the following lemma.

Lemma 1 Let $t > r$ and $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$ be defined as in Algorithm 2. Then

$$\sum_{i=1}^{m} \mathbf{u}_i^t f_i\left(\frac{1}{t - r} \sum_{s=r+1}^{t} \mathbf{x}^s\right) \leq \frac{1}{t - r} \sum_{i=1}^{m} \sum_{s=r+1}^{t} \mathbf{u}_i^s f_i(\mathbf{x}^s).$$

Proof First, we invoke the property that, by virtue of being feasible for dual program (8), $\mathbf{x}_t$ minimizes $L(\mathbf{x}, \mathbf{u}_t)$, to conclude that

$$f_0\left(\frac{1}{t - r} \sum_{s=r+1}^{t} \mathbf{x}^s\right) - \sum_{i=1}^{m} \mathbf{u}_i^t f_i\left(\frac{1}{t - r} \sum_{s=r+1}^{t} \mathbf{x}^s\right)$$

$$\geq f_0(\mathbf{x}^t) - \sum_{i=1}^{m} \mathbf{u}_i^t f_i(\mathbf{x}^t). \tag{9}$$

An application of Proposition 6 yields

$$f_0(\mathbf{x}^t) - \sum_{i=1}^{m} \mathbf{u}_i^t f_i(\mathbf{x}^t) \geq f_0(\mathbf{x}^s) - \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^s) \tag{10}$$

for each s satisfying $r < s < t$. Now, averaging (10) over s, applying (9), and rearranging terms, we obtain

$$\sum_{i=1}^{m} \mathbf{u}_i^t f_i\left(\frac{1}{t - r} \sum_{s=r+1}^{t} \mathbf{x}^s\right) \leq \frac{1}{t - r} \sum_{i=1}^{m} \sum_{s=r+1}^{t} \mathbf{u}_i^s f_i(\mathbf{x}^s)$$

$$+ f_0\left(\frac{1}{t - r} \sum_{s=r+1}^{t} \mathbf{x}^s\right)$$

$$- \frac{1}{t - r} \sum_{s=r+1}^{t} f_0(\mathbf{x}^s). \qquad ■$$

Since $f_0(\mathbf{x})$ is convex, we obtain

$$\sum_{i=1}^{m} \mathbf{u}_i^t f_i\left(\frac{1}{t - r} \sum_{s=r+1}^{t} \mathbf{x}^s\right) \leq \frac{1}{t - r} \sum_{i=1}^{m} \sum_{s=r+1}^{t} \mathbf{u}_i^s f_i(\mathbf{x}^s).$$

We can now establish the convex analogue of Proposition 4.

Proposition 8 [complementarity] The sequences $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$ satisfy

$$\lim_{s\to\infty} \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^s) = 0.$$

Furthermore, with the subsequence $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$ of $\{\mathbf{u}^s\}_{s=0}^{\infty}$ and the average sequence $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ derived from $\{\mathbf{x}^s\}_{s=0}^{\infty}$ defined as above,

$$\lim_{l\to\infty} \sum_{i=1}^{m} \mathbf{u}_i^{s_{l+1}} f_i(\bar{\mathbf{x}}^{l+1}) = 0.$$

Proof We can use an argument identical to the one provided in the proof of Proposition 4 for the linear programming case to conclude that

$$\lim_{s\to\infty} \sum_{i=1}^{m} \mathbf{u}_i^{s+1} f_i(\mathbf{x}_i^{s+1}) = 0, \tag{11}$$

and, moreover, that

$$\sqrt{\frac{n}{k}}\, \alpha_s \max_{1\leq i\leq n} \{\mathbf{u}_i^{s+1}\} > \sum_{i=1}^{m} \mathbf{u}_i^{s+1} f_i(\mathbf{x}^{s+1}),$$

where $\lim_{s\to\infty} \alpha_s = 0$. Now suppose we choose S large enough to guarantee that

$$\sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^s) < \epsilon_0 \tag{12}$$

for all $s \geq S$. Then we apply Lemma 1 with $r = s_l$ and $t = s_{l+1}$ to (12) to conclude that

$$\sum_{i=1}^{m} \mathbf{u}_i^{s_{l+1}} f_i\left(\frac{1}{s_{l+1}-s_l} \sum_{s=s_l+1}^{s_{l+1}} \mathbf{x}^s\right) = \sum_{i=1}^{m} \mathbf{u}_i^{s_{l+1}} f_i(\bar{\mathbf{x}}^{l+1}) \le \epsilon_0 .$$

Proof of Theorem 2 The inequalities of the first property of the theorem follow from Propositions 5 and 6. The convergence to the optimal value follows from the complementarity of $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ and $\{\mathbf{u}^{s_l}\}_{l=0}^{\infty}$, which is the conclusion of Proposition 8. The second property of the theorem follows from the orthogonality in the limit of corresponding elements of the sequences $\{f_0(\mathbf{x}^s)\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$, which is established in (11). In particular, since $\mathbf{x}^s$ and $\mathbf{u}^s$ constitute a feasible solution to (8) and the values of the dual solutions are monotone increasing, we have that

$$f_0(\mathbf{x}^*) = f_0(\mathbf{x}^*) - \sum_{i=1}^{m} \mathbf{u}_i^* f_i(\mathbf{x}^*) \le f_0(\mathbf{x}^s) - \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^s).$$

Because the values are monotone increasing, we have

$$\begin{aligned} f_0(\mathbf{x}^*) &= f_0(\mathbf{x}^*) - \sum_{i=1}^{m} \mathbf{u}_i^* f_i(\mathbf{x}^*) \\ &= \lim_{s\to\infty} [f_0(\mathbf{x}^s) - \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^s)] \\ &= \lim_{s\to\infty} f_0(\mathbf{x}^s) - \lim_{s\to\infty} \sum_{i=1}^{m} \mathbf{u}_i^s f_i(\mathbf{x}^s) \\ &= \lim_{s\to\infty} f_0(\mathbf{x}^s). \end{aligned}$$

The third property is the content of Proposition 7. The fourth property of the theorem follows from the strict monotonicity of the value of the sequence of dual solutions established in Proposition 6 and Remark 3. ∎

Concluding remarks

Since the logarithmic modified barrier function method exhibits different behavior from the classical barrier function method, it is instructive to study the relationship between the two methods. For a discussion of classical barrier functions, the reader should see Fiacco and McCormick [1].

It is well known that problem (1) is equivalent to the unconstrained nonsmooth problem

$$\mathbf{x}^* \in \operatorname{argmin}\{\varphi(\mathbf{x}) \mid \mathbf{x} \in \Re^n\},$$

where $\varphi(\mathbf{x}) \equiv \max_{0\le i\le m} \varphi_i(\mathbf{x})$, $\varphi_0(\mathbf{x}) \equiv f_0(\mathbf{x}) - f_0(\mathbf{x}^*)$, and $\varphi_i(\mathbf{x}) \equiv -f_i(\mathbf{x})$ for $i = 1, 2, \cdots, m$. To simplify our consideration at this point, we assume that $f_0(\mathbf{x}^*) = 0$, and therefore

$$\varphi(\mathbf{x}) = \max\{f_0(\mathbf{x}), -f_i(\mathbf{x}), i = 1, \cdots, m\}$$

and

$$\varphi(\mathbf{x}^*) = \min_{\mathbf{x}\in\Re^n} \varphi(\mathbf{x}) = 0.$$

A basic idea of barrier methods is to replace the nonsmooth function $\varphi(\mathbf{x})$ with a sequence of smooth functions whose minimizers converge to $\mathbf{x}^*$.

One way to view this optimization is like the problem of finding a good smooth approximation for the nonsmooth function $\varphi(\mathbf{x})$. It is this viewpoint that motivated Polyak's original work on modified barrier functions. Under suitable nondegeneracy assumptions, the modified barrier function provides an exact smooth approximation to $\varphi(\mathbf{x})$ for any fixed $k > 0$ when $\mathbf{u}$ is fixed at $\mathbf{u}^*$. That is, the optimal solutions and optimal values of the modified barrier function and $\varphi(\mathbf{x})$ agree:

$$\min_{\mathbf{x}\in\Re^n} F(\mathbf{x}, \mathbf{u}^*; k) = f_0(\mathbf{x}^*) = \min_{\mathbf{x}\in\Re^n} \varphi(\mathbf{x}) = \varphi(\mathbf{x}^*) = 0.$$

Moreover, for any fixed $k > 0$ and for any sequence $\{\mathbf{u}^s\}_{s=0}^{\infty}$ converging to $\mathbf{u}^*$, there exists a sequence $\{\mathbf{x}^s\}_{s=0}^{\infty}$ converging to $\mathbf{x}^*$ with

$$\mathbf{x}^s = \operatorname*{argmin}_{\mathbf{x}\in\Re^n} F(\mathbf{x}, \mathbf{u}^s; k).$$

Irrespective of nondegeneracy assumptions,

$$\lim_{s\to\infty} [F(\mathbf{x}^s, \mathbf{u}^s; k) - \varphi(\mathbf{x}^s)] = 0.$$

On the other hand, the classical barrier function

$$F(\mathbf{x}; k) = f_0(\mathbf{x}) - \frac{1}{k} \sum_{i=1}^{n} \ln[f_i(\mathbf{x})]$$

is also a smooth approximation to $\varphi(\mathbf{x})$, but in this case

$$\lim_{s\to\infty} [F(\mathbf{x}^s; k) - \varphi(\mathbf{x}^s)] = \infty$$

for any fixed $k > 0$ and any sequence $\{\mathbf{x}^s\}_{s=0}^{\infty}$ converging to $\mathbf{x}^*$. So the difference between $\varphi(\mathbf{x})$ and its smooth approximation $F(\mathbf{x}; k)$ based on the classical barrier function diverges as $\mathbf{x}$ approaches $\mathbf{x}^*$ if the barrier parameter $k > 0$ is fixed. In contrast, the difference between $\varphi(\mathbf{x})$ and the smooth approximation $F(\mathbf{x}, \mathbf{u}; k)$ based on the modified barrier function converges to zero as $\mathbf{x}^s$ approaches $\mathbf{x}^*$ and $\mathbf{u}^s$ approaches $\mathbf{u}^*$ for any fixed value of the barrier parameter $k > 0$.

Although the classical barrier function method does not have the same convergence properties as the modified barrier function method, a variation on the classical barrier function method does, as we indicate in the following.

Suppose we consider the modified barrier function for the standard form of a linear program, as presented above in the section on convergence for linear programming problems. The program

$$(\mathrm{P})\qquad \begin{aligned} &\min\ \mathbf{u}^{0^\top}\mathbf{x}\\ &\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}^0\\ &\mathbf{x} \geq 0 \end{aligned}$$

has an equivalent barrier formulation

$$(\mathrm{P})\qquad \begin{aligned} &\min\ \mathbf{u}^{0^\top}\mathbf{x}\\ &\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}^0\\ &\frac{1}{k}\ln(k\mathbf{x}_i + 1) \geq 0, \qquad i = 1, 2, \cdots, n. \end{aligned}$$

Solving the problem (P) is equivalent to finding a saddle-point of the classical Lagrangian $L(\mathbf{x}, \mathbf{y}, \mathbf{u})$:

$$\begin{aligned} &\min_{\mathbf{x}\in\Re^n}\max_{\mathbf{y}\in\Re^m}\max_{\mathbf{u}\in\Re^n_+} L(\mathbf{x}, \mathbf{y}, \mathbf{u})\\ &\quad = \min_{\mathbf{x}\in\Re^n}\max_{\mathbf{y}\in\Re^m}\max_{\mathbf{u}\in\Re^n_+}\left[\mathbf{u}^{0^\top}\mathbf{x} - \mathbf{y}^\top(\mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}^0) - \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i\ln(k\mathbf{x}_i + 1)\right]\\ &\quad = \max_{\mathbf{u}\in\Re^n_+}\max_{\mathbf{y}\in\Re^m}\left\{\mathbf{y}^\top\mathbf{A}\mathbf{x}^0 + \min_{\mathbf{x}\in\Re^n}\left[(\mathbf{u}^0 - \mathbf{A}^\top\mathbf{y})^\top\mathbf{x} - \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i\ln(k\mathbf{x}_i + 1)\right]\right\}. \end{aligned}$$

The minimum is attained at $(\mathbf{u}^0 - \mathbf{A}^\top\mathbf{y})_i = \mathbf{u}_i/(k\mathbf{x}_i + 1)$. Solving for $\mathbf{x}$, we have

$$\mathbf{x}_i = \frac{\mathbf{u}_i}{k(\mathbf{u}^0 - \mathbf{A}^\top\mathbf{y})_i} - \frac{1}{k},$$

so the original problem has been transformed to

$$\max_{\mathbf{u}\in\Re^n_+}\max_{\mathbf{y}\in\Re^m}\left[\left(\mathbf{x}^0 + \frac{\mathbf{e}}{k}\right)^\top\mathbf{A}^\top\mathbf{y} + \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i\ln(\mathbf{u}^0 - \mathbf{A}^\top\mathbf{y})_i + \frac{1}{k}\mathbf{e}^\top(\mathbf{u} - \mathbf{u}^0) - \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i\ln\mathbf{u}_i\right],$$

where $\mathbf{e} \in \Re^n$ is a vector of ones. The last two terms are constant for any fixed $\mathbf{u}$. Using the symmetric form representation of the problem used in the section on convergence for linear programming problems, we have the following equivalent formulation:

$$\mathbf{x}^{0^\top}\mathbf{u}^0 + \max_{\mathbf{u}\in\Re^n_+}\left\{\frac{1}{k}\mathbf{e}^\top\mathbf{u} - \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i\ln\mathbf{u}_i - \min_{\mathbf{w}\in U_+}\left[\left(\mathbf{x}^0 + \frac{\mathbf{e}}{k}\right)^\top\mathbf{w} - \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i\ln\mathbf{w}_i\right]\right\}.$$

One can prove that optimizing this maximization problem in $\mathbf{u}$ and $\mathbf{w}$ can be accomplished by finding multipliers $\mathbf{u}$ so that the optimal solution to the minimization in $\mathbf{w}$ is also $\mathbf{u}$. Because of the equivalence examined above, the arguments presented in the section on linear programming problems suffice to prove this. Thus, the problem of optimizing the modified barrier function for a fixed $\mathbf{u}$ is equivalent to optimizing a weighted logarithmic barrier function together with an "entropy function" term

$$\left(\frac{1}{k}\mathbf{e}^\top\mathbf{u} - \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i\ln\mathbf{u}_i\right)$$

in $\mathbf{u}$. These comments motivate the following modified barrier function method.

Algorithm 3 The modified barrier function method for linear programming is as follows:

1. Initialization:

   $s = 0;\ 0 < \mathbf{u}^0 \in U_+;$

2. Dual update:

   $$\mathbf{w}^{s+1} = \operatorname*{argmin}_{\mathbf{w}\in U_+} F(\mathbf{w}, \mathbf{u}^s; k) \equiv \operatorname*{argmin}_{\mathbf{w}\in U_+}\left[\left(\mathbf{x}^0 + \frac{\mathbf{e}}{k}\right)^\top\mathbf{w} - \frac{1}{k}\sum_{i=1}^{n}\mathbf{u}_i^s\ln\mathbf{w}_i\right]$$

3. Dual update:

   $\mathbf{u}^{s+1} = \mathbf{w}^{s+1};$

4. Iteration:

   $s = s + 1;$

   Goto 2.

Here we have represented the linear programming problem as a problem with a logarithmic barrier function and a shift in the linear objective function. It is important to note that the sequence $\{\mathbf{u}^s\}_{s=0}^{\infty}$ generated by this algorithm is identical to that generated by Algorithm 1.

The sequence $\{\mathbf{x}^s\}_{s=0}^{\infty}$ is also generated implicitly. The optimality conditions for step (2) require that

$$\mathbf{x}^0 + \frac{\mathbf{e}}{k} - \frac{\mathbf{u}^s}{k\mathbf{w}^{s+1}} = \mathbf{z}^{s+1} \in V^{\perp}.$$

Rearranging terms, we have

$$\mathbf{x}^{s+1} = \mathbf{x}^0 - \mathbf{z}^{s+1} = \frac{\mathbf{u}^s}{k\mathbf{w}^{s+1}} - \frac{\mathbf{e}}{k},$$

or, equivalently,

$$\mathbf{w}^{s+1} = \frac{\mathbf{u}^s}{k\mathbf{x}^{s+1} + 1}.$$

This, together with the fact that, given a fixed $\mathbf{u}^s$, there is a unique pair $\mathbf{x}^{s+1} \in X_k$ and $\mathbf{u}^{s+1} \in U_+$ satisfying these relationships, implies that Algorithms 1 and 3 generate the same sequences $\{\mathbf{x}^s\}_{s=0}^{\infty}$ and $\{\mathbf{u}^s\}_{s=0}^{\infty}$. The convergence results of the section on linear programming problems carry over to this problem formulation, since it is in fact an equivalent problem. This emphasizes the importance of the weights on the barriers in the modified barrier function approach to linear programming. The weights, and the appropriate shift to either the barriers or the objective function, allows us to obtain a convergent algorithm while keeping the barrier parameter $k > 0$ fixed.

The work presented here raises several questions about the modified barrier function method. The use of the arithmetic–geometric-mean inequality to prove that the sequence $\{\bar{\mathbf{x}}^l\}_{l=1}^{\infty}$ is nonnegative is hedging against oscillation in the components of the primal solution. This raises the question of whether such an oscillation in the components of the primal solution can occur. This can happen only if the corresponding duals are converging to zero at different rates. The algorithm as stated is not practical, in the sense that it requires the exact optimum to the modified barrier function. We must find a suitable way to relax this requirement. The algorithm behaves in practice far better than this analysis predicts. We view the results of this paper as a very preliminary validation of the convergence of the modified barrier function method in the presence of degeneracy and with the barrier parameter fixed at any positive level.

Acknowledgment
We are grateful to Professor M. Powell, whose investigations and comments stimulated this work.

References
1. A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968.
2. N. Karmarkar, "A New Polynomial-Time Algorithm for Linear Programming," *Combinatorica* **4,** 373–395 (1984).
3. M. J. D. Powell, "A Method for Nonlinear Constraints in Minimization Problems," *Optimization*, Academic Press, Inc., New York, 1969.
4. K. Jittorntrum and M. R. Osborne, "A Modified Barrier Function Method with Improved Rate of Convergence for Degenerate Problems," *J. Austr. Math. Soc. Ser. B* **21,** 305–329 (1980).
5. K. Jittorntrum and M. R. Osborne, "Trajectory Analysis and Extrapolation in Barrier Function Methods," *J. Austr. Math. Soc. Ser. B* **20,** 39–58 (1978).
6. R. M. Freund, "Theoretical Efficiency of a Shifted-Barrier-Function Algorithm for Linear Programming," *Linear Algebra & Appl.* **152,** 19–41 (1991).
7. P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "Shifted Barrier Methods for Linear Programming," *Technical Report SOL88-9*, Department of Operations Research, Stanford University, Stanford, CA, 1988.
8. A. R. Conn, N. I. M. Gould, and Ph. L. Toint, "A Globally Convergent Lagrangian Barrier Algorithm for Optimization with General Inequality Constraints and Simple Bounds," *Research Report RC-18049*, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 1992.
9. E. Kranich, "Interior Point Methods for Mathematical Programming: A Bibliography," *Discussion Paper 171*, Institute of Economy and Operations Research, FernUniversität, Hagen, Germany, 1991.
10. R. Polyak, "Modified Barrier Functions (Theory and Methods)," *Math. Programming* **54,** 177–222 (1992).
11. R. Polyak, "Modified Barrier Functions in Linear Programming (Submitted to Computational Optimization and Applications)," *Technical Report 17790*, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 1992.
12. A. Ben-Tal, M. Tsibulevskii, and I. Yusefovich, "Modified Barrier Methods for Constrained and Minimax Problems," Technical report, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion (Israel Institute of Technology), Haifa, 1992.
13. D. L. Jensen, R. Polyak, and R. Schneur, "Numerical Experience with Modified Barrier Functions for Linear Programming," *Technical Report 18415*, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 1992.
14. M. Powell, "Some Convergence Properties of a Shifted Log Barrier Method for Linear Programming," *Technical Report NA7*, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England, 1992.
15. P. Wolfe, "A Duality Theorem for Non-Linear Programming," *Quart. Appl. Math.* **19,** 239–244 (1961).
16. J. R. Eriksson, "Algorithms for Entropy and Mathematical Programming," *Linköping Studies in Science and Technology Dissertations*, Department of Mathematics, Linköping University, S-58183 Linköping, Sweden 63, 1981.
17. C. G. Han, P. M. Pardalos, and Y. Ye, "On Potential Reduction Algorithms for Some Entropy Optimization Problems," *Technical Report CS-91-02*, Department of Computer Science, The Pennsylvania State University, University Park, PA, 1991.
18. M. Avriel, *Nonlinear Programming: Analysis and Methods*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1976.

Received January 11, 1993; accepted for publication March 16, 1994

David L. Jensen *IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 (DJENSEN at WATSON, djensen@watson.ibm.com).* Dr. Jensen received a B.A. in mathematics from Colgate University in 1977. In 1985 he received a doctorate in operations research and industrial engineering from Cornell University. He taught in the Applied Mathematics Department at Stony Brook University as an Assistant Professor from 1983 until he joined the Mathematical Sciences Department of the IBM Thomas J. Watson Research Center in 1987. Dr. Jensen manages the Mathematical Optimization Department in the Optimization Center. He is responsible for the quadratic programming component of the Optimization Subroutine Library (OSL) and has contributed to the network component of OSL and to the Graphical User Interface for OSL. More recently, he has worked on mixed-integer quadratic programming, interior-point methods for nonlinear programming, and client/server support for mathematical programming. His research interests include combinatorial optimization, quadratic programming, network optimization, and interior-point algorithms for nonlinear programming.

Roman A. Polyak *Department of Operations Research, George Mason University, Fairfax, Virginia 22030 (rpolyak@gmuvax.gmu.edu).* Dr. Polyak received the Ph.D. degree in mathematics in 1966 from Moscow Central Institute of Mathematics and Economics at the USSR Academy of Sciences. He was a professor in mathematics at the Continuing Education Graduate School for Engineers and Researchers in Kiev. After emigration from the former Soviet Union in 1988, he was a visiting scientist at the Mathematical Sciences Department of the IBM Thomas J. Watson Research Center until the end of 1992. Since January 1993, he has been a Research Professor at the Department of Operations Research at George Mason University. During his years at the Thomas J. Watson Research Center, Dr. Polyak developed the theory of modified barrier and distance functions and associated methods for linear and nonlinear programming. He was also a principal investigator of the joint IBM and NASA study "Optimization in Structural Design." Dr. Polyak is an author and coauthor of six monographs and chapters of books; he has published more than fifty papers in refereed professional journals. His areas of expertise are linear and nonlinear programming, game theory, variational inequalities, and mathematical economics.