Mathematical Programming 59 (1993) 71-85
North-Holland

Restricted simplicial decomposition for convex constrained problems

Jose A. Ventura

Department of Industrial and Management Systems Engineering, The Pennsylvania State University, University Park, PA, USA

Donald W. Hearn

Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA

Received 22 June 1987
Revised 21 June 1991

The strategy of Restricted Simplicial Decomposition is extended to convex programs with convex constraints. The resulting algorithm can also be viewed as an extension of the (scaled) Topkis-Veinott method of feasible directions in which the master problem involves optimization over a simplex rather than the usual line search. Global convergence of the method is proven and conditions are given under which the master problem will be solved a finite number of times. Computational testing with dense quadratic problems confirms that the method dramatically improves the Topkis-Veinott algorithm and that it is competitive with the generalized reduced gradient method.

1. Introduction

In a recent paper (Hearn, Lawphongpanich and Ventura, 1987) the strategy of Restricted Simplicial Decomposition (RSD) has been developed for nonlinear programs with linear constraints. This technique alternately solves a linear programming (LP) subproblem and a nonlinear master problem which has simple constraints, i.e., those which define a simplex. This method offers modularity in the solving of a nonlinear program: the choice of algorithms for both the master and subproblem can be made on the basis of problem structure or any other appropriate criteria. For example, in the cited reference, a combination of the projected Newton method (Bertsekas, 1982) and the primal simplex algorithm is shown to be effective on test problems from the Colville (1968) study as well as on large-scale nonlinear network flow problems.

The objective of this paper is to extend the concept of RSD to the class of nonlinear programs with (quasi-)convex constraints. The convexity of the feasible region is necessary to guarantee that the successive simplices generated by the algorithm are feasible. We denote the algorithm by RSDCC. Just as RSD can be viewed as an extension of the Frank-Wolfe feasible direction method in which the

Correspondence to: Prof. D.W. Hearn, Department of Industrial and Systems Engineering, University of Florida, 303 Weil Hall, Gainesville, FL 32611-2083, USA.
This research was supported in part by NSF Grants ECS-8516365 and DDM-8814075.

usual line search is replaced by optimization over a simplex, RSDCC may be viewed as a generalization of the (scaled) Topkis–Veinott (1967) method. It also alternates between an LP direction-finding subproblem and nonlinear optimization on a simplex. As will be shown, the proof of global convergence can be established with arguments similar to those of Topkis and Veinott. Similar to the result in Hearn, Lawphongpanich and Ventura (1985), we also give conditions under which RSDCC converges after a *finite* number of major cycles. When these conditions are met, RSDCC inherits the *local* convergence rate of the algorithm chosen to solve the master problem. Thus, if the projected Newton algorithm is used, the local convergence will be superlinear. While this finiteness result applies in theory only if the binding constraints are linear, it suggests that the overall number of iterations may be reduced even when nonlinear constraints are binding. Computational results with quadratically constrained quadratic programs confirm this expectation.

2. Problem formulation and the RSDCC algorithm

Consider the following problem

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_j(x) \leq 0, \quad j \in N_q, \\ & Ax \leq b, \end{aligned} \tag{2.1}$$

where $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, A is an $m \times n$ real matrix, and N_q is the set of integers from 1 to q.

In addition to the notation above, we will employ $\nabla f(x)$ and $\nabla g_j(x)$ to represent gradients of the functions $f(x)$ and $g_j(x)$, $\bar{1}$ will be a column vector of ones, $\bar{0}$ a column vector of zeros, and the inner product of two vectors will be denoted by concatenation. Define the following sets:

$$S' = \{x\colon g_j(x) \leq 0, j \in N_q\},$$

$$S'' = \{x\colon Ax \leq b\},$$

and consider the following assumptions:

Assumption 2.1. $f(x)$ is a continuously differentiable pseudoconvex function.

Assumption 2.2. $g_j(x)$ is a continuously differentiable convex function for $j \in N_q$.

Assumption 2.3. The feasible region of problem (2.1), denoted as S ($S = S' \cap S''$), is bounded.

Assumption 2.4. For each $x \in S$, the gradients in the set $\{\nabla g_j(x)\colon g_j(x) = 0, j \in N_q\}$ are linearly independent.

Note that the last assumption rules out the case of nonlinear equality constraints, since the possibility of replacing $g_j(x) = 0$ by $g_j(x) \leq 0$ and $-g_j(x) \leq 0$ is not allowed.

First, we summarize the standard Topkis and Veinott algorithm.

Topkis–Veinott Algorithm.

Step 0. Let x^0 be a feasible point and set $k=0$.

Step 1. (subproblem). If $\|\nabla f(x^k)\|=0$, x^k is a solution; terminate. Otherwise, let (z^k, d^k) be an optimal solution to the following problem:

$$\begin{aligned} \min \quad & z \\ \text{s.t.} \quad & \nabla f(x^k)d - z \leqslant 0, \\ & \nabla g_j(x^k)d - z \leqslant -g_j(x^k) \quad \text{for } j \in N_q, \\ & Ad - z\bar{1} \leqslant b - Ax^k, \\ & -\bar{1} \leqslant d \leqslant \bar{1}. \end{aligned}$$

If $z^k \geqslant 0$, x^k is a solution; terminate. Otherwise, go to Step 2.

Step 2 (master problem). Let

$$\alpha^k \in \arg\min\{f(x^k + \alpha d^k)\colon 0 \leqslant \alpha \leqslant \alpha_{\max}\},$$

where

$$\alpha_{\max} = \sup\{\alpha\colon g_j(x^k + \alpha d^k) \leqslant 0, j \in N_q, A(x^k + \alpha d^k) \leqslant b\}.$$

Set $x^{k+1} = x^k + \alpha^k d^k$, increase k by 1 and go to Step 1.

The RSDCC algorithm modifies the subproblem in Step 1 by scaling the constraints induced by the objective function and the set of nonlinear constraints of the original problem. For each of these functions, the scaling factor is the norm of the gradient at the current iterate and it multiplies the term z. This is intended to balance the effect of the objective function and the set of near-binding constraints in the generation of the descent direction. In addition, the box constraints are defined by a parameter, $\mu > 0$, that bounds the size of d. The box constraints have a center, denoted as $\bar{x}$, that changes only when the progress of the algorithm has been satisfactory. Under conditions to be given in Section 4, $\bar{x}$ eventually becomes fixed and the number of major cycles is finite. In RSDCC the line search of the master problem above is replaced by optimization over a simplex. This enhances the performance of the algorithm by impeding "zig-zagging" when the number of cycles is infinite.

A restriction parameter, $r \geqslant 1$ and integer, is chosen by the user to control the size of the master problem. When $r=1$, the master problem reduces to optimization on a simplex of dimension 1, i.e., a line search (Hearn, Lawphongpanich and Ventura, 1987).

RSDCC Algorithm.

Step 0. Let x^0 be a feasible point, $r \geqslant 1$ and integer, and $\mu > 0$.

Set $\bar{x}=x^0$, $1^0=-\mu\bar{1}$, $u^0=\mu\bar{1}$, $W^0=W_x^0=\{x^0\}$, $W_s^0=\emptyset$, and $k=0$.

Step 1 (subproblem). If $\|\nabla f(x^k)\|=0$, x^k is a solution; terminate. Otherwise, let (z^k, d^k) be an optimal solution to the following problem:

$$\begin{aligned} \min\quad & z \\ \text{s.t.}\quad & \nabla f(x^k)d-\|\nabla f(x^k)\|z\leqslant 0, \\ & \nabla g_j(x^k)d-\|\nabla g_j(x^k)\|z\leqslant -g_j(x^k) \quad \text{for } j\in N_1, \\ & Ad\leqslant b-Ax^k, \\ & 1^k\leqslant d\leqslant u^k. \end{aligned}$$

If $z^k\geqslant 0$, x^k is a solution; terminate. Otherwise, let

$$\bar{\alpha}=\min\{1, \sup\{\alpha: g_j(x^k+\alpha d^k)\leqslant 0, j\in N_q, A(x^k+\alpha d^k)\leqslant b\}\}$$

and let $y^k=x^k+\bar{\alpha}d^k$.

(i) If $|W_s^k|<r$, set $W_s^{k+1}=W_s^k\cup\{y^k\}$ and $W_x^{k+1}=W_x^k$.

(ii) If $|W_s^k|=r$, let y^k replace the element of W_s^k that has minimal weight in the expression of x^k as a convex combination of the elements of W^k to obtain W_s^{k+1}. Let $W_x^{k+1}=\{x^k\}$.

Set $W^{k+1}=W_s^{k+1}\cup W_x^{k+1}$ and go Step 2.

Step 2 (master problem). Let $x^{k+1}\in\arg\min\{f(x): x\in \mathrm{H}(W^{k+1})\}$, where $\mathrm{H}(W^{k+1})$ is the convex hull of W^{k+1}.

Purge W_s^{k+1} and W_x^{k+1} of all elements with zero weight in the expression of x^{k+1} as a convex combination of the elements of W^{k+1} and go to Step 3.

Step 3. Define $\bar{x}$, 1^{k+1} and u^{k+1} as follows:

(i) If $\|\bar{x}-x^{k+1}\|_\infty\leqslant\frac{1}{2}\mu$, set $1^{k+1}=-\mu\bar{1}+(\bar{x}-x^{k+1})$ and $u^{k+1}=\mu\bar{1}+(\bar{x}-x^{k+1})$.

(ii) If $\|\bar{x}-x^{k+1}\|_\infty>\frac{1}{2}\mu$, set $\bar{x}=x^{k+1}$, $1^{k+1}=-\mu\bar{1}$ and $u^{k+1}=\mu\bar{1}$.

Increase k by 1 and go to Step 1.

3. Global Convergence

The proof of global convergence of the RSDCC algorithm follows from similar arguments in Topkis and Veinott (1967) and Zangwill (1969).

Theorem 3.1 (Motzkin's theorem). *Let B be an $m\times n$ matrix and C a $p\times n$ matrix with B being nonvacuous. Then exactly one of the following two systems has a solution*:

System 1: $Bx<0$ *and* $Cx\leqslant 0$ *for some* $x\in\mathbb{R}^n$.

System 2: $B^{\mathrm{T}}v+C^{\mathrm{T}}w=0$, $v\geqslant 0$, $w\geqslant 0$, for some nonzero $v\in\mathbb{R}^m$ and some $w\in\mathbb{R}^p$.

Proof. See, e.g., Mangasarian (1969). □

Lemma 3.1. *Let x^k be a feasible solution to problem* (2.1) *and define*

$$I_1^k = \{j \in N_q : g_j(x^k) = 0\} \quad \textit{and} \quad I_2^k = \{j \in N_m : a^j x^k = b_j\},$$

where a^j is the jth row of A. If (z^k, d^k) is an optimal solution to the subproblem, then

(i) *$z^k \geq 0$ if and only if x^k is an optimal solution to problem* (2.1).

(ii) *$z^k < 0$ if and only if $f(x^{k+1}) < f(x^k)$.*

Proof. It can be easily verified that $z^k > 0$ is never a solution to the subproblem since $(z^k, d^k) = (0, \bar{0})$ is always a feasible point and has a lower objective function value. Now, since $1^k < \bar{0}$ and $\mu^k > \bar{0}$, $z^k = 0$ holds if and only if there is no solution to the system: $\nabla f(x^k) d^k < 0$, $\nabla g_j(x^k) d^k < 0$ for all $j \in I_1^k$, and $a^j d^k \leq 0$ for all $j \in I_2^k$. By Theorem 3.1, this system has no solution if and only if x^k is a Fritz-John point. Since $f(x)$ is pseudoconvex, x^k is a global optimal solution.

If $z^k < 0$, then the solution of the subproblem is such that $\nabla g_j(x^k) d^k < 0$ for all $j \in I_1$ and $a^j d^k \leq 0$ for all $j \in I_2$. This and the fact that $g_j(x^k) < 0$ for all $j \notin I_1^k$ and $a^j x^k < b_j$ for all $j \notin I_2^k$ imply that $x^k + \alpha d^k$ is feasible for all $\alpha \in (0, \bar{\alpha}]$. Thus, d^k is a feasible direction. Also, $\nabla f(x^k) d^k < 0$; hence, d^k is an improving direction. Furthermore, $x^k + \alpha d^k \in \mathrm{H}(W^{k+1}) \subseteq S$, for all $\alpha \in (0, \bar{\alpha}]$. Thus, since x^{k+1} solves the master problem, we conclude that $f(x^{k+1}) < f(x^k)$. Reversing this argument completes the proof. □

Lemma 3.2. *The sequence $\{(x^k, d^k)\}$ generated by RSDCC cannot admit an infinite subsequence K with the following properties:*

(i) *$x^k \to x^\infty$ for $k \in K$.*

(ii) *$d^k \to d^\infty$ for $k \in K$.*

(iii) *$x^k + \alpha d^k \in S$ for all $\alpha \in (0, \delta]$, for each $k \in K$ and for some $\delta > 0$.*

(iv) *$\nabla f(x^\infty) d^\infty < 0$.*

Proof (by contradiction). Assume that there exists a such subsequence K. By condition (iv), there exists a $\tau > 0$ such that $\nabla f(x^\infty) d^\infty = -\tau$. Since $x^k \to x^\infty$ and $d^k \to d^\infty$ for $k \in K$, and since $f(x)$ is continuously differentiable, there exists a $\delta' > 0$ such that, for sufficiently large $k \in K$,

$$\nabla f(x^k + \alpha d^k) d^k \leftarrow -\tfrac{1}{2}\tau \quad \text{for all } \alpha \in (0, \delta']. \tag{3.1}$$

Now, let $\bar{\delta} = \min\{\delta', \delta\} > 0$. Consider $k \in K$ sufficiently large. By condition (iii) and the fact that x^{k+1} solves the master problem over the set $\mathrm{H}(W^{k+1})$, which includes the point $(x^k + \bar{\delta} d^k)$, we must have $f(x^{k+1}) \leq f(x^k + \bar{\delta} d^k)$. By the mean value theorem,

$$f(x^k + \bar{\delta} d^k) = f(x^k) + \bar{\delta}\, \nabla f(\hat{x}^k) d^k,$$

where $\hat{x}^k$ is in the segment $(x^k, x^k + \bar{\delta} d^k)$. By (3.1), it then follows that

$$f(x^{k+1}) < f(x^k) - \tfrac{1}{2}\tau\bar{\delta} \quad \text{for } k \in K \text{ sufficiently large.} \tag{3.2}$$

Since the RSDCC algorithm generates a sequence of points with strictly decreasing objective function values, $\lim_{k\to\infty} f(x^k) = f(x^\infty)$. In particular, both $f(x^{k+1})$ and $f(x^k)$ approach $f(x^\infty)$ as $k \in K$ approaches ∞. Thus, from (3.2), we have $f(x^\infty) \leq f(x^\infty) - \frac{1}{2}\tau\bar{\delta}$, which is impossible since τ, $\bar{\delta} > 0$. Therefore, there could not be a subsequence with properties (i) through (iv). □

Theorem 3.2. *Consider problem* (2.1) *under Assumptions* 2.1–2.4. *Suppose that the sequence* $\{x^k\}$ *is generated by RSDCC. Then, any accumulation point of* $\{x^k\}$ *is an optimal solution.*

Proof (by contradiction). Assume that there exists a convergent subsequence $\{x^k\}_K$ with limit x^∞, not an optimal solution, and let z^∞ be the optimal function value of the subproblem. By Lemma 3.1, there exists an $\tau > 0$ such that $z^\infty = -\tau$. In addition, define $I_\infty = \{j \in N_q : g_j(x^\infty) = 0\}$. Assumption 2.4 guarantees the existence of a $\delta > 0$ such that $\delta = \min\{\|\nabla g_j(x^\infty)\| : j \in I_\infty\}$. For $k \in K$, consider the subproblem of RSDCC, and let (z^k, d^k) be an optimal solution. Since $\{d^k\}_K$ is bounded, there exists a subsequence $\{d^k\}_{K'}$ with limit d^∞. Furthermore, since $f(x)$ and all $g_j(x)$, for $j \in N_q$, are continuously differentiable, and $x^k \to x^\infty$, it follows that for $k \in K'$ sufficiently large, $z^k < -\frac{1}{2}\tau$ and $\|\nabla g_j(x^k)\| > \frac{1}{2}\delta$ for all $j \in I_\infty$. By definition of the subproblem

$$\nabla f(x^k) d^k \leq \|\nabla f(x^k)\| z^k < -\tfrac{1}{2}\|\nabla f(x^k)\| \tau, \tag{3.3}$$

$$g_j(x^k) + \nabla g_j(x^k) d^k \leq \|\nabla g_j(x^k)\| z^k < -\tfrac{1}{4}\delta\tau \quad \text{for } j \in I_\infty. \tag{3.4}$$

By the continuous differentiability of $f(x)$ and the fact that $\|\nabla f(x^\infty)\| \neq 0$, (3.3) implies that

$$\nabla f(x^\infty) d^\infty < 0.$$

Since all $g_j(x)$ are continuously differentiable, from (3.4) there exists a $\Gamma > 0$ such that the following inequality holds for each $\alpha \in (0, \Gamma]$ and $k \in K'$ sufficiently large:

$$g_j(x^k) + \nabla g_j(x^k + \alpha d^k) d^k < -\tfrac{1}{8}\delta\tau \quad \text{for } j \in I_\infty. \tag{3.5}$$

Now, let $\alpha \in (0, \Gamma]$. By the mean value theorem, and since $g_j(x^k) \leq 0$,

$$\begin{aligned} g_j(x^k + \alpha d^k) &= g_j(x^k) + \alpha \nabla g_j(x^k + \beta_{jk}\alpha d^k) d^k \\ &= (1-\alpha) g_j(x^k) + \alpha[g_j(x^k) + \nabla g_j(x^k + \beta_{jk}\alpha d^k) d^k] \\ &\qquad \text{for } j \in I_\infty, \end{aligned} \tag{3.6}$$

where $\beta_{jk} \in (0, 1]$. Since $\beta_{jk}\alpha \in (0, \Gamma]$, from (3.5) and (3.6) it follows that $g_j(x^k + \alpha d^k) \leq -\frac{1}{8}\alpha\delta\tau < 0$ for $j \in I_\infty$ and $k \in K'$ sufficiently large.

For $j \in I_\infty$ there exists a negative $\delta' = \max\{g_j(x^\infty) : j \in N_q, j \notin I_\infty\}$. Thus $g_j(x^k) < \frac{1}{2}\delta'$ for $j \notin I_\infty$ and $k \in K'$ sufficiently large, and by continuity there exists a $\Gamma' > 0$ such that for each $\alpha \in (0, \Gamma']$, $g_j(x^k + \alpha d^k) \leq 0$. It also holds that $A(x^k + \alpha d^k) \leq b$ for

$\alpha \in (0,1]$ and $k \in K'$. This shows that $x^k + \alpha d^k$ is feasible for each $\alpha \in (0, \min\{\Gamma, \Gamma', 1\})$ for all $k \in K'$ sufficiently large. Therefore, the sequence $\{(x^k, d^k)\}_{K'}$ satisfies conditions (i) through (iv) in Lemma 3.2. But this is a contradiction and x^∞ must be an optimal solution. □

4. Finite convergence

In order to show finite convergence of RSDCC, we need to consider the following additional assumptions:

Assumption 4.1. Problem (2.1) has a unique solution, denoted as x^*.

Assumption 4.2. x^* is in the interior of S' and $\|\nabla g_j(x^*)\|$ is bounded from above for $j \in N_q$.

When Assumption 4.2 holds, there exists $\varepsilon_1, \varepsilon_2 > 0$ such that

(i) $\varepsilon_1 = \min\{-g_j(x^*)/\|\nabla g_j(x^*)\|: j \in N_q\}$

(ii) ε_2 is the radius of the largest open ball $B(x^*, \varepsilon_2)$ around x^* such that $B(x^*, \varepsilon_2)$ is in S'.

The following four lemmas give conditions on the size of μ (selected at Step 0) with respect to ε_1 and ε_2 that are necessary to prove the final finiteness result.

Lemma 4.1. *Assume that $6\mu n^{1/2} \leqslant \varepsilon_1$. There exists an index π_1 such that for all $k \geqslant \pi_1$, if (z^k, d^k) solved the subproblem, then $d^k \in \arg\min\{\nabla f(x^k)d: Ad \leqslant b - Ax^k, 1^k \leqslant d \leqslant u^k\}$.*

Proof. For all $k \geqslant 0$,

$$\begin{aligned} &\|d^k\|_\infty \leqslant \max\{\|1^k\|_\infty, \|u^k\|_\infty\} \leqslant \mu + \|\bar{x} - x^k\|_\infty \leqslant \mu + \tfrac{1}{2}\mu = \tfrac{3}{2}\mu, \\ &\|d^k\| \leqslant \tfrac{3}{2}\mu n^{1/2}. \end{aligned} \tag{4.1}$$

Therefore, using the fact that $\nabla f(x^k) \neq 0$,

$$-\|\nabla f(x^k)\| z^k \leqslant -\nabla f(x^k) d^k \leqslant \|\nabla f(x^k)\| \, \|d^k\|$$

implies

$$-z^k = |z^k| \leqslant \|d^k\| \leqslant \tfrac{3}{2}\mu n^{1/2}. \tag{4.2}$$

Since $g_j(x)$ for $j \in N_q$ is continuously differentiable, from Assumption 4.2, there must exist an index π_1, such that for all $k \geqslant \pi_1$ the following holds:

$$\frac{-g_j(x^k)}{\|\nabla g_j(x^k)\|} > \tfrac{1}{2}\varepsilon_1 \quad \text{for } j \in N_q. \tag{4.3}$$

From (4.1)–(4.3) it follows that

$$\begin{aligned}\nabla g_j(x^k)d^k - \|\nabla g_j(x^k)\|z^k &\leqslant \|\nabla g_j(x^k)\|\,\|d^k\| - \|\nabla g_j(x^k)\|z^k \\ &= \|\nabla g_j(x^k)\|(\|d^k\| - z^k) \\ &\leqslant 3\mu n^{1/2}\|\nabla g_j(x^k)\| \\ &\leqslant \tfrac{1}{2}\varepsilon_1\|\nabla g_j(x^k)\| \\ &< -g_j(x^k) \quad \text{for } j\in N_q.\end{aligned}$$

Therefore, the constraints $\nabla g_j(x^k)d - \|\nabla g_j(x^k)\|z \leqslant -g_j(x^k)$ for $j\in N_q$ become inactive in the subproblem. This implies that the constraint $\nabla f(x^k)d - \|\nabla f(x^k)\|z \leqslant 0$ will always be active since the variable z forms the objective function. Thus, if (z^k, d^k) solves the subproblem, d^k must solve $\min\{\nabla f(x^k)d\colon Ad \leqslant b - Ax^k,\ 1^k \leqslant d \leqslant u^k\}$. □

Lemma 4.2. *There exists an index π_2, such that for all $k \geqslant \pi_2$, $\|\bar{x} - x^k\|_\infty \leqslant \frac{1}{2}\mu$.*

Proof. Since x^* is unique, by Assumption 4.2, there exists an open ball $B(x^*, \frac{1}{4}\mu)$ around x^* and an index Γ, such that $x^k \in B(x^*, \frac{1}{4}\mu)$ for all $k \geqslant \Gamma$. Now, consider the following two cases:

(i) If $\bar{x}$ does not change for $k \geqslant \Gamma$, then $\pi_2 = \Gamma$.

(ii) Otherwise, there exists an index $\tau > \Gamma$ such that $\|\bar{x} - x^\tau\|_\infty > \frac{1}{2}\mu$. In this case, Step 3(i) of RSDCC sets $\bar{x} = x^\tau$. This implies that for all $k \geqslant \tau$,

$$\|x^\tau - x^k\|_\infty \leqslant \|x^\tau - x^*\|_\infty + \|x^* - x^k\|_\infty < \tfrac{1}{4}\mu + \tfrac{1}{4}\mu = \tfrac{1}{2}\mu$$

and then $\pi_2 = \tau$. □

Lemma 4.3. *Set $\pi_3 = \max\{\pi_1, \pi_2\}$, where π_1 and π_2 are as defined by Lemmas 4.1 and 4.2, respectively. For all $k \geqslant \pi_3$, the points $x^k + d^k$ are extreme points of the following polytope:*

$$S(\bar{x}) = \{w\colon Aw \leqslant b,\ -\mu\bar{1} + \bar{x} \leqslant w \leqslant \mu\bar{1} + \bar{x}\}.$$

Proof. From Lemma 4.2, $\bar{x}$ does not change for all $k \geqslant \pi_3$. In addition, Lemma 4.1 states that the subproblem of RSDCC is equivalent to

$$\min\{\nabla f(x^k)d\colon Ad \leqslant b - Ax^k,\ l^k \leqslant d \leqslant u^k\}, \tag{4.4}$$

where $l^k = -\mu\bar{1} + (\bar{x} - x^k)$ and $u^k = \mu\bar{1} + (\bar{x} - x^k)$.

By performing the change of variables $w = x^k + d$, (4.4) becomes

$$\min\{\nabla f(x^k)(w - x^k)\colon Aw \leqslant b,\ -\mu\bar{1} + \bar{x} \leqslant w \leqslant \mu\bar{1} + \bar{x}\}$$

and the feasible region is $S(\bar{x})$. □

Lemma 4.4. *Given that* $k \geq \pi_3$ *(as defined by Lemma 4.3), if* $\mu(\frac{1}{4}+n^{1/2}) \leq \varepsilon_2$, *then* $S(\bar{x})$ *is in* S'.

Proof. Since $\pi_3 \geq \pi_2$, Lemma 4.2 imples that $x^k \in B(x^*, \frac{1}{4}\mu)$ for all $k \geq \pi_3$, which further implies that for all $w \in S(\bar{x})$,

$$\|x^* - w\| \leq \|x^* - \bar{x}\| + \|\bar{x} - w\| \leq \tfrac{1}{4}\mu + \mu n^{1/2} = \mu(\tfrac{1}{4} + n^{1/2}) \leq \varepsilon_2,$$

and, by definition of ε_2, $S(\bar{x})$ is in S'. □

The above lemmas show that if $\{x^k\}$ is the sequence generated by RSDCC and μ is small enough, then there exists an index π_3 such that for all $k \geq \pi_3$, the vector $\bar{x}$ does not change, and the subproblem generates directions $\{d^k\}$ such that $x^k + d^k$ are extreme points of the polytope $S(\bar{x})$. In addition, this polytope contains x^*, the optimal solution of (2.1). Therefore, problem (2.1) could also be stated as follows

$$\min\{f(x) \colon x \in \mathrm{H}(a_1, a_2, \ldots, a_N)\},$$

where $\{a_1, a_2, \ldots, a_N\}$ is the set of extreme points of $S(\bar{x})$. Define

$$I^*(\bar{x}) = \{a_j \colon \nabla f(x^*)(a_j - x^*) = 0, j \in N_N\}.$$

We refer to $\mathrm{H}(I^*(\bar{x}))$ as the optimal face of problem (2.1) with respect to $S(\bar{x})$.

The following lemma and theorems are similar to the ones for the linearly constrained case (Hearn, Lawphongpanich and Ventura, 1985) and, therefore, many of the proofs are omitted.

Lemma 4.5. *There exists an index* $\pi_4 \geq \pi_3$ *(as defined by Lemma 4.3) such that for any* $k \geq \pi_4$ *the following hold*:

(i) $\nabla f(x^k)(a_j - x^k) > 0$ *for all* $a_j \notin I^*(\bar{x})$,

(ii) $\min\{\nabla f(x^k)d \colon d \in S(\bar{x}) - x^k\} = \min\{\nabla f(x^k)d \colon d \in I^*(\bar{x}) - x^k\}$. □

Theorem 4.1. *If RSDCC generates an infinite sequence* $\{x^k\}$ *converging to* x^*, *then there exists an integer* $\pi \geq \pi_4$ *(as defined by Lemma 4.3) such that for all* $k \geq \pi$ *the following hold*:

(i) *The incoming grid point at iteration k,* y^k, *belongs to* $I^*(\bar{x})$.

(ii) W_s^k *is a subset of the smallest manifold containing* $I^*(\bar{x})$, *denoted as* $\mathrm{M}(I^*(\bar{x}))$.

Proof. When $k \geq \pi_4$, Lemma 4.4 shows that $S(\bar{x})$ is in S', which implies that $x^k + d^k$ is feasible. Then, (i) follows directly from Lemma 4.5(ii).

To prove (ii), assume that for some $k > \pi_4$, W_s^k is not in $\mathrm{M}(I^*(\bar{x}))$, in particular the elements $a_1, \ldots, a_p$ from W_s^k do not belong to $\mathrm{M}(I^*(\bar{x}))$. Then, using similar arguments to Lemma 4.5 (i), at the end of Step 2,

$$\nabla f(x^{k+1})(a_j - x^{k+1}) \geq 0 \quad \text{for } j \in N_p,$$

which further implies that the weight of a_j for $j \in N_p$ in the expression of x^{k+1} as a convex combination of the elements of W^{k+1} must all be zero. Therefore, we have that $a_1, \ldots, a_p$ satisfy the column dropping criteria in Step 2, which means that W_s^{k+1} is in $\mathrm{M}(I^*(\bar{x}))$. Moreover, (i) insures that only elements from $I^*(\bar{x})$ will be added to the set W_s^k in subsequent iterations, and the desired result is obtained by letting π equal $k+1$. □

Theorem 4.2. *In RSDCC, the set* $\mathrm{H}(W^k)$ *is a simplex for all* $k \geq 0$. □

Theorem 4.3. *If* $r \geq \dim I^*(x)+1$, *RSDCC converges to* x^* *after a finite number of major cycles.* □

5. A computational test

This section presents the results of testing the RSDCC algorithm on randomly generated medium- and large-scale quadratically constrained quadratic problems. Each test problem was solved by RSDCC with different values of r and, for comparison purposes, by the GRGA code developed by Abadie and Carpentier (1969) and Abadie (1975). The RSDCC computer program was written in double precision FORTRAN, compiled using the FORTRAN 77 compiler, and run on a VAX 8550 under the VAX/VMS operating system. The technique employed to solve the master problem was the projected Newton method of Bertsekas (1982) employed in the original RSD algorithm (Hearn, Lawphongpanich and Ventura, 1987). The maximum number of projections per iteration was set to $r+1$; however, since all test problems have a convex quadratic objective function, only one or two projections were usually required to determine a near optimal solution of the master problem with a relative duality gap $(\min\{\nabla f(x^k)(x^k-y^k)/|\nabla f(x^k)y^k|: y^k \in W^k\})$ less than or equal to 10^{-8}. At each projection, a Cholesky factorization (Forsythe and Moler, 1967) is employed to compute the search direction, and the stepsize for the line search is chosen by an Armijo-like rule (Bertsekas, 1976). The linear subproblem of RSDCC was solved by the linear programming subroutine LPSUBRS developed by Gill et al. (1983).

In our implementation of RSDCC, the box size μ that bounds the direction generated by the linear subproblem changes with the problem size and iteration number. For problems with up to 40 variables, the strategy was $\mu = \frac{1}{8}$ for the first 50 iterations, $\mu = \frac{1}{16}$ for the following 50 iterations, etc. until $\mu \leq 10^{-4}$; then it was kept constant for the remaining iterations. For problems with more than 40 variables, the same approach was used but starting with $\mu = \frac{1}{2}$.

In Step 1, the stepsize $\bar{\alpha}$ used to determine the incoming extreme point y^k is computed as follows. If $g_j(x^k+d^k) \leq 0$, $j \in N_q$, then $\bar{\alpha} = 1$. Otherwise, the bisection method in the interval (0, 1) is used. The number of evaluations of the constraint functions is reduced by analyzing first the constraints that are most infeasible at $\bar{\alpha} = 1$. The Insert sorting method (Knuth, 1973) puts the constraints in nonincreasing

Table 1

Summary of computational results for quadratic problems

Problem	Method	$RE \leq 10^{-2}$			$RE \leq 10^{-4}$			$RE \leq 10^{-6}$		
		Iter.	Proj.	CPU	Iter.	Proj.	CPU	Iter.	Proj.	CPU
	GRGA	24.6		10.2	113.8		21.3	197.6		32.2
$n=30$	$r=15$	13.2	19.2	2.8	36.4	67.8	10.1	113.8	276.8	36.1
$m=10$	$r=10$	13.2	19.2	2.7	36.6	71.0	10.0	115.8	281.8	36.2
	$r=5$	13.2	19.2	2.8	39.6	81.0	11.0	129.2	307.0	39.7
	$r=3$	13.2	19.2	2.7	50.8	93.6	14.2	147.6	308.6	43.8
	$r=1$	19.8	22.2	4.5	171.4	176.0	43.8	409.0	435.8	96.5
	GRGA	112.2		67.0	161.2		81.1	195.0		90.7
$n=40$	$r=15$	23.0	34.8	8.7	45.0	81.2	19.4	98.0	205.8	49.0
$m=15$	$r=10$	23.6	36.0	9.0	45.2	82.2	19.3	99.0	199.2	49.7
	$r=5$	23.6	36.0	9.1	45.6	84.2	19.5	105.6	222.2	52.5
	$r=3$	23.6	36.0	9.1	46.4	84.8	20.0	119.2	237.4	57.4
	$r=1$	28.4	36.0	11.1	182.0	228.6	80.0	317.0	358.8	147.7
	GRGA	90.0		116.7	230.8		199.1	307.8		235.2
$n=50$	$r=15$	18.8	30.6	13.5	44.2	84.6	36.1	96.6	200.4	91.9
$m=20$	$r=10$	18.8	30.6	13.5	44.2	84.6	36.4	97.0	202.8	93.3
	$r=5$	20.6	32.2	14.5	46.2	88.4	37.6	104.2	222.2	99.3
	$r=3$	20.6	32.2	14.6	63.2	115.4	51.8	131.2	252.8	121.7
	$r=1$	26.4	33.0	18.7	212.0	261.2	174.5	421.0	472.0	364.1
	GRGA	250.4		529.8	318.0		630.8	389.6		695.6
$n=60$	$r=15$	16.0	29.0	22.4	42.2	99.2	74.7	109.2	285.4	255.9
$m=25$	$r=10$	16.0	29.0	22.9	42.6	100.4	75.0	126.2	309.8	263.0
	$r=5$	16.2	29.0	23.9	51.6	128.8	88.9	144.6	310.8	296.8
	$r=3$	18.0	31.4	35.5	90.0	145.2	148.9	179.4	311.6	296.6
	$r=1$	65.2	75.4	130.9	200.6	212.2	350.3	400.8	424.0	607.1
	GRGA	127.2		554.2	273.0		912.1	324.2		1043.1
$n=70$	$r=15$	23.4	40.0	47.7	71.4	138.4	194.0	220.0	411.6	505.2
$m=30$	$r=10$	25.6	40.6	53.1	79.6	158.8	214.3	266.8	438.6	526.2
	$r=5$	26.6	41.0	55.5	111.2	187.0	284.2	283.2	419.4	589.4
	$r=3$	34.4	47.4	75.5	144.0	157.8	328.6	360.4	538.6	662.3
	$r=1$	111.2	125.2	215.6	400.2	429.6	722.8	*	*	*

* The indicated relative error was not achieved at the 500th iteration in at least one of the problems.

order of their function value at $\bar{\alpha} = 1$ $(g_j(x^k + d^k), j \in N_q)$. This order saves computational effort because often only the first constraint analyzed will be active at y^k. The bisection method terminates when the search interval is less than or equal to 10^{-4} for the first 50 iterations, 10^{-5} for the following 50 iterations, etc. When the tolerance becomes 10^{-8}, it is kept constant.

The quadratically constrained quadratic test problems have the following

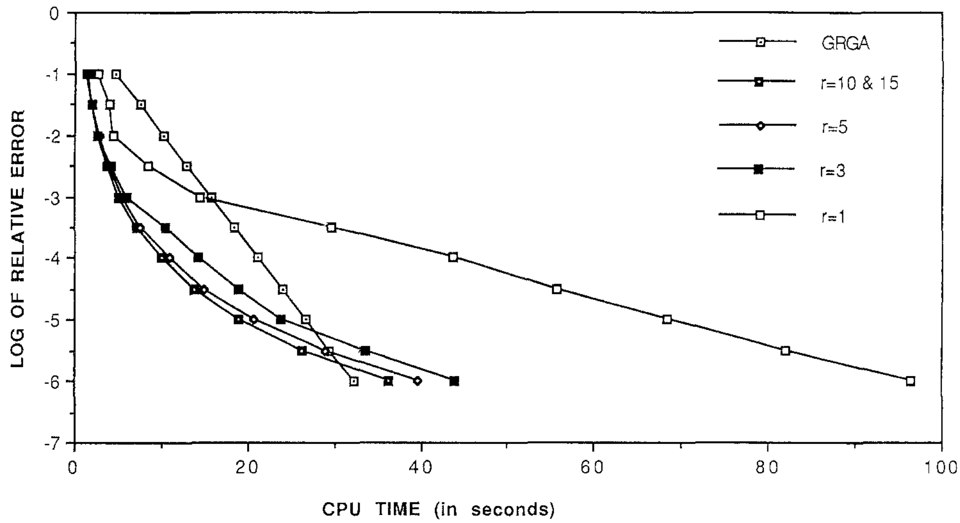

Fig. 1. Computational results for quadratic problems of size 30 * 10.

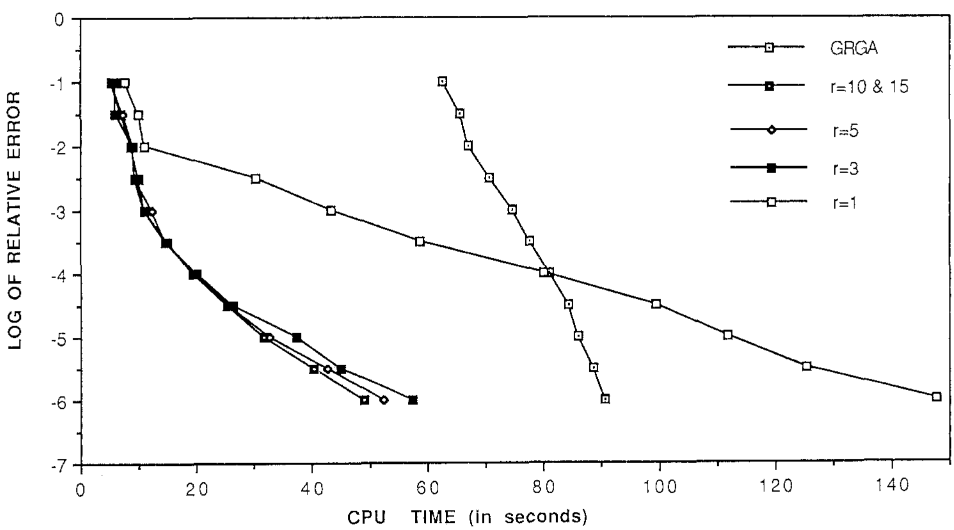

Fig. 2. Computational results for quadratic problems of size 40 * 15.

formulation:

$$\begin{aligned} \min \quad & f(x) = \tfrac{1}{2}x^{\mathrm{T}}Q_0x + c_0^{\mathrm{T}}x + d_0, \\ \text{s.t.} \quad & g_j(x) = \tfrac{1}{2}x^{\mathrm{T}}Q_jx + c_j^{\mathrm{T}}x + d_j \leqslant 0 \quad \text{for } j \in N_q. \end{aligned} \tag{5.1}$$

Data for the test problems was generated from uniform distributions represented by $U(a, b)$, where a and b are the lower and upper bounds, respectively, of the uniform random variable. Dense matrices Q_j were generated by the formula

$$Q_j = 2w_jw_j^{\mathrm{T}} + \mathrm{diag}(p_j),$$

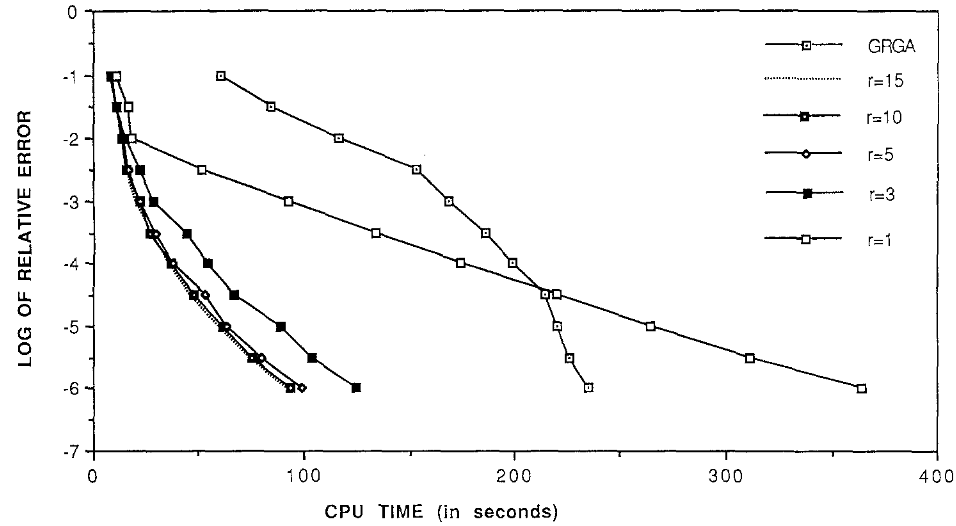

Fig. 3. Computational results for quadratic problems of size 50 * 20.

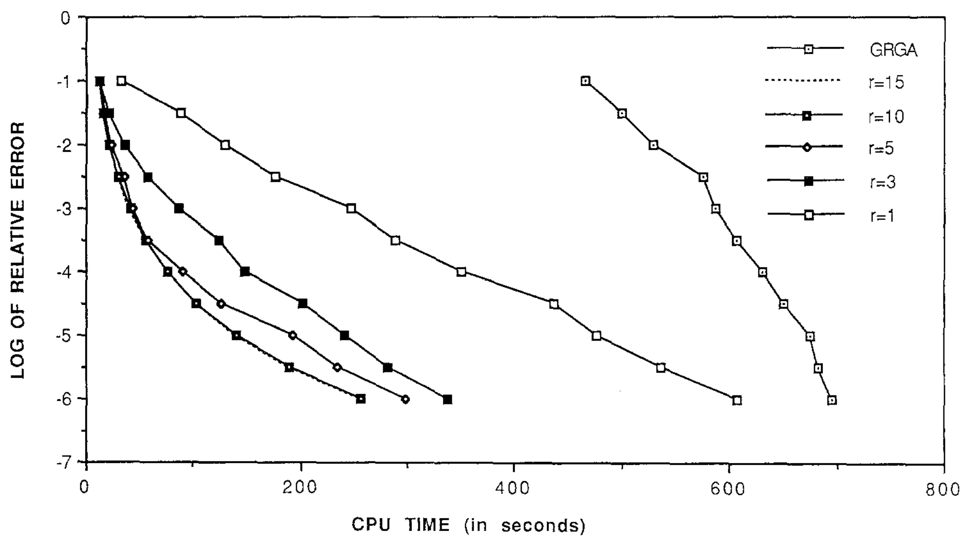

Fig. 4. Computational results for quadratic problems of size 60 * 25.

where

$$w_j = [w_{j(i)}], \quad w_{j(i)} \in U(0, 1), \qquad \text{for all } j,$$

$$p_0 = [p_{0(i)}], \quad p_{0(i)} \in U(1, 3),$$

$$p_j = [p_{j(i)}], \quad p_{j(i)} \in U(1, 4), \qquad \text{for } j \in N_q.$$

In addition, vectors c_j, scalars d_j, and the initial feasible solution x_0 were generated

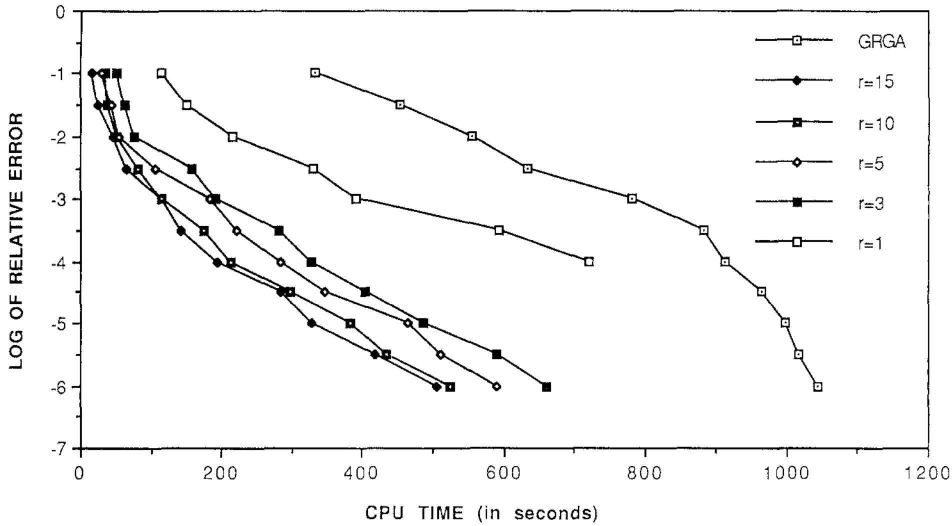

Fig. 5. Computational results for quadratic problems of size 70 * 30.

from the uniform distributions shown below.

$$c_0=[c_{0(i)}],\quad c_{0(i)}\in U(-5,-2),$$

$$c_j=[c_{j(i)}],\quad c_{j(i)}\in U(-3,-1),\qquad \text{for } j\in N_q,$$

$$d_0\in U(50,100),$$

$$x_{0(i)}\in U(-1.5,1.5),$$

$$d_j=\tfrac{1}{2}x_0^{\mathrm{T}}Q_jx_0+c_j^{\mathrm{T}}x_0+e_j,\quad e_j\in U(-2,0),\qquad \text{for } j\in N_q.$$

Note that the choice of d_j causes the initial solution to be interior to the feasible region, and in the problems generated, a check showed that 20–30% of the constraints were binding at the optimal solution.

Problems of five different sizes were generated. The problems ranged from 30 variables and 10 constraints to 70 variables and 30 constraints. For each problem size, five different test problems were run. Each problem was solved by RSDCC with $r=1$, 3, 5, 10 and 15, and by GRGA. Table 1 and Figures 1 to 5 present a summary of the computational results. For each problem size, Table 1 shows the average number of iterations, average number of master problem projections, and CPU time in seconds required to achieve the indicated relative error (RE $=[f(x^k)-f(x^*)]/|f(x^*)|$, where x^* is the optimal solution) of 10^{-2}, 10^{-4}, and 10^{-6}. The figures display the contrasting convergent behavior of the five different runs of RSDCC and GRGA. In the graphs, the x-axis represents CPU time, and the y-axis is log(RE), representing the number of correct digits of the current solution, i.e., $\log(\mathrm{RE})=-4$ indicates that the first four digits of the objective function value are correct. It is interesting to observe from Figures 2 to 5 that for the four largest problems RSDCC with $r\geqslant 3$ is far superior to RSDCC with $r=1$, the scaled

Topkis–Veinott algorithm, and it outperforms GRGA significantly. Only for the smallest problem (see Figure 1), does GRGA outperform RSDCC ($r \geq 3$) and then only for RE $\leq 10^{-5}$.

The main contribution of the RSDCC algorithm as an improvement of a feasible direction method seems substantiated by this computational test. The performance of the algorithm generally improves as the problem size and r increase. In particular, there is a decided improvement with $r > 1$ versus $r = 1$. In other words, on these test problems, "zig-zagging" is substantially reduced by the generalization of the line searches to optimization over simplices.

Acknowledgement

The authors gratefully acknowledge the constructive comments of the referees.

References

J. Abadie and J. Carpentier, "Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints," in: R. Fletcher, ed., *Optimization* (Academic Press, 1969).

J. Abadie, "Methods du gradient redirect generalise: Le code GRGA," Note HI 1756/00, Electricite de France (Paris, France, 1975).

D.P. Bertsekas, "on the Goldstein–Levitin–Polyak gradient projection method," *IEEE Transactions on Automatic Control* AC-21 (1976) 174–184.

D.P. Bertsekas, "Projected Newton methods for optimization problems with simple constraints," *SIAM Journal of Control and Optimization* 20 (1982) 221–246.

A.R. Colville, "A comparative study of nonlinear programming codes," in: H.W. Kuhn, ed., *Proceedings of the Princeton Symposium on Mathematical Programming* (Princeton University Press, Princeton, NJ, 1970).

G.E. Forsythe and C.B. Moler, *Computer Solution of Linear Algebraic Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1967).

P.E. Gill, W. Murray, M.A. Saunders and M.H. Wright, "User's guide for SOL/NPSOL+: A Fortran package for nonlinear programming," Technical Report SOL 83-12, Systems Optimization Laboratory, Stanford University (Stanford, CA, 1983).

D.W. Hearn, S. Lawphongpanich and J.A. Ventura, "Restricted simplicial decomposition: computation and extensions," *Mathematical Programming Study* 31 (1987) 99–118.

D.W. Hearn, S. Lawphongpanich and J.A. Ventura, "Finiteness in restricted simplicial decomposition," *Operations Research Letters* 4 (1985) 125–130.

D.E. Knuth, *The Art of Computer Programming, Vol. 3* (Addison-Wesley, Reading, MA, 1973).

O.L. Mangasarian, *Nonlinear Programming* (McGraw-Hill, New York, 1969).

D.M. Topkis and A.F. Veinott, "On the convergence of some feasible direction algorithms for nonlinear programming," *SIAM Journal of Control* 5 (1967) 268–279.

W.I. Zangwill, *Nonlinear Programming: a Unified Approach* (Prentice-Hall, Englewood Cliffs, NJ, 1969).