

NASA/CR-1998-208458
ICASE Report No. 98-31



A Globally Convergent Augmented Lagrangian Pattern Search Algorithm for Optimization with General Constraints and Simple Bounds

*Robert Michael Lewis
ICASE, Hampton, Virginia*

*Virginia Torczon
The College of William & Mary, Williamsburg, Virginia*

*Institute for Computer Applications in Science and Engineering
NASA Langley Research Center
Hampton, VA*

Operated by Universities Space Research Association



National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center
under Contract NAS1-97046

August 1998

A GLOBALLY CONVERGENT AUGMENTED LAGRANGIAN PATTERN SEARCH ALGORITHM FOR OPTIMIZATION WITH GENERAL CONSTRAINTS AND SIMPLE BOUNDS

ROBERT MICHAEL LEWIS * AND VIRGINIA TORCZON †

Abstract. We give a pattern search adaptation of an augmented Lagrangian method due to Conn, Gould, and Toint. The algorithm proceeds by successive bound constrained minimization of an augmented Lagrangian. In the pattern search adaptation we solve this subproblem approximately using a bound constrained pattern search method. The stopping criterion proposed by Conn, Gould, and Toint for the solution of this subproblem requires explicit knowledge of derivatives. Such information is presumed absent in pattern search methods; however, we show how we can replace this with a stopping criterion based on the pattern size in a way that preserves the convergence properties of the original algorithm. In this way we proceed by successive, inexact, bound constrained minimization without knowing exactly how inexact the minimization is. So far as we know, this is the first provably convergent direct search method for general nonlinear programming.

Key words. augmented Lagrangian, constrained optimization, direct search, nonlinear programming, pattern search

Subject classification. Applied and Numerical Mathematics

1. Introduction. In this paper we consider the extension of pattern search methods to nonlinearly constrained minimization. We will consider problems of the form

$$(1.1) \quad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c(x) = 0 \\ & \ell \leq x \leq u, \end{array}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c(x) = (c_1(x), \dots, c_m(x))$. We allow the possibility that some of the variables are unbounded either above or below by permitting $\ell_j, u_j = \pm\infty$, $j \in \{1, \dots, n\}$. This formulation assumes that any general inequality constraints have been converted into equality constraints by the introduction of non-negative slack variables, leaving bounds as the only explicit inequality constraints.

The pattern search method that we will discuss here is an adaptation of an augmented Lagrangian method due to Conn, Gould, and Toint [6]. The latter method is the basis for the subroutine **AUGLG** in the **LANCELOT** optimization package [7]. The method of Conn, Gould, and Toint involves successive bound constrained minimization of an augmented Lagrangian. Since pattern search methods have recently been extended to bound constrained minimization [19, 21], an adaptation of the augmented Lagrangian method

*ICASE, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia, 23681-2199, buckaroo@icase.edu. This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-97046 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23681-2199.

†Department of Computer Science, College of William & Mary, P. O. Box 8795, Williamsburg, Virginia 23187-8795, va@cs.wm.edu. This research was supported by the National Science Foundation under Grant CCR-9734044 and by the National Aeronautics and Space Administration under NASA Contract No. NAS1-97046 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, Virginia 23681-2199.

of Conn, Gould, and Toint to pattern search naturally suggests itself. Furthermore, the multiplier update of Algorithm 1 in [6] does not involve information about derivatives of the objective or constraints, so the augmented Lagrangian approach is consistent with the derivative-free nature of pattern search algorithms.

Since there exist broad classes of pattern search methods for unconstrained [20, 34, 35, 36] and bound constrained minimization [19, 21], it seems to us natural to first extend pattern search methods to nonlinearly constrained minimization via algorithms that proceed by successive unconstrained or bound constrained minimization, such as the augmented Lagrangian method we discuss here. In the absence of information about derivatives of the objective and constraints, it is difficult to design pattern search algorithms for general nonlinearly constrained minimization that produce only feasible directions or feasible iterates. This is due to the fact that a pattern in a pattern search algorithm would need to include a sufficiently rich set of search directions to capture any feasible improvement in the objective. When nonlinear constraints are present, it is not clear how to design such a pattern without first-order information.

We will show that despite the absence of an explicit estimation of any derivatives (a characteristic of pattern search methods), our pattern search augmented Lagrangian approach exhibits all of the first-order convergence properties of the original algorithm of Conn, Gould, and Toint. This at first is surprising, since the original algorithm allows its subproblems to be solved approximately, and the stopping criterion for the solution of the subproblems is based on the magnitude of a measure of first-order stationarity for bound constrained minimization. This information is not explicitly available in a direct search method. However, as we discuss in §5.1, there is a correlation between the size of the pattern in bound constrained pattern search and the amount of local feasible descent. Using this correlation we are able to establish convergence even without explicit knowledge of derivatives. That is, we are able to proceed by successive, inexact minimization of the augmented Lagrangian via pattern search methods, even without knowing exactly how inexact the minimization is.

This is the main contribution of the work presented here, and shows how one can use pattern search in a practical algorithm for nonlinear programming. Otherwise, the extension of pattern search to constrained minimization by means of the augmented Lagrangian approach of Conn, Gould, and Toint is straightforward, due to the strength and generality of the convergence analysis presented in [6].

The question of treating general nonlinear constraints with direct search minimization algorithms has a long history, beginning with the original work on direct search methods. Rosenbrock, in [28], proposed treating constraints using his rotating directions method by redefining the objective near the boundary of the feasible region in a way that would tend to keep the iterates feasible, a form of penalization. Similar ideas for modifying the objective in the case of bound constraints are discussed by Spendley, Hext, and Himsworth [30] and Nelder and Mead [24] in connection with their simplex-based methods. In these approaches the objective is given a suitably large value (in the case of minimization) at all infeasible points.

More systematic approaches to penalization have also appeared. The treatment of inequality constraints via exact, non-smooth penalization (though not by that name) appears as early as the work of Hooke and Jeeves [15]. More recently, Kearsley and Glowinski [13, 16] have applied pattern search methods to equality constrained problems arising in control via exact, non-smooth penalization. Weisman's MINIMAL algorithm [14] applies the pattern search algorithm of Hooke and Jeeves to a non-smooth quadratic penalty function and incorporates an element of random search. Davies and Swann [8], in connection with applying the pattern search method of Hooke and Jeeves to constrained optimization, recommend the use of the reciprocal barrier method of Carroll [5, 11].

A direct search method for constrained minimization that has proven very popular in application is

M. J. Box's Complex method [3], which was originally developed to address difficulties encountered with Rosenbrock's method. In this algorithm, the objective is sampled at a broader set of points than in the simplex-based methods as a way to avoid premature termination. There is also an element of random search involved. The ACSIM algorithm of Dixon [10] is a sophisticated direct search algorithm, combining ideas from the Nelder-Mead simplex method and the Complex method with elements of hem-stitching and quadratic modeling to accelerate convergence.

In the special case of bound constraints, Spendley also suggested the expedient of simply setting to the corresponding bound any variable that was tending to go infeasible [29]. In [17], Keefer proposed a hybrid, feasible iterates algorithm for bound constrained minimization that uses the algorithm of Nelder-Mead in the interior of the feasible region and the method of Hooke and Jeeves at the boundary, since the pattern in the algorithm of Hooke and Jeeves conforms in a natural way to the boundary of the feasible region. In the case of linear constraints there is the algorithm of May [22], which is an extension of Mifflin's derivative-free unconstrained minimization method in [23]. This algorithm also takes into account the particular geometry of the feasible region.

Others have proposed modifications of the method of Hooke and Jeeves along the lines of feasible directions algorithms. These methods involve a limited calculation of sensitivity information to compute feasible directions at the boundary of the feasible region if the algorithm appears to have stalled. Klingman and Himmelblau [18] give an algorithm with a simple construction of a suitable feasible direction. The method of Glass and Cooper [12] is more sophisticated, and computes a new search direction by solving a linear programming problem involving a linear approximation of the objective and constraints, just as one would in a derivative-based feasible directions algorithm.

Finally, we note the flexible tolerance method of Paviani and Himmelblau [14, 25]. This algorithm, based on the method of Nelder and Mead, alternatively attempts to reduce the objective and constraint violation, depending on the extent to which the iterates are infeasible.

These proposals for direct search algorithms for constrained minimization, while they have often proven effective, have not been accompanied by any convergence analysis. A notable exception is May's algorithm for linearly constrained minimization [22]; his sufficient decrease criterion for accepting steps enables him to prove global convergence. More recently, provably convergent, feasible iterate pattern search algorithms for bound constrained and linearly constrained minimization were developed in [19, 21]; we apply the analysis for bound constrained pattern search methods in the present work.

2. The augmented Lagrangian method of Conn, Gould, and Toint. We begin by reviewing the augmented Lagrangian approach in [6]. To facilitate comparison of the pattern search approach with the original algorithm, we will adhere to the notation of [6] throughout.

The augmented Lagrangian is

$$(2.1) \quad \Phi(x; \lambda, S, \mu) = f(x) + \sum_{i=1}^m \lambda_i c_i(x) + \frac{1}{2\mu} \sum_{i=1}^m s_{ii} c_i(x)^2.$$

The vector $\lambda = (\lambda_1, \dots, \lambda_m)^T$ is the Lagrange multiplier estimate, μ is the penalty parameter, and the entries s_{ii} of the diagonal matrix S are positive weights. The equality constraints of (1.1) are incorporated in the augmented Lagrangian Φ while the simple bounds are left explicit. For a particular choice of multiplier $\lambda^{(k)}$, penalty parameter $\mu^{(k)}$, and scaling $S^{(k)}$, we define

$$\Phi^{(k)}(x) = \Phi(x; \lambda^{(k)}, S^{(k)}, \mu^{(k)}).$$

Given an iterate $x^{(k)}$, we define

$$\nabla_x \Phi^{(k)} = \nabla_x \Phi(x^{(k)}; \lambda^{(k)}, S^{(k)}, \mu^{(k)}).$$

Conn, Gould, and Toint define the first-order Lagrange multiplier update to be

$$(2.2) \quad \bar{\lambda}(x, \lambda, S, \mu) = \lambda + Sc(x)/\mu.$$

This is a form of the Hestenes-Powell multiplier update for the augmented Lagrangian (2.1). For the purposes of a pattern search augmented Lagrangian approach, which assumes no explicit knowledge of derivative information, one appears to have no choice other than some variant of the Hestenes-Powell multiplier update. All other multiplier update formulae (such as those discussed in [1, 32]) require information about derivatives.

The projection onto the convex set $B = \{ x \mid \ell \leq x \leq u \}$ will be denoted by P ; it is defined component-wise by

$$(P[x])_i = \begin{cases} \ell_i & \text{if } x_i \leq \ell_i \\ u_i & \text{if } x_i \geq u_i \\ x_i & \text{otherwise.} \end{cases}$$

Given $x \in B$ and a vector v , we define

$$P(x, v) = x - P[x - v].$$

Unless otherwise noted, we use $\|\cdot\|$ to denote the Euclidean vector norm or its induced matrix norm.

We base our augmented Lagrangian pattern search method on Algorithm 1 of [6]. The original algorithm follows.

Step 0 [Initialization]. An initial vector of Lagrange multiplier estimates $\lambda^{(0)}$ is given. The positive constants $\eta_0, \mu_0, \omega_0, \tau < 1, \gamma_1 < 1, \omega_* \ll 1, \eta_* \ll 1, \alpha_\omega, \beta_\omega, \alpha_\eta$, and β_η are specified. The diagonal matrices S_1 and S_2 , for which $0 < S_1^{-1} \leq S_2 < \infty$, are given (the inequalities are to be understood element-wise for the diagonal elements). Set $\mu^{(0)} = \mu_0$, $\alpha^{(0)} = \min(\mu^{(0)}, \gamma_1)$, $\omega^{(0)} = \omega_0(\alpha^{(0)})^{\alpha_\omega}$, $\eta^{(0)} = \eta_0(\alpha^{(0)})^{\alpha_\eta}$, and $k = 0$.

Step 1 [Inner iteration]. Define a scaling matrix $S^{(k)}$ for which $S_1^{-1} \leq S^{(k)} \leq S_2$. Find $x^{(k)} \in B$ such that

$$(2.3) \quad \|P(x^{(k)}, \nabla_x \Phi^{(k)})\| \leq \omega^{(k)}.$$

If

$$\|c(x^{(k)})\| \leq \eta^{(k)},$$

execute Step 2. Otherwise, execute Step 3.

Step 2 [Test for convergence and update Lagrange multiplier estimates]. If $\|P(x^{(k)}, \nabla_x \Phi^{(k)})\| \leq \omega^*$ and $\|c(x^{(k)})\| \leq \eta_*$, stop. Otherwise, set

$$\lambda^{(k+1)} = \bar{\lambda}(x^{(k)}, \lambda^{(k)}, S^{(k)}, \mu^{(k)})$$

$$\mu^{(k+1)} = \mu^{(k)}$$

$$\alpha^{(k+1)} = \min(\mu^{(k+1)}, \gamma_1)$$

$$\omega^{(k+1)} = \omega^{(k)}(\alpha^{(k+1)})^{\beta_\omega}$$

$$\eta^{(k+1)} = \eta^{(k)}(\alpha^{(k+1)})^{\beta_\eta},$$

increment k by one and go to Step 1.

Step 3 [Reduce the penalty parameter]. Set

$$\begin{aligned}\lambda^{(k+1)} &= \lambda^{(k)} \\ \mu^{(k+1)} &= \tau\mu^{(k)} \\ \alpha^{(k+1)} &= \min(\mu^{(k+1)}, \gamma_1) \\ \omega^{(k+1)} &= \omega_0(\alpha^{(k+1)})^{\alpha_\omega} \\ \eta^{(k+1)} &= \eta_0(\alpha^{(k+1)})^{\alpha_\eta},\end{aligned}$$

increment k by one and go to Step 1.

3. Bound constrained pattern search algorithms. We next review the relevant features of the general pattern search method for the bound constrained problem

$$(3.1) \quad \begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && \ell \leq x \leq u. \end{aligned}$$

As noted in [19], a number of “classical” pattern search algorithms are suitable for bound constrained minimization, including

- coordinate search with fixed step lengths [26],
- evolutionary operation using composite designs ([2] and [4, 31]),
- the original pattern search method of Hooke and Jeeves [15], and
- the multidirectional search algorithm ([33, 34] and [9])

For a further discussion, see [19, 21].

3.1. The pattern. The index j will denote the iteration in a pattern search method. A pattern $P^{(j)}$ is a matrix $P^{(j)} \in \mathbf{Z}^{n \times p_j}$, where $p_j > n + 1$. There is no upper bound on p_j . We partition the pattern into components

$$P^{(j)} = [\Gamma^{(j)} \quad L^{(j)}].$$

We require that $\Gamma^{(j)} \in \mathbf{Z}^{n \times r_j}$ belong to a finite set of matrices $\mathbf{\Gamma}$ and that $L^{(j)} \in \mathbf{Z}^{n \times (p_j - r_j)}$ contains at least one column, a column of zeroes. The inclusion of a column of zeroes is simply a formalism to allow for a zero step, i.e., $x^{(j+1)} = x^{(j)}$.

The matrices $\Gamma^{(j)}$ must satisfy certain conditions, discussed more fully in [19, 21], that ensure that near the boundary of the bound constrained feasible region we always have a set of generators for any possible tangent cone. This, in turn, means that we can capture any feasible improvement in the objective.

For the purposes of this discussion, the reader may assume that $\Gamma^{(j)} = [I \quad -I]$, or, more generally,

$$(3.2) \quad \Gamma^{(j)} = [D^{(j)} \quad -D^{(j)}],$$

where

$$(3.3) \quad D^{(j)} = \text{diag}(d_i^{(j)}), \quad i = 1, \dots, n.$$

This was the prescription for the pattern given in [19]. In [21] this condition is relaxed so that a bound constrained pattern search algorithm can behave like a pattern search algorithm for unconstrained minimization in the interior of the feasible region or in a subspace of unbounded variables.

At iteration j , given $\Delta^{(j)} \in \mathbb{R}$ with $\Delta^{(j)} > 0$, we define a *trial step* to be a vector of the form $s_i^{(j)} = \Delta^{(j)} c_i^{(j)}$ for some $i \in \{1, \dots, p_j\}$, where $c_i^{(j)}$ denotes the i th column of $P^{(j)}$ (i.e., $P^{(j)} = [c_1^{(j)} \dots c_{p_j}^{(j)}]$). We call a trial step $s_i^{(j)}$ *feasible* for (3.1) if $(x^{(j)} + s_i^{(j)}) \in B = \{x \mid \ell \leq x \leq u\}$. At iteration j , a *trial point* is any point of the form $x_i^{(j)} = x^{(j)} + s_i^{(j)}$, where $x^{(j)}$ is the current iterate.

3.2. The bound constrained exploratory moves. Pattern search methods proceed by conducting a series of exploratory moves about the current iterate $x^{(j)}$ to choose a new iterate $x^{(j+1)} = x^{(j)} + s^{(j)}$, for some feasible step $s^{(j)}$ determined during the course of the exploratory moves. The following hypotheses on the result of the bound constrained exploratory moves allow a broad choice of exploratory moves while ensuring the properties required to prove convergence. By abuse of notation, if A is a matrix, $y \in A$ means that the vector y is a column of A .

1. $s^{(j)} \in \Delta^{(j)} P^{(j)} \equiv \Delta^{(j)} [\Gamma^{(j)} L^{(j)}]$.
2. $(x^{(j)} + s^{(j)}) \in B = \{ x \mid \ell \leq x \leq u \}$.
3. If $\min \{ f(x^{(j)} + y) \mid y \in \Delta^{(j)} \Gamma^{(j)}, x^{(j)} + y \in B \} < f(x^{(j)})$,
then $f(x^{(j)} + s^{(j)}) < f(x^{(j)})$.

FIG. 3.1. Hypotheses on the result of the bound constrained exploratory moves.

3.3. The bound constrained pattern search method. Fig. 3.2 states the generalized pattern search method for minimization with bound constraints. To define a particular pattern search method, we must specify the pattern $P^{(j)}$, the bound constrained exploratory moves to be used to produce a feasible step $s^{(j)}$, and the algorithms for updating $P^{(j)}$ and $\Delta^{(j)}$.

Let $x^{(0)} \in B$ and $\Delta^{(0)} > 0$ be given.

For $j = 0, 1, \dots$,

- a) Compute $f(x^{(j)})$.
- b) Determine a step $s^{(j)}$ using a bound constrained exploratory moves algorithm.
- c) If $f(x^{(j)} + s^{(j)}) < f(x^{(j)})$, then $x^{(j+1)} = x^{(j)} + s^{(j)}$. Otherwise $x^{(j+1)} = x^{(j)}$.
- d) Update $P^{(j)}$ and $\Delta^{(j)}$.

FIG. 3.2. The Generalized Pattern Search Method for Bound Constrained Problems.

3.4. The updates. The aim of the update of $\Delta^{(j)}$ is to force a strict reduction in f . An iteration with $f(x^{(j)} + s^{(j)}) < f(x^{(j)})$ is *successful*; otherwise, the iteration is *unsuccessful*. Note that to accept a step we only require *simple*, as opposed to *sufficient*, decrease. We cannot increase or decrease $\Delta^{(j)}$ in an arbitrary manner (as is detailed more fully in [19, 21]), but for the purposes of analyzing the augmented Lagrangian pattern search algorithm, the update of $\Delta^{(j)}$ can be summarized as

$$(3.4) \quad \text{If } f(x^{(j)} + s^{(j)}) < f(x^{(j)}) \text{ then } \Delta^{(j+1)} \geq \Delta^{(j)}.$$

$$(3.5) \quad \text{If } f(x^{(j)} + s^{(j)}) \geq f(x^{(j)}) \text{ then } \Delta^{(j+1)} < \Delta^{(j)}.$$

If an iteration is successful it may be possible to increase the step length parameter $\Delta^{(j)}$, but $\Delta^{(j)}$ is not allowed to decrease. Whereas if an iteration is unsuccessful, the step length parameter $\Delta^{(j)}$ must be decreased. Again, we refer the reader to [19, 21] for the details.

4. The pattern search augmented Lagrangian method. At iteration k of the original augmented Lagrangian algorithm described in §2, we approximately solve the subproblem

$$(4.1) \quad \begin{aligned} &\text{minimize} && \Phi^{(k)}(x) \\ &\text{subject to} && \ell \leq x \leq u. \end{aligned}$$

The degree to which this subproblem must be solved is given by (2.3). We adapt Algorithm 1 in [6] to pattern search by solving the bound constrained subproblem using a bound constrained pattern search method. However, pattern search methods do not have recourse to derivatives or explicit approximations thereof. For that reason we must replace the stopping criterion (2.3) with one that is appropriate to a pattern search method.

We replace (2.3) with a new criterion on the size of the pattern. As we discuss in §5, we retain the convergence properties of the original Conn, Gould and Toint algorithm because the size of the pattern and the stationarity condition (2.3) are correlated, even though we do not have explicit control of $\| P(x^{(k)}, \nabla_x \Phi^{(k)}) \|$.

We now state the augmented Lagrangian pattern search algorithm. At iteration k in the outermost loop of the algorithm, we will denote by $\{x^{(k,j)}\}$ the sequence of iterates produced in the solution of (4.1) via a bound constrained pattern search algorithm. We also assume that there exists d^* such that $|d_i^{(k,j)}| \leq d^*$ for all k , where the $d_i^{(k,j)}$ are the diagonal entries in (3.3). This uniformity in the pattern search algorithms used in the successive minimization of the augmented Lagrangian is not at all restrictive. An obvious way in which to accomplish this is simply to choose for all k the same set Γ in the definition of the pattern search algorithms (see §3.1).

In order to relate the stopping criterion in the pattern search solution of the subproblems to the multiplier estimates and the penalty parameter, we introduce the function

$$\theta(\lambda, \mu) = (1 + \|\lambda\| + 1/\mu)^{-1}.$$

We note that any function $\theta(\lambda, \mu)$ such that

$$\theta(\lambda, \mu) = O((\|\lambda\| + 1/\mu)^{-1})$$

as $(\|\lambda\| + 1/\mu) \rightarrow \infty$ will suffice for the purposes of proving convergence.

Step 0 [Initialization]. An initial vector of Lagrange multiplier estimates $\lambda^{(0)}$ is given. The positive constants $\eta_0, \mu_0, \omega_0, \tau < 1, \gamma_1 < 1, \delta_* \ll 1, \eta_* \ll 1, \alpha_\omega, \beta_\omega, \alpha_\eta$, and β_η are specified. The diagonal matrices S_1 and S_2 , for which $0 < S_1^{-1} \leq S_2 < \infty$, are given (the inequalities are to be understood element-wise for the diagonal elements). Set $\mu^{(0)} = \mu_0$, $\alpha^{(0)} = \min(\mu^{(0)}, \gamma_1)$, $\omega^{(0)} = \omega_0(\alpha^{(0)})^{\alpha_\omega}$, $\delta^{(0)} = \theta(\lambda^{(0)}, \mu^{(0)})\omega^{(0)}$, $\eta^{(0)} = \eta_0(\alpha^{(0)})^{\alpha_\eta}$, and $k = 0$.

Step 1 [Inner iteration]. Define a scaling matrix $S^{(k)}$ for which $S_1^{-1} \leq S^{(k)} \leq S_2$. Apply the bound constrained pattern search method to

$$(4.2) \quad \begin{array}{ll} \text{minimize} & \Phi^{(k)}(x) \\ \text{subject to} & \ell \leq x \leq u \end{array}$$

to find $x^{(k)} = x^{(k,j)} \in B$ such that the pattern is sufficiently small,

$$(4.3) \quad \Delta^{(k,j)} \leq \delta^{(k)},$$

and we do *not* find an acceptable step in the part of the pattern $P^{(k,j)}$ corresponding to $\Gamma^{(k,j)}$,

$$(4.4) \quad f(x^{(k,j)} + s^{(k,j)}) \geq f(x^{(k,j)}) \text{ for all } s^{(k,j)} \in \Delta^{(k,j)}\Gamma^{(k,j)}.$$

The latter is the case, for instance, in the event of an unsuccessful step.

If

$$\|c(x^{(k)})\| \leq \eta^{(k)},$$

execute Step 2. Otherwise, execute Step 3.

Step 2 [Test for convergence and update Lagrange multiplier estimates]. If $\delta^{(k)} \leq \delta^*$ and $\|c(x^{(k)})\| \leq \eta_*$, stop. Otherwise, set

$$\begin{aligned}\lambda^{(k+1)} &= \bar{\lambda}(x^{(k)}, \lambda^{(k)}, S^{(k)}, \mu^{(k)}) \\ \mu^{(k+1)} &= \mu^{(k)} \\ \alpha^{(k+1)} &= \min(\mu^{(k+1)}, \gamma_1) \\ \omega^{(k+1)} &= \omega^{(k)} (\alpha^{(k+1)})^{\beta_\omega} \\ \delta^{(k+1)} &= \theta(\lambda^{(k+1)}, \mu^{(k+1)}) \omega^{(k+1)} \\ \eta^{(k+1)} &= \eta^{(k)} (\alpha^{(k+1)})^{\beta_\eta},\end{aligned}$$

increment k by one and go to Step 1.

Step 3 [Reduce the penalty parameter]. Set

$$\begin{aligned}\lambda^{(k+1)} &= \lambda^{(k)} \\ \mu^{(k+1)} &= \tau \mu^{(k)} \\ \alpha^{(k+1)} &= \min(\mu^{(k+1)}, \gamma_1) \\ \omega^{(k+1)} &= \omega_0 (\alpha^{(k+1)})^{\alpha_\omega} \\ \delta^{(k+1)} &= \theta(\lambda^{(k+1)}, \mu^{(k+1)}) \omega^{(k+1)} \\ \eta^{(k+1)} &= \eta_0 (\alpha^{(k+1)})^{\alpha_\eta},\end{aligned}$$

increment k by one and go to Step 1.

Note that we have replaced the stopping criterion (2.3) for the inner iteration of Algorithm 1 in [6] with (4.3)–(4.4), which are stopping criteria based on the size of the pattern, because we do not assume explicit information about the derivatives. The remaining modifications to Algorithm 1 in [6] are to correctly manage the sequence $\{\delta^{(k)}\}$, which controls the stopping criteria we have introduced. The question now remains: having removed an exact specification of how inexact the solution of the subproblem can be (i.e., (2.3)), are the weaker conditions we have introduced (i.e., (4.3)–(4.4)) sufficient to guarantee that (2.3) will be satisfied asymptotically? An answer in the affirmative is provided in the next section.

5. Convergence analysis. We now discuss the convergence properties of the augmented Lagrangian pattern search algorithm. As we shall see, altering the original algorithm by solving the bound constrained subproblem via pattern search does leaves the convergence properties of the original algorithm almost entirely unchanged.

In [6], Conn, Gould, and Toint call a component of $x^{(k)}$ *floating* if

$$\ell_i < x_i^{(k)} - (\nabla_x \Phi^{(k)})_i < u_i.$$

For a convergent subsequence $\{x^{(k)}\}$, $k \in K$, with limit point x^* they define the index set

$$I_1 = \{i \mid x_i^{(k)} \text{ are floating for all } k \in K \text{ sufficiently large and } \ell_i < x_i^* < u_i\},$$

and let $\hat{A}(x)$ denote the corresponding columns of the Jacobian of $c(x)$, where $A(x)$ is the entire Jacobian of $c(x)$.

The following assumptions are made in [6].

AS1. *The functions $f(x)$ and $c(x)$ are twice continuously differentiable for all $x \in B$.*

AS2. The iterates $\{x^{(k)}\}$ considered lie within a closed, bounded domain Ω .

AS3. The matrix $\hat{A}(x^*)$ has column rank no smaller than m at any limit point x^* of the sequences $\{x^{(k)}\}$ considered in this paper.

In addition, in order to be assured that a bound constrained pattern search algorithm applied to the subproblem (4.2) will find an iterate satisfying (4.3)–(4.4), we assume the following.

PS1. For a given k , the set $B \cap \{x \mid \Phi^{(k)}(x) \leq \Phi^{(k)}(x^{(k,0)})\}$ is compact.

That is, we assume compactness of the set of $x \in B$ for which the augmented Lagrangian is less than the value of the augmented Lagrangian at the point at which we begin the solution of the subproblem. This is not a particularly restrictive assumption, as we discuss further in the context of inequality constrained minimization in §6, but it is necessary to ensure convergence of any pattern search method applied to the subproblem (4.2).

Under hypothesis (PS1), we are assured that in the inner iteration (the pattern search minimization of the bound constrained augmented Lagrangian),

$$\liminf_{j \rightarrow +\infty} \Delta^{(k,j)} = 0$$

(see [19, 21]), so the termination criterion (4.3) eventually will be satisfied. Moreover, the update rules (3.4)–(3.5) only allow $\Delta^{(k,j)}$ to be decreased at unsuccessful steps, where (4.4) holds. Thus both termination criteria (4.3) and (4.4) will eventually be satisfied, the pattern search solution of the augmented Lagrangian subproblem will halt, and the overall iteration of the pattern search augmented Lagrangian algorithm is well-defined.

5.1. The relationship between the pattern size and stationarity. For convenience, let

$$q^{(k,j)} = P(x^{(k,j)}, \nabla_x \Phi^{(k)}(x^{(k,j)})).$$

The following result is the key to analyzing the augmented Lagrangian pattern search method.

PROPOSITION 5.1. *There exists $C_{5.1}$, independent of k , such that*

$$\|P(x^{(k)}, \nabla_x \Phi^{(k)})\| \leq C_{5.1} \omega^{(k)}$$

for all k .

Proof. Given k , $x^{(k)} \equiv x^{(k,j)}$ for some j . By design we have $d^* > 0$ such that $|d_i^{(k,j)}| \leq d^*$ for all i, j , and k , where $d_i^{(k,j)}$ is as defined in (3.2).

First suppose

$$(5.1) \quad \Delta^{(k,j)} \geq \frac{\|q^{(k,j)}\|_\infty}{d^*}.$$

Then (5.1), (4.3), and the rule for updating $\delta^{(k)}$ in either Step 2 or Step 3 give us

$$\|q^{(k,j)}\|_\infty \leq d^* \Delta^{(k,j)} \leq d^* \delta^{(k)} \leq d^* \omega^{(k)}$$

and so

$$(5.2) \quad \|q^{(k,j)}\| \leq n^{\frac{1}{2}} d^* \omega^{(k)}.$$

On the other hand, suppose

$$\Delta^{(k,j)} < \frac{\|q^{(k,j)}\|_\infty}{d^*}.$$

The proof of Proposition 5.2 in [19] shows that if $\Delta^{(k,j)} < \|q^{(k,j)}\|_\infty/d^*$, then there is a trial step $s_i^{(k,j)} \in \Delta^{(k,j)}\Gamma^{(k,j)}$ such that $x^{(k,j)} + s_i^{(k,j)} \in B$ and

$$(5.3) \quad \nabla_x \Phi^{(k)}(x^{(k,j)})^T s_i^{(k,j)} < -n^{-\frac{1}{2}} \|q^{(k,j)}\| \|s_i^{(k,j)}\|.$$

The stopping criterion (4.4) means that

$$(5.4) \quad 0 \leq \Phi^{(k)}(x^{(k)} + s_i^{(k,j)}) - \Phi^{(k)}(x^{(k)}).$$

At the same time we have

$$(5.5) \quad \Phi^{(k)}(x^{(k)} + s_i^{(k,j)}) - \Phi^{(k)}(x^{(k)}) = \nabla_x \Phi^{(k)}(\xi)^T s_i^{(k,j)}$$

for some ξ in the line segment $(x^{(k)}, x^{(k)} + s_i^{(k,j)})$ connecting $x^{(k)}$ and $x^{(k)} + s_i^{(k,j)}$. Thus from (5.4), (5.5), and (5.3) we obtain

$$\begin{aligned} 0 &\leq \Phi^{(k)}(x^{(k)} + s_i^{(k,j)}) - \Phi^{(k)}(x^{(k)}) \\ &= \nabla_x \Phi^{(k)}(x^{(k)})^T s_i^{(k,j)} + (\nabla_x \Phi^{(k)}(\xi) - \nabla_x \Phi^{(k)}(x^{(k)}))^T s_i^{(k,j)} \\ &\leq -n^{-\frac{1}{2}} \|q^{(k,j)}\| \|s_i^{(k,j)}\| + \|\nabla_x \Phi^{(k)}(\xi) - \nabla_x \Phi^{(k)}(x^{(k)})\| \|s_i^{(k,j)}\|, \end{aligned}$$

which yields

$$(5.6) \quad \|q^{(k,j)}\| \leq n^{\frac{1}{2}} \|\nabla_x \Phi^{(k)}(\xi) - \nabla_x \Phi^{(k)}(x^{(k)})\|.$$

Applying the mean-value theorem again, for some $\zeta \in (x^{(k)}, \xi)$ we have

$$\nabla_x \Phi^{(k)}(\xi) - \nabla_x \Phi^{(k)}(x^{(k)}) = \nabla_{xx}^2 \Phi^{(k)}(\zeta)(\xi - x^{(k)}),$$

so

$$(5.7) \quad \begin{aligned} \|\nabla_x \Phi^{(k)}(\xi) - \nabla_x \Phi^{(k)}(x^{(k)})\| &\leq \|\nabla_{xx}^2 \Phi^{(k)}(\zeta)\| \|\xi - x^{(k)}\| \\ &\leq \|\nabla_{xx}^2 \Phi^{(k)}(\zeta)\| \|s_i^{(k,j)}\|. \end{aligned}$$

Now,

$$\nabla_{xx}^2 \Phi^{(k)}(\zeta) = \nabla_{xx}^2 f(\zeta) + \sum_{l=1}^m \lambda_l^{(k)} \nabla^2 c_l(\zeta) + \frac{1}{\mu^{(k)}} (\nabla c(\zeta) S \nabla c(\zeta)^T + \sum_{l=1}^m s_u c_l(\zeta) \nabla^2 c_l(\zeta)).$$

By construction, $\omega^{(k)} \rightarrow 0$, so $\delta^{(k)} \rightarrow 0$, so by (AS2), ζ lies in a compact subset that is independent of k . Furthermore, the bound $S^{(k)} \leq S_2$ is independent of k . Thus we can find M , independent of k , such that

$$\|\nabla_{xx}^2 \Phi^{(k)}(\zeta)\| \leq M + M \|\lambda^{(k)}\| + M \frac{1}{\mu^{(k)}}.$$

Returning to (5.7) we have

$$(5.8) \quad \|\nabla_x \Phi^{(k)}(\xi) - \nabla_x \Phi^{(k)}(x^{(k)})\| \leq M \left(1 + \|\lambda^{(k)}\| + \frac{1}{\mu^{(k)}}\right) \|s_i^{(k,j)}\|.$$

Thus from (5.6), (5.8), the fact that $s_i^{(k,j)} \in \Delta^{(k,j)}\Gamma^{(k,j)}$, and (4.3) we have

$$\begin{aligned} \|q^{(k,j)}\| &\leq n^{\frac{1}{2}} \|\nabla_x \Phi^{(k)}(\xi) - \nabla_x \Phi^{(k)}(x^{(k)})\| \\ &\leq n^{\frac{1}{2}} M \left(1 + \|\lambda^{(k)}\| + \frac{1}{\mu^{(k)}}\right) \|s_i^{(k,j)}\| \\ &\leq n^{\frac{1}{2}} d^* M \left(1 + \|\lambda^{(k)}\| + \frac{1}{\mu^{(k)}}\right) \Delta^{(k,j)} \\ &\leq n^{\frac{1}{2}} d^* M \left(1 + \|\lambda^{(k)}\| + \frac{1}{\mu^{(k)}}\right) \delta^{(k)}. \end{aligned}$$

Finally, the rule for updating $\delta^{(k)}$ in either Step 2 or Step 3 gives us

$$(5.9) \quad \|q^{(k,j)}\| \leq n^{\frac{1}{2}} d^* M \omega^{(k)}.$$

Combining (5.2) and (5.9) yields the proposition. \square

5.2. Convergence results. Proposition 5.1 means that the asymptotic behavior of $\|P(x^{(k)}, \nabla_x \Phi^{(k)})\|$ in the augmented Lagrangian pattern search algorithm is like that of the same quantity in the original algorithm. This, in turn, allows us to piggy-back the convergence analysis for the augmented Lagrangian pattern search algorithm on that for the original augmented Lagrangian algorithm in [6]. Because of Proposition 5.1 the original proofs of all these results still hold.

The first convergence result corresponds to Theorem 4.4 and Lemma 4.3 in [6]. Let

$$g_L(x; \lambda) = \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla c_i(x),$$

which is the gradient of the Lagrangian with respect to the constraints $c_i(x)$ only.

THEOREM 5.2. *Assume that (AS1) holds. Let x^* be any limit point of the sequence $\{x^{(k)}\}$ generated by the augmented Lagrangian pattern search algorithm for which (AS2) and (AS3) hold and let K be the set of indices of an infinite subsequence of the $x^{(k)}$ whose limit is x^* . Then*

(i) $c(x^*) = 0$.

(ii) x^* is a Karush-Kuhn-Tucker point (first-order stationary point) for the problem (1.1), λ^* is the corresponding vector of Lagrange multipliers, and the sequence $\{\bar{\lambda}(x^{(k)}), \lambda^{(k)}, S^{(k)}, \mu^{(k)}\}$ converges to λ^* for $k \in K$.

(iii) There are positive constants a_1, a_2, s_1 and an integer k_0 such that

$$\|\bar{\lambda}(x^{(k)}), \lambda^{(k)}, S^{(k)}, \mu^{(k)} - \lambda^*\| \leq a_1 \omega^{(k)} + a_2 \|x^{(k)} - x^*\|$$

and

$$\|c(x^{(k)})\| \leq s_1 (a_1 \omega^{(k)} \mu^{(k)} + \mu^{(k)} \|\lambda^{(k)} - \lambda^*\| + a_2 \mu^{(k)} \|x^{(k)} - x^*\|)$$

for all $k \geq k_0$, ($k \in K$).

(iv) The gradients $\nabla_x \Phi^{(k)}$ converge to $g_L(x^*; \lambda^*)$ for $k \in K$.

As in [6], under additional assumptions we obtain stronger results. Following [6], if J_1 and J_2 are any index sets, and $H_L(x^*, \lambda^*)$ is the Hessian of the Lagrangian, then $H_L(x^*, \lambda^*)_{[J_1, J_2]}$ is the matrix formed by taking the rows and columns of $H_L(x^*, \lambda^*)$ indexed by J_1 and J_2 , respectively, while $A(x^*)_{[J_1]}$ is the matrix formed by taking the columns of $A(x^*)$ indexed by J_1 . We then make the following assumptions.

AS4. *The second derivatives of the functions $f(x)$ and the $c_i(x)$ are Lipschitz continuous at all points within Ω .*

AS5. *Suppose that (x^*, λ^*) is a Karush-Kuhn-Tucker point for the problem (1.1) and that*

$$\begin{aligned} J_1 &= \{ i \mid (g_L(x^*; \lambda^*))_i = 0 \text{ and } \ell_i < x_i^* < u_i \} \\ J_2 &= \{ i \mid (g_L(x^*; \lambda^*))_i = 0 \text{ and } (x_i^* = \ell_i \text{ or } x_i^* = u_i) \}. \end{aligned}$$

Then we assume that the matrix

$$\begin{bmatrix} H_L(x^*, \lambda^*)_{[J, J]} & (A(x^*)_{[J]})^T \\ A(x^*)_{[J]} & 0 \end{bmatrix}$$

is nonsingular for all sets J , where J is any set made up from the union of J_1 and any subset of J_2 .

The next result from [6], which also holds for the augmented Lagrangian pattern search algorithm, is Lemma 5.1. This result relates the convergence of the iterates to the error in the multipliers, a relationship characteristic of augmented Lagrangian methods [1, 32]. Again, the proof in [6] holds for the pattern search variant because of Proposition 5.1.

LEMMA 5.3. *Suppose that (AS1) holds. Let $\{x^{(k)}\} \subset B$, $k \in K$, be a subsequence which converges to the Karush-Kuhn-Tucker point x^* for which (AS2), (AS4), and (AS5) hold, and let λ^* be the corresponding vector of Lagrange multipliers. Assume that $\{\lambda^{(k)}\}$, $k \in K$, is any sequence of vectors, that $\{S^{(k)}\}$, $k \in K$, is any sequence of diagonal matrices satisfying $0 < S_1^{-1} \leq S^{(k)} \leq S_2 < \infty$, and that $\{\mu^{(k)}\}$, $k \in K$, form a nonincreasing sequence of positive scalars, so that the product $\mu^{(k)} \|\lambda^{(k)} - \lambda^*\|$ converges to zero as k increases. Now, suppose further that*

$$\|P(x^{(k)}, \nabla_x \Phi^{(k)})\| \leq \omega^{(k)},$$

where the $\omega^{(k)}$ are positive scalar parameters which converge to zero as $k \in K$ increases. Then there are positive constants $\bar{\mu}$, a_3 , a_4 , a_5 , a_6 , and s_1 and an integer value k_0 so that if $\mu^{(k_0)} \leq \bar{\mu}$ then

$$(5.10) \quad \|x^{(k)} - x^*\| \leq a_3 \omega^{(k)} + a_4 \mu^{(k)} \|\lambda^{(k)} - \lambda^*\|$$

$$\|\bar{\lambda}(x^{(k)}, \lambda^{(k)}, S^{(k)}, \mu^{(k)}) - \lambda^*\| \leq a_5 \omega^{(k)} + a_6 \mu^{(k)} \|\lambda^{(k)} - \lambda^*\|$$

and

$$(5.11) \quad \|c(x^{(k)})\| \leq s_1(a_5 \omega^{(k)} \mu^{(k)} + (\mu^{(k)} + a_6(\mu^{(k)})^2) \|\lambda^{(k)} - \lambda^*\|)$$

for all $k \geq k_0$, ($k \in K$).

The following is Corollary 5.2 in [6].

COROLLARY 5.4. *Suppose that the conditions of Lemma 5.3 hold and that $\hat{\lambda}^{(k+1)}$ is any Lagrange multiplier estimate for which*

$$\|\hat{\lambda}^{(k+1)} - \lambda^*\| \leq a_{16} \|x^{(k)} - x^*\| + a_{17} \omega^{(k)}$$

for some positive constants a_{16} and a_{17} and all $k \in K$ sufficiently large. Then there are positive constants $\bar{\mu}$, a_3 , a_4 , a_5 , a_6 , s_1 and an integer value k_0 so that if $\mu^{(k_0)} \leq \bar{\mu}$ then (5.10),

$$\|\hat{\lambda}^{(k+1)} - \lambda^*\| \leq a_5 \omega^{(k)} + a_6 \mu^{(k)} \|\lambda^{(k)} - \lambda^*\|,$$

and (5.11) hold for all $k \geq k_0$, ($k \in K$).

We also inherit the following result indicating that we may generally expect the penalty parameter to remain bounded away from zero. This is Theorem 5.3 in [6]. Taken together with the convergence of the multiplier estimates, this means that the stopping tolerance for the inexact minimization of the augmented Lagrangian is decreasing at the same rate as in the original algorithm. However, in §6 of [6] the authors show that in the case of non-unique limit points one can have $\mu^{(k)} \rightarrow 0$, in which case the stopping tolerance δ^k decreases more like $(\mu^{(k)})^2$.

THEOREM 5.5. *Suppose that the iterates $\{x^{(k)}\}$ of the augmented Lagrangian pattern search algorithm converge to the single limit point x^* , that (AS1), (AS2), (AS4), and (AS5) hold, and that α_η and β_η satisfy $\alpha_\eta < \min(1, \alpha_\omega)$ and $\beta_\eta < \min(1, \beta_\omega)$. Then there is a constant $\mu > 0$ such that $\mu^{(k)} > \mu$ for all k .*

The proof of Theorem 5.5 makes use of the fact that $\| P(x^{(k)}, \nabla_x \Phi^{(k)}) \| = O(\omega^k)$, whereas the proofs of the preceding convergence results require only that

$$\| P(x^{(k)}, \nabla_x \Phi^{(k)}) \| \rightarrow 0.$$

Finally, we have the following result on the rate of convergence of the outer iteration, corresponding to Theorem 5.5 in [6].

THEOREM 5.6. *Under the assumptions of Theorem 5.5, the iterates $x^{(k)}$ and the Lagrange multiplier estimates $\bar{\lambda}^{(k)}$ of the augmented Lagrangian pattern search algorithm are at least R -linearly convergent with R -factor at most $\hat{\mu}^{\min(\beta_\omega, \beta_\eta)}$, where $\hat{\mu} = \min[\gamma_1, \mu]$ and where μ is the smallest value of the penalty parameter generated by the algorithm in question.*

6. Application to inequality constrained minimization. Special consideration is due to the general problem

$$(6.1) \quad \begin{aligned} & \text{minimize} && f(y) \\ & \text{subject to} && g(y) \leq 0 \\ & && \ell \leq y \leq u, \end{aligned}$$

converted into the form (1.1) via the introduction of non-negative slack variables:

$$(6.2) \quad \begin{aligned} & \text{minimize} && f(y) \\ & \text{subject to} && g(y) + z = 0 \\ & && \ell \leq y \leq u \\ & && z \geq 0. \end{aligned}$$

The augmented Lagrangian associated with (6.2) is

$$(6.3) \quad \Phi(y, z; \lambda, S, \mu) = f(y) + \lambda^T (g(y) + z) + \frac{1}{2\mu} \sum_{i=1}^m s_{ii} (g_i(y) + z_i)^2.$$

Explicit equality constraints may also be present in (6.1); we ignore them here for brevity.

The introduction of slacks increases the dimension of the bound constrained subproblem that we must solve at each outer iteration. Such increases in dimension usually cause a degradation in performance for pattern search methods. However, we can avoid this increase in dimension because of the simple way in which the slacks z enter into (6.3). A standard approach [1, 27] is to note that given y , we can minimize $\Phi(y, z; \lambda, S, \mu)$ explicitly in z for $z \geq 0$. Doing so leads to a subproblem problem in y alone:

$$(6.4) \quad \begin{aligned} & \text{minimize} && \Phi(y, z(y); \lambda, S, \mu) \\ & \text{subject to} && \ell \leq y \leq u, \end{aligned}$$

where

$$\Phi(y, z(y); \lambda, S, \mu) = f(y) + \frac{\mu}{2} \sum_{i=1}^m \frac{1}{s_{ii}} (\max(0, \lambda_i + \frac{s_{ii}}{\mu} g_i(y))^2 - \lambda_i^2).$$

The multiplier update formula (2.2) is also modified:

$$\bar{\lambda}_i(x, \lambda, S, \mu) = \max(0, \lambda_i + s_{ii} c_i(x)/\mu), \quad i = 1, \dots, m.$$

See [1] for further discussion. The reduced augmented Lagrangian $\Phi(y, z(y); \lambda, S, \mu)$ has Lipschitz first derivatives. Moreover, if the feasible region for the original problem (6.1) is compact (e.g., if there are upper and lower bounds on all the components of y), then the feasible region for (6.4) is also compact, so we may be assured of convergence of a bound constrained pattern search algorithm applied to (6.4).

REFERENCES

- [1] D. P. BERTSEKAS, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, NY, 1982.
- [2] G. E. P. BOX, *Evolutionary operation: A method for increasing industrial productivity*, *Applied Statistics*, 6 (1957), pp. 81–101.
- [3] M. J. BOX, *A new method of constrained optimization and a comparison with other methods*, *The Computer Journal*, 8 (1965), pp. 42–52.
- [4] M. J. BOX, D. DAVIES, AND W. H. SWANN, *Non-Linear Optimization Techniques*, ICI Monograph No. 5, Oliver & Boyd, Edinburgh, 1969.
- [5] C. W. CARROLL, *The created response surface technique for optimizing nonlinear restrained systems*, *Operations Research*, 9 (1961).
- [6] A. R. CONN, N. I. M. GOULD, AND P. L. TOINT, *A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds*, *SIAM Journal on Numerical Analysis*, 28 (1991), pp. 545–572.
- [7] ———, *LANCELOT: A Fortran package for large-scale nonlinear optimization (release A)*, vol. 17 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1992.
- [8] D. DAVIES AND W. H. SWANN, *Review of constrained optimization*, in *Optimization*, R. Fletcher, ed., Academic Press, 1969, pp. 187–202.
- [9] J. E. DENNIS, JR. AND V. TORCZON, *Direct search methods on parallel machines*, *SIAM Journal on Optimization*, 1 (1991), pp. 448–474.
- [10] L. C. W. DIXON, *ACSIM—An accelerated constrained simplex technique*, *Computer Aided Design*, 5 (1973), pp. 22–32.
- [11] A. V. FIACCO AND G. P. MCCORMICK, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Classics in Applied Mathematics, SIAM, Philadelphia, 1990.
- [12] H. GLASS AND L. COOPER, *Sequential search: A method for solving constrained optimization problems*, *Journal of the Association for Computing Machinery*, 12 (1965), pp. 71–82.
- [13] R. GLOWINSKI AND A. J. KEARSLEY, *On the simulation and control of some friction constrained problems*, *SIAM Journal on Optimization*, 5 (1995), pp. 681–694.
- [14] D. M. HIMMELBLAU, *Applied Nonlinear Programming*, McGraw-Hill, 1972.
- [15] R. HOOKE AND T. A. JEEVES, *Direct search solution of numerical and statistical problems*, *Journal of the Association for Computing Machinery*, 8 (1961), pp. 212–229.
- [16] A. J. KEARSLEY, *The use of optimization techniques in the solution of partial differential equations from science and engineering*, PhD thesis, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 1996; available as Tech. Rep. 96-11, Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005-1892.
- [17] D. L. KEEFER, *Simpat: Self-bounding direct search method for optimization*, *Industrial and Engineering Chemistry Process Design and Development*, 12 (1973), pp. 92–99.
- [18] W. R. KLINGMAN AND D. M. HIMMELBLAU, *Nonlinear programming with the aid of a multiple-gradient summation technique*, *Journal of the Association for Computing Machinery*, 11 (1964), pp. 400–415.
- [19] R. M. LEWIS AND V. J. TORCZON, *Pattern search algorithms for bound constrained minimization*, Tech. Rep. 96-20, Institute for Computer Applications in Science and Engineering, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia 23681-2199, 1996. To appear in *SIAM J. on*

Optimization.

- [20] —, *Rank ordering and positive bases in pattern search algorithms*, Tech. Rep. 96-71, Institute for Computer Applications in Science and Engineering, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia 23681-2199, 1996. In revision for *Mathematical Programming*.
- [21] —, *Pattern search methods for linearly constrained minimization*, Tech. Rep. 98-3, Institute for Computer Applications in Science and Engineering, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia 23681-2199, January 1998. Submitted to *SIAM Journal on Optimization*.
- [22] J. H. MAY, *Linearly Constrained Nonlinear Programming: A Solution Method That Does Not Require Analytic Derivatives*, PhD thesis, Yale University, December 1974.
- [23] R. MIFFLIN, *A superlinearly convergent algorithm for minimization without evaluating derivatives*, *Mathematical Programming*, 9 (1975), pp. 100–117.
- [24] J. A. NELDER AND R. MEAD, *A simplex method for function minimization*, *The Computer Journal*, 7 (1965), pp. 308–313.
- [25] D. PAVIANI AND D. M. HIMMELBLAU, *Constrained nonlinear optimization by heuristic programming*, *Operations Research*, 17 (1969), pp. 872–882.
- [26] E. POLAK, *Computational Methods in Optimization: A Unified Approach*, Academic Press, New York, 1971.
- [27] R. T. ROCKAFELLAR, *A dual approach to solving nonlinear programming problems by unconstrained optimization*, *Mathematical Programming*, 5 (1973), pp. 354–373.
- [28] H. H. ROSENBROCK, *An automatic method for finding the greatest or least value of a function*, *The Computer Journal*, 3 (1960), pp. 175–184.
- [29] W. SPENDLEY, *Nonlinear least squares fitting using a modified simplex minimization method*, in *Optimization*, R. Fletcher, ed., Academic Press, 1969, pp. 259–270.
- [30] W. SPENDLEY, G. R. HEXT, AND F. R. HIMSWORTH, *Sequential application of simplex designs in optimisation and evolutionary operation*, *Technometrics*, 4 (1962), pp. 441–461.
- [31] W. H. SWANN, *Direct search methods*, in *Numerical Methods for Unconstrained Optimization*, W. Murray, ed., Academic Press, London and New York, 1972, pp. 13–28.
- [32] R. A. TAPIA, *Diagonalized multiplier methods and quasi-Newton methods for constrained optimization*, *Journal of Optimization Theory and Applications*, 22 (1977), pp. 135–194.
- [33] V. TORCZON, *Multi-Directional Search: A Direct Search Algorithm for Parallel Machines*, PhD thesis, Department of Mathematical Sciences, Rice University, Houston, Texas, 1989; available as Tech. Rep. 90-07, Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005-1892.
- [34] —, *On the convergence of the multidirectional search algorithm*, *SIAM Journal on Optimization*, 1 (1991), pp. 123–145.
- [35] —, *PDS: Direct search methods for unconstrained optimization on either sequential or parallel machines*, Tech. Rep. TR92-09, Rice University, Department of Computational and Applied Mathematics, Houston, Texas 77005–1892, 1992. In revision.
- [36] —, *On the convergence of pattern search algorithms*, *SIAM Journal on Optimization*, 7 (1997), pp. 1–25.