Mathematical Programming 43 (1989) 277-303
North-Holland

A ROBUST SEQUENTIAL QUADRATIC PROGRAMMING METHOD

J.V. BURKE*

Department of Mathematics, GN-50, University of Washington, Seattle, WA 98195, USA

S-P. HAN**

Department of Mathematics, University of Illinois, Urbana, IL 61801, USA

Received 28 December 1984
Revised manuscript received 25 January 1988

The sequential quadratic programming method developed by Wilson, Han and Powell may fail if the quadratic programming subproblems become infeasible, or if the associated sequence of search directions is unbounded. This paper considers techniques which circumvent these difficulties by modifying the structure of the constraint region in the quadratic programming subproblems. Furthermore, questions concerning the occurrence of an unbounded sequence of multipliers and problem feasibility are also addressed.

Key words: Sequential quadratic programming, nonlinear programming.

1. Introduction

In this paper the sequential quadratic programming (SQP) method, developed by Wilson [29], Han [14, 15], and Powell [20, 21] (also see [1, 11, 23]), for solving nonlinear programming problems is extended to problems in which the quadratic programming sub-problems may be infeasible. Our approach is based on the method developed in [3]. Recall that the SQP method iteratively approximates the nonlinear program

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \quad g(x)\leq 0,\ h(x)=0 \tag{1.1}$$

by quadratic programs (QP) of the form

$$\begin{aligned}\min_{d\in\mathbb{R}^n}\ & \nabla f(x)^{\mathrm T}d+\tfrac{1}{2}d^{\mathrm T}Hd\\ \text{subject to}\ & g(x)+g'(x)d\leq 0, \\ & h(x)+h'(x)d=0,\end{aligned} \tag{1.2}$$

* Work supported in part by the National Science Foundation under Grant No. DMS-8602399 and by the Air Force Office of Scientific Research under Grant No. ISSA-860080.
** Work supported in part by the National Science Foundation under Grant No. DMS-8602419.

where $f:\mathbb{R}^n \to \mathbb{R}$, $g:\mathbb{R}^n \to \mathbb{R}^{m_1}$, and $h:\mathbb{R}^n \to \mathbb{R}^{m_2}$ are all continuously differentiable on $\mathbb{R}^n$, and $H \in \mathbb{R}^{n\times n}$ is symmetric positive definite. The iteration then has the form

$$\bar{x} := x + \lambda d$$

where d solves (1.2) and λ is a step length chosen to reduce the value of a merit function for (1.1) (e.g. see [14, 20, 21]).

A serious limitation of the SQP approach is the requisite consistency of the QP subproblems (1.2). The goal of this paper is to describe a modification to this method wherein the QP subproblem (1.2) is altered in a way which guarantees that the associated constraint region is nonempty for each $x \in \mathbb{R}^n$ and for which a reasonably robust convergence theory can be established.

Within the framework of the SQP method, Powell [21] suggests an approach that is similar to the one that we will describe. The modification that Powell considers is designed to guarantee the nonemptiness of the constraint region associated with the QP subproblem at each $x \in \mathbb{R}^n$. This modification also produces search directions that are descent directions for the associated l_1 merit function. In [25, 26], Schittkowski provides an in-depth investigation into the computational behavior of these methods. His work indicates that in practice these methods work quite well. However, there does not yet exist a satisfactory convergence theory for these algorithms.

Another, quite different, QP-based approach to solving Problem (1.1) has been developed by Fletcher [8, 9, 10]. It is known as the QL method or the Sl$_1$QP method. In Fletcher's approach one employs a trust region strategy to locate stationary points of the l_1 merit function that are feasible to (1.1). Fletcher has shown that, under certain local hypotheses, his method and the SQP method generate identical iterates [9].

Our method is similar to the methods of Powell [21] and Fletcher [8, 9, 10] in that it can overcome some difficulties associated with the infeasibility of the QP subproblems (1.2). A feature which is uncommon to the other methods is that even when Problem (1.1) is itself infeasible our method can sometimes extract useful information about the problem from our calculated points. This is because the intent of the method is to search for stationary points of (1.1) in a broader sense. A point will be called a stationary point for (1.1) if it is either a Kuhn–Tucker point, a Fritz John point (see the discussion after Corollary 5.1), or is a stationary point of the distance function

$$\phi(x) := \operatorname{dist}\left[\begin{pmatrix} g(x) \\ h(x) \end{pmatrix} \middle| \mathbb{R}^{m_1}_{-} \times \{0\}_{\mathbb{R}^{m_2}}\right]$$

which is not feasible for (1.1) (the distance function ϕ is defined in Section 2). The motivation for our approach is based on the point of view that problem (1.1) is composed of two problems. The first is the feasibility problem and the second is the problem of minimizing f. In this paper, the goal of primary importance is the attainment of feasibility. The consequences of this perspective are developed in the next two sections wherein we discuss the modified SQP subproblems and a rule for updating the penalty parameter of the merit function used in choosing a step length. In Section 4 the proposed model algorithm is stated. In Section 5 we study properties of the updating rule for the penalty parameter in order to ascertain the causes of its potential unboundedness. The convergence theory for the method is presented in Section 6. Section 7 contains a brief discussion on questions concerning implementation. The paper is concluded in Section 8 by studying two examples. The examples are included in order to demonstrate situations in which the algorithm proposed in this paper succeeds while either the QL algorithm or the methods proposed by Powell [21] can fail. Moreover, the examples are interesting in their own right since they are constructed to illustrate dramatically certain deficiencies in the techniques described in [8, 21].

The notation that we employ is standard. However, a partial list of definitions is provided for the reader's convenience.

- $f'(x; d) := \lim_{\lambda \downarrow 0}(f(x+\lambda d)-f(x))/\lambda$.
- $g'(x)$ is the Fréchet derivative of g at x.
- For $g:\mathbb{R}^n \to \mathbb{R}^m$ we define $\mathrm{Ran}(g) \subset \mathbb{R}^n$ to be the range of g.
- For $C \subset \mathbb{R}^m$, $\overline{\mathrm{co}}\, C$ is closed convex hull of C, $\mathrm{int}\, C$ is the interior of C, and $N(x \mid C) := \{x^*: \langle x^*, y-x\rangle \leq 0$ for *all* $y \in C\}$ is the normal cone to C at a point $x \in C$.
- Let $\|\cdot\|$ denote the norm on $\mathbb{R}^n$; then B denotes the associated closed unit ball, $\|\cdot\|_0$ denotes the associated dual norm, i.e. $\|x\|_0 := \sup\{\langle x, y\rangle : \|y\| \leq 1\}$, and B^0 denotes the closed unit ball of the dual norm.
- $K := \mathbb{R}^{m_1}_- \times \{0\}_{\mathbb{R}^{m_2}}$ and $K^0 := \mathbb{R}^{m_1}_+ \times \mathbb{R}^{m_2}$.
- If $C \subset \mathbb{R}^n$, then

$$\mathrm{dist}(x \mid C) := \inf\{\|x-y\| : y \in C\},$$

and

$$\mathrm{dist}_0(x \mid C) := \inf\{\|x-y\|_0 : y \in C\}.$$

- For $f:\mathbb{R}^n \to \mathbb{R}$ and $C \subset \mathbb{R}^n$, set

$$\mathrm{argmin}\{f(x): x \in C\} := \{\bar{x} \in C: f(\bar{x}) = \inf\{f(x): x \in C\}\}.$$

- For $f:\mathbb{R}^n \to \mathbb{R}$ locally Lipschitzian, let ∂f denote its Clarke subdifferential.
- The vector $e \in \mathbb{R}^n$ is the vector of ones, $e := (1, 1, \ldots, 1)^{\mathrm{T}}$.

2. The modified SQP subproblems

The modification to (1.2) that we consider is based upon the point of view that the attainment of feasibility is the objective of primary importance, whereas the minimization of f is secondary. With this perspective in mind, we associate with every point in $\mathbb{R}^n$ a set of search directions that reflects the desire to attain feasibility. Specifically, for each $x \in \mathbb{R}^n$ we associate a set of directions, $D(x, \sigma, \beta)$, where $D(x, \sigma, \beta) \subset \beta B$ and each $d \in D(x, \sigma, \beta)$ is a descent direction for the distance function

$$\phi(x) := \text{dist}\left[\begin{pmatrix} g(x) \\ h(x) \end{pmatrix} \middle| K\right] \tag{2.1}$$

with

$$K := \mathbb{R}^{m_1}_- \times \{0\}_{\mathbb{R}^{m_2}}.$$

Here the mapping $\text{dist}(\binom{y}{z}|K)$ is defined by

$$\text{dist}\left(\begin{pmatrix} y \\ z \end{pmatrix} \middle| K\right) := \inf\left\{\left\|\begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} x \\ 0 \end{pmatrix}\right\| : x_i \leq 0,\ i = 1, \ldots, m_1\right\}$$

with $\|\cdot\|$ being any given norm on $\mathbb{R}^{m_1+m_2}$. The parameter σ is required in the designation of the set $D(x, \sigma, \beta)$ in order to assure the existence of Kuhn–Tucker multipliers for the modified subproblem (see Lemma 2.2). The parameters σ and β have the relation $0 < \sigma < \beta$. In Section 7 we derive the following practical formulas for $\phi(x)$ in the case where the norm on $\mathbb{R}^{m_1+m_2}$ is chosen to be the l^1, l^2 or l^∞ norm:

$$l^1\text{-norm:} \quad \phi_1(x) = \sum_{i=1}^{m_1} g_i(x)_+ + \sum_{j=1}^{m_2} |h_j(x)|,$$

$$l^2\text{-norm:} \quad \phi_2(x) = \left[\sum_{i=1}^{m_1} (g_i(x)_+)^2 + \sum_{j=1}^{m_2} (h_j(x))^2\right]^{1/2},$$

$$l^\infty\text{-norm:} \quad \phi_\infty(x) = \max\{g_i(x)_+, |h_j(x)|: i = 1, \ldots, m_1, j = 1, \ldots, m_2\},$$

where $\zeta_+ := \max\{0, \zeta\}$ for all $\zeta \in \mathbb{R}$. The mapping ϕ is an example of a convex composite function [2] since the distance function, $\text{dist}(y\,|\,K)$, is convex [4, Theorem 3.1]. Moreover, due to its nature, the distance function for a cone such as K has a very rich geometric structure. This structure was explored in [4], to which we make frequent reference.

Let us now turn to the construction of the sets $D(x, \sigma, \beta)$. We begin by associating with each $x \in \mathbb{R}^n$ and $\sigma > 0$ a set of residuals $R(x, \sigma) \subset \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, where $\binom{z_1}{z_2} \in R(x, \sigma)$ if and only if there exists $\bar{d} \in \mathbb{R}^n$ and $z \in \mathbb{R}^{m_1}_-$ such that

$$z_1 = g(x) + g'(x)\bar{d} - z,$$

$$z_2 = h(x) + h'(x)\bar{d},$$

$$\|\bar{d}\| \leq \sigma,$$

and $\binom{z_1}{z_2}$ satisfies the equation

$$\left\|\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right\| = \inf\left\{\text{dist}\left[\begin{pmatrix} g(x) + g'(x)d \\ h(x) + h'(x)d \end{pmatrix} \middle| K\right] \middle| \|d\| \leq \sigma\right\}.$$

Note that

$$\bar{d} \in \operatorname{argmin}\left\{ \operatorname{dist}\left[\begin{pmatrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{pmatrix} \,\middle|\, K \right] \,\middle|\, \|d\| \leq \sigma \right\}.$$

By definition, each element $\binom{z_1}{z_2}$ in $R(x, \sigma)$ is an optimal residual for the problem

$$\min_{d \in \mathbb{R}^n} \operatorname{dist}\left[\begin{pmatrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{pmatrix} \,\middle|\, K \right]$$

$$\text{subject to } \|d\| \leq \sigma,$$

which is itself equivalent to the problem

$$\min_{d \in \mathbb{R}^n, y \in \mathbb{R}^{m_1}} \left\| \begin{pmatrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{pmatrix} - \begin{pmatrix} y \\ 0 \end{pmatrix} \right\| \qquad (2.2)$$

$$\text{subject to } \|d\| \leq \sigma, \; y \leq 0.$$

Note that the introduction of the constraint $\|d\| \leq \sigma$ ensures that the set $R(x, \sigma)$ is always well defined and nonempty. Moreover, $R(x, \sigma)$ always contains exactly one element whenever the norm employed in $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ is strictly convex (cf. Vlasov [28] or [13]). An alternative way of viewing the residual set $R(x, \sigma)$ is that $R(x, \sigma)$ consists of those vectors $\binom{z_1}{z_2}$ in $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ of least norm for which the system

$$g(x)+g'(x)d \leq z_1,$$

$$h(x)+h'(x)d = z_2,$$

and

$$\|d\| \leq \sigma$$

is consistent.

Next, let the mappings $r_1 : \mathbb{R}^n \to \mathbb{R}^{m_1}$ and $r_2 : \mathbb{R}^n \to \mathbb{R}^{m_2}$ be such that

$$\begin{pmatrix} r_1(x) \\ r_2(x) \end{pmatrix} \in R(x, \sigma)$$

for every $x \in \mathbb{R}^n$. That is, $\binom{r_1(x)}{r_2(x)}$ is a selection from the point-to-set map $R(x, \sigma)$. Practically speaking, one can choose the norms on $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ and $\mathbb{R}^n$ so that (2.2) is either a linear or quadratic program. If one were to choose a specific algorithm to solve this program, then, given $x \in \mathbb{R}^n$ and $\sigma > 0$, a residual $\binom{r_1(x)}{r_2(x)} \in R(x, \sigma)$ would be identified at the solution. In this way, the choice of algorithm to solve (2.2) can determine a selection $\binom{r_1(x)}{r_2(x)}$. Finally, we define $D(x, \sigma, \beta)$ to be the set

$$D(x, \sigma, \beta) := \{ d \in \mathbb{R}^n \mid \|d\| \leq \beta, \, g(x)+g'(x)d \leq r_1(x), \, h(x)+h'(x)d = r_2(x) \}$$

where $\beta > \sigma$. Clearly $D(x, \sigma, \beta)$ is nonempty, since, in particular,

$$D(x, \sigma, \beta) \cap \operatorname{argmin}\left\{ \operatorname{dist}\left[\begin{matrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{matrix} \,\middle|\, K \right] \,\middle|\, \|d\| \leq \sigma \right\} \neq \emptyset.$$

In the analysis to follow the parameters σ and β will play a central role. The relationship $\sigma<\beta$ is required in the proof of existence of Kuhn–Tucker multipliers for our modified QP subproblems (Lemma 2.2). In particular, the condition $\sigma<\beta$ insures that the constraint $\|d\|\leq\beta$ satisfies the Slater constraint qualification [18, page 78].

We now provide two examples illustrating these concepts. The first of these examples demonstrates the need for the parameter σ.

Examples. (1) Let $g:\mathbb{R}\to\mathbb{R}^2$ be given by

$$g(x):=\begin{pmatrix} x^2+1 \\ x \end{pmatrix}$$

for all $x\in\mathbb{R}$, and $m_2=0$. Let $\mathbb{R}^2$ be equipped with the 2-norm. In order to illustrate the purpose of the parameters σ and β, we first consider the case in which the norm constraint in the definition of both $R(x,\sigma)$ and $D(x,\sigma,\beta)$ is absent. In this case, we have

$$\inf\{\operatorname{dist}(g(x)+g'(x)d\,|\,K)\,|\,d\in\mathbb{R}\}=\inf\{\|(g(x)+g'(x)d)_+\|:\ d\in\mathbb{R}\}$$
$$=\begin{cases}(4x^2+1)^{-1}\left\|\begin{matrix}1-x^2\\-2x(1-x^2)\end{matrix}\right\|, & \text{if } -1\leq x<0,\\ 0, & \text{otherwise},\end{cases}$$

with

$$R(x,\infty):=\begin{cases}\left\{(4x^2+1)^{-1}\begin{pmatrix}1-x^2\\-2x(1-x^2)\end{pmatrix}\right\}, & \text{if } -1\leq x\leq 0,\\ \left\{\begin{pmatrix}0\\0\end{pmatrix}\right\}, & \text{otherwise},\end{cases}$$

and

$$D(x,\infty,\infty)=\begin{cases}\left[-\dfrac{(x^2+1)}{2x},-x\right], & \text{if } x\leq -1,\\ \left\{-\dfrac{(2x^3+3x)}{4x^2+1}\right\}, & \text{if } -1\leq x<0,\\ (-\infty,0], & \text{if } x=0,\\ \left(-\infty,-\dfrac{x^2+1}{2x}\right], & \text{if } 0<x\leq 1,\\ (-\infty,-x], & \text{if } 1<x.\end{cases}$$

Note that the feasible region for (1.1) with this choice of g is empty. However, the set of search directions $D(x,\infty,\infty)$ always points towards the origin. This makes sense since the image of the origin under g is the closest point in the range of g to K with respect to the 2-norm.

The most significant feature of the example is the local behavior of $D(x, \infty, \infty)$ near $x = 0$. Observe that

$$\lim_{x \downarrow 0} \operatorname{dist}(0 | D(x, \infty, \infty)) = +\infty.$$

This behavior is unacceptable within the framework of the convergence theory provided in Section 6. In particular, it necessitates the placement of explicit bounds on the choice of search direction. In our setting this is done by introducing the parameters σ and β. If we now set $\sigma = 1$ and take $\beta \geq \sigma$ we find that

$$D(x, \sigma, \beta) = \begin{cases} [1, \min\{-x, \beta\}] & \text{if } x \leq -1, \\ \left\{-\dfrac{2x^3+3x}{4x^2+1}\right\} & \text{if } -1 \leq x < 0, \\ [-\beta, 0] & \text{if } x = 0, \\ [-\beta, -1] & \text{if } 0 < x. \end{cases}$$

Observe that, in general, $D(x, \sigma, \beta_1) \supset D(x, \sigma, \beta_2)$ whenever $\beta_1 > \beta_2$. However, as occurs in this example when $-1 \leq x < 0$, it is possible that $D(x, \infty, \infty) = D(x, \sigma, \beta) = D(x, \sigma, \sigma)$.

(2) Let $h : \mathbb{R} \to \mathbb{R}^2$ be the mapping

$$h(x) := \begin{pmatrix} 1 - e^x \\ x \end{pmatrix}$$

and $m_1 = 0$. Let $\mathbb{R}^2$ be equipped with the 2-norm. Again, as in example 1, it is instructive to begin with the case in which the bound constraint is absent. In this case we have

$$\begin{aligned} \inf\{\operatorname{dist}(h(x) + h'(x)d | K) : d \in \mathbb{R}\} &= \inf\{\|(h(x) + h'(x)d)\| : d \in \mathbb{R}\} \\ &= \inf\left\{\left\|\begin{pmatrix} 1 - (1+d)\, e^x \\ x + d \end{pmatrix}\right\| : d \in \mathbb{R}\right\} \\ &= \frac{1 + (x-1)\, e^x}{1 + e^{2x}} \left\|\begin{pmatrix} 1 \\ e^x \end{pmatrix}\right\| \end{aligned}$$

with

$$R(x, \infty) = \left\{\frac{1 + (x-1)\, e^x}{1 + e^{2x}} \begin{pmatrix} 1 \\ e^x \end{pmatrix}\right\}$$

and

$$D(x, \infty, \infty) = \{(1 + e^{2x})^{-1}(e^x - e^{2x} - x)\}. \tag{2.3}$$

Here the feasible region for (1.1) is the singleton $\{0\}$. Furthermore, the constraints in (1.2) are inconsistent for every $x \in \mathbb{R} \backslash \{0\}$, and the mapping h does not satisfy the

Mangasarian–Fromowitz constraint qualification [18, Definition 5.1] at any point in $\mathbb{R}$. Now, given σ and β satisfying $0<\sigma\leq\beta$, one can employ (2.3) to show that

$$D(x,\sigma,\beta)=\begin{cases}\left\{\min\left\{\sigma,\dfrac{e^x-e^{2x}-x}{1+e^{2x}}\right\}\right\}, & \text{if } x\leq 0,\\ \left\{\max\left\{-\sigma,\dfrac{e^x-e^{2x}-x}{1+e^{2x}}\right\}\right\}, & \text{if } 0<x.\end{cases}$$

This example is particularly pertinent to the discussion since all of the modified subproblems considered by Powell [21] and Schittkowski [25, 26] return a search direction $d=0$ at every point $x\in\mathbb{R}$, regardless of the choice of C^1 objective function f. On the other hand, the method described in this paper converges to the solution $x=0$ from any starting point. □

We now show that the directions in $D(x,\sigma,\beta)$ are indeed descent directions for ϕ.

Lemma 2.1. *Let $0<\sigma\leq\beta$ and choose $\bar{d}\in D(x,\sigma,\beta)$. Then*

$$\phi'(x;\bar{d})\leq\Delta(x,\sigma)\leq 0,$$

where

$$\Delta(x,\sigma):=\inf\left\{\operatorname{dist}\left[\begin{pmatrix}g(x)+g'(x)d\\ h(x)+h'(x)d\end{pmatrix}\middle| K\right]\middle|\ \|d\|\leq\sigma\right\}-\phi(x).$$

Moreover, $\Delta(x,\sigma)=0$ if and only if x is a stationary point of ϕ in the sense that $0\in\partial\phi(x)$, where $\partial\phi(x)$ is the Clarke subdifferential of ϕ at x.

Proof. Let $\bar{d}\in D(x,\sigma,\beta)$. Then

$$\begin{aligned}\operatorname{dist}\left[\begin{pmatrix}g(x)+g'(x)\bar{d}\\ h(x)+h'(x)\bar{d}\end{pmatrix}\middle| K\right]&\leq\operatorname{dist}\left[\begin{pmatrix}r_1(x)\\ r_2(x)\end{pmatrix}\middle| K\right]\\ &\quad+\operatorname{dist}\left[\begin{pmatrix}g(x)+g'(x)\bar{d}-r_1(x)\\ h(x)+h'(x)\bar{d}-r_2(x)\end{pmatrix}\middle| K\right]\\ &\leq\left\|\begin{pmatrix}r_1(x)\\ r_2(x)\end{pmatrix}\right\|\\ &=\inf\left\{\operatorname{dist}\left[\begin{pmatrix}g(x)+g'(x)d\\ h(x)+h'(x)d\end{pmatrix}\middle| K\right]\middle|\ \|d\|\leq\sigma\right\}\end{aligned}\tag{2.4}$$

where the first inequality is a consequence of the sublinearity of the distance function $\operatorname{dist}(y|K)$ (cf. [4, Theorem 3.1]); a mapping $l:\mathbb{R}^m\to\mathbb{R}$ is sublinear if and only if $l(x+y)\leq l(x)+l(y)$ and $l(\alpha x)=\alpha l(x)$ for all $x,y\in\mathbb{R}^m$ and $\alpha\geq 0$. Also, by [4, Lemma 4.1],

$$\phi'(x;d)\leq\operatorname{dist}\left[\begin{pmatrix}g(x)+g'(x)d\\ h(x)+h'(x)d\end{pmatrix}\middle| K\right]-\phi(x)\tag{2.5}$$

for all $d \in \mathbb{R}^n$. By combining (2.4) and (2.5) we have

$$\phi'(x; d) \leq \Delta(x, \sigma) \leq 0.$$

If $\Delta(x, \sigma) = 0$ then $\bar{d} = 0$ is an element of the set

$$\operatorname{argmin}\left\{ \operatorname{dist}\left[\begin{pmatrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{pmatrix} \middle| K \right] : \|d\| \leq \sigma \right\}.$$

Consequently, the constraint $\|d\| \leq \sigma$ is superfluous. Therefore, the last satement of the lemma follows from [4, Theorem 4.2]. □

Having defined the sets $D(x, \sigma, \beta)$, we now describe our proposed modification to the subproblem (1.2). The subproblem (1.2) is simply replaced by the convex program

$$\begin{aligned} Q(x, H, \sigma, \beta): \quad & \min_{d \in \mathbb{R}^n} \nabla f(x)^{\mathrm{T}} d + \tfrac{1}{2} d^{\mathrm{T}} H d \\ & \text{subject to } \; g(x)+g'(x)d \leq r_1(x), \\ & \qquad\qquad h(x)+h'(x)d = r_2(x), \\ & \qquad\qquad \|d\| \leq \beta. \end{aligned}$$

These convex programs have the following properties.

Lemma 2.2. *Let $x \in \mathbb{R}^n$, $0 < \sigma < \beta$, and $H \in \mathbb{R}^{n \times n}$ be symmetric and positive definite. (1) The convex program $Q(x, H, \sigma, \beta)$ has a unique solution d where d satisfies the following Kuhn-Tucker conditions: There exist vectors $u \in \mathbb{R}^{m_1}$ and $v \in \mathbb{R}^{m_2}$ such that*

(a) $g(x)+g'(x)d \leq r_1(x)$, $h(x)+h'(x)d = r_2(x)$, *and* $\|d\| \leq \beta$,
(b) $u \geq 0$,
(c) $0 \in \nabla f(x) + Hd + g'(x)^{\mathrm{T}} u + h'(x)^{\mathrm{T}} v + N(d|\beta B)$ *where* $N(d|\beta B) = \{z^* \in \mathbb{R}^n : \langle z^*, d\rangle = \beta \|z^*\|_0\}$ *is the normal cone to* βB *at* d, *and* (2.6)
(d) $u^{\mathrm{T}}(g(x)+g'(x)d - r_1(x)) = 0$.

(2) If $d = 0$ is the solution to $Q(x, H, \sigma, \beta)$, then x is a stationary point of ϕ.

(3) If x is such that $g(x) \leq 0$, $h(x) = 0$, and $d = 0$ solves $Q(x, H, \sigma, \beta)$, then x is a Kuhn-Tucker point for (1.1).

Proof. (1) Since H is symmetric and positive definite, this follows from the elementary theory of convex programming (cf. Rockafellar [24, Theorem 28.2]).

(2) Substituting $d = 0$ into (2.6a) we find that $g(x) \leq r_1(x)$ and $h(x) = r_2(x)$. Hence, as in (2.4),

$$\phi(x) \leq \inf\left\{ \operatorname{dist}\left[\begin{pmatrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{pmatrix} \middle| K \right] : d \in \mathbb{R}^n \right\}.$$

Therefore, $\Delta(x, \sigma) = 0$ and the result follows from Lemma (2.1).

(3) The hypotheses imply that $r_1(x) = 0$ and $r_2(x) = 0$. Also, since $d = 0$, $N(d|\beta B) = \{0\}$. Plugging this information into (2.6) yields the result. □

We conclude this section with a result concerning the continuity properties of the mapping $\Delta : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$.

Proposition 2.1. *Let $\Delta : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ be as defined in Lemma 2.1. Then Δ is continuous on $\mathbb{R}^n \times \mathbb{R}_+$.*

Proof. It is sufficient to show that the mapping $\phi : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ defined by

$$\phi(x, \sigma) := \inf\left\{ \operatorname{dist}\left[\begin{pmatrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{pmatrix} \middle| K \right] : \|d\| \le \sigma \right\},$$

is continuous on $\mathbb{R}^n$, since, by [4, Theorem 5.3], ϕ is continuous on $\mathbb{R}^n$. Moreover, since for each $x \in \mathbb{R}^n$ the mapping $\phi(x, \cdot) : \mathbb{R}_+ \to \mathbb{R}$ is convex, finite valued, and continuous at $\sigma = 0$, we need only show that $\phi(\cdot, \sigma) : \mathbb{R}^n \to \mathbb{R}$ is continuous on $\mathbb{R}^n$ for every $\sigma \in \mathbb{R}_+$. To this end, note that for every $x \in \mathbb{R}^n$ there is a $d_x \in \mathbb{R}^n$ such that

$$\phi(x, \sigma) = \operatorname{dist}\left[\begin{pmatrix} g(x)+g'(x)d_x \\ h(x)+h'(x)d_x \end{pmatrix} \middle| K \right]$$

and $\|d_x\| \le \sigma$. Thus, for $x, y \in \mathbb{R}^n$, we have

$$\begin{aligned} \phi(x, \sigma) &\le \operatorname{dist}\left[\begin{pmatrix} g(x)+g'(x)d_y \\ h(x)+h'(x)d_y \end{pmatrix} \middle| K \right] \\ &\le \phi(y, \sigma) + \left\| \begin{pmatrix} g(x) \\ h(x) \end{pmatrix} - \begin{pmatrix} g(y) \\ h(y) \end{pmatrix} \right\| + \left\| \begin{pmatrix} g'(x) \\ h'(x) \end{pmatrix} - \begin{pmatrix} g'(y) \\ h'(y) \end{pmatrix} \right\| \sigma. \end{aligned}$$

Hence, by symmetry,

$$\|\phi(x, \sigma) - \phi(y, \sigma)\| \le \left\| \begin{pmatrix} g(x) \\ h(x) \end{pmatrix} - \begin{pmatrix} g(y) \\ h(y) \end{pmatrix} \right\| + \left\| \begin{pmatrix} g'(x) \\ h'(x) \end{pmatrix} - \begin{pmatrix} g'(y) \\ h'(y) \end{pmatrix} \right\| \sigma$$

for every x and y in $\mathbb{R}^n$, whereby the continuity of $\phi(\cdot, \sigma)$ is established. □

3. Updating the penalty parameter

As in [14], the procedure that will be described generates iterates of the form

$$x_{i+1} := x_i + \lambda_i d_i,$$

where d_i is the solution to $Q(x_i, H_i, \sigma_i, \beta_i)$ for an appropriate choice of $H_i \in \mathbb{R}^{n\times n}$. The step length λ_i is determined by a line search routine applied to the exact penalty function

$$P_\alpha(x) := f(x) + \alpha \operatorname{dist}\left[\begin{pmatrix} g(x) \\ h(x) \end{pmatrix} \middle| K \right]$$

for a suitable choice of penalty parameter $\alpha = \alpha_i$. The procedure for choosing α_i is crucial to the success of the method. In particular, α_i must be chosen so that d_i solving $Q(x_i, H_i, \sigma_{ij}\beta_i)$ is a descent direction for P_{α_i} at x_i.

Lemma 3.1. *Let $d \in \mathbb{R}^n$ be the solution to $Q(x, H, \sigma, \beta)$ for some $x \in \mathbb{R}^n$ and some symmetric and positive definite matrix $H \in \mathbb{R}^{n \times n}$. Then the directional derivative $P'_\alpha(x; d)$ satisfies the inequality*

$$
\begin{aligned}
P'_\alpha(x; d) &\leqslant \nabla f(x)^{\mathrm{T}} d + \alpha \Delta(x, \sigma) \\
&\leqslant -d^{\mathrm{T}} H d + \left\| \begin{pmatrix} u \\ v \end{pmatrix} \right\|_0 \left\| \begin{pmatrix} z \\ h(x) \end{pmatrix} - \begin{pmatrix} r_1(x) \\ r_2(x) \end{pmatrix} \right\| + \alpha \Delta(x, \sigma) \qquad (3.1)
\end{aligned}
$$

where z is any element from the set of residuals

$$
R_0(x) := \left\{ z \;\middle|\; \begin{array}{c} z = g(x) - y \text{ for some } y \leqslant 0 \text{ such that} \\ \left\| \begin{pmatrix} z \\ h(x) \end{pmatrix} \right\| = \operatorname{dist}\left[\begin{pmatrix} g(x) \\ h(x) \end{pmatrix} \middle| K \right] \end{array} \right\}
$$

and $\binom{u}{v}$ is any element from the set of multipliers

$$
M(x, H, \sigma, \beta) := \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \;\middle|\; \begin{array}{c} u \geqslant 0,\ u^{\mathrm{T}}(g(x) + g'(x)d - r_1(x)) = 0 \\ 0 \in \nabla f(x) + Hd + g'(x)^{\mathrm{T}} u + h'(x)^{\mathrm{T}} v + N(d \mid \beta B) \end{array} \right\}.
$$

Remark. In Section 7 it is shown that $g(x)_+ \in R_0(x)$ whenever the norm on $\mathbb{R}^{m_1+m_2}$ is any one of the l_p-norms. In particular, if the l_2-norm is used, then $R_0(x) = \{g(x)_+\}$. However, since we wish to develop the theory without identifying a specific norm structure, the above definition of $R_0(x)$ is required.

Proof. The set $M(x, H, \sigma, \beta)$ is nonempty by part 1 of Lemma 2.2. Moreover, if $d = 0$, the result holds, trivially. Thus, suppose that $d \neq 0$ and let $\binom{u}{v} \in M(x, H, \sigma, \beta)$. First observe that

$$
P'_\alpha(x; d) \leqslant \nabla f(x)^{\mathrm{T}} d + \alpha \Delta(x, \sigma)
$$

by Lemma 2.1. Hence we need only establish the second inequality in (3.1). By part (c) in (2.6), there is a $z \in N(d \mid \beta B)$ such that

$$
\nabla f(x) = -[Hd + g'(x)^{\mathrm{T}} u + h'(x)^{\mathrm{T}} v + z].
$$

This, along with part (d) of (2.6), allows us to rewrite $\nabla f(x)^{\mathrm{T}} d + \alpha \Delta(x, \sigma)$ as

$$
-d^{\mathrm{T}} H d - \beta \|z\|_0 + \begin{pmatrix} u \\ v \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} g(x) - r_1(x) \\ h(x) - r_2(x) \end{pmatrix} + \alpha \Delta(x, \sigma).
$$

Finally, note that if $z \in R_0(x)$, then $z = g(x) - y$ for some $y \leqslant 0$. Hence $u^{\mathrm{T}} g(x) \leqslant u^{\mathrm{T}}(g(x) - y)$. The result now follows from the definition of $\|\cdot\|_0$. $\square$

Our objective in the choice of α is to ensure the validity of the inequality

$$
P'_\alpha(x; d) \leqslant \nabla f(x)^{\mathrm{T}} d + \alpha \Delta(x, \sigma) \leqslant -d^{\mathrm{T}} H d. \qquad (3.2)
$$

As was noted in the proof of the above lemma, the first of these inequalities is valid for any $d \in D(x, \sigma, \beta)$. For the second inequality we first observe that if $\Delta(x, \sigma) = 0$, then

$$0 \in \operatorname{argmin}\left\{ \operatorname{dist}\left[\begin{matrix} g(x)+g'(x)d \\ h(x)+h'(x)d \end{matrix} \middle| K \right] : \|d\| \leq \beta \right\}.$$

Consequently, the selection $\binom{r_1(x)}{r_2(x)}$ can be chosen so that $r_1(x) \in R_0(x)$ and $r_2(x) = h(x)$. If the selection is so defined, then (3.2) follows from (3.1) whenever $\Delta(x, \sigma) = 0$. On the other hand, when $\Delta(x, \sigma) \neq 0$, the second inequality in (3.2) is valid as long as

$$\frac{\nabla f(x)^T d + d^T H d}{-\Delta(x, \sigma)} \leq \alpha. \tag{3.3}$$

Thus, by redefining the selection $\binom{r_1(x)}{r_2(x)}$ if necessary, it is always possible to choose α so that (3.2) holds. Obviously there is a danger in choosing α in this way, since $\Delta(x, \beta_0)$ can be arbitrarily close to zero. Section 5 is devoted to a study of conditions that assure the boundedness of the expression on the left of inequality (3.3).

Remark. It is interesting to note that when $\binom{r_1(x)}{r_2(x)} = 0$, inequality (3.2) is valid whenever $\alpha \geq \operatorname{dist}_0(0 \mid M(x, H, \sigma, \beta))$. If, moreover, the constraint $\|d\| \leq \beta$ is inactive, then $Q(x, H, \sigma, \beta)$ reduces to (1.2) and we have recovered the well known condition given in [14].

4. The model algorithm

Initialization: Choose $x_0 \in \mathbb{R}^n$, $\alpha_0 > 0$, $0 < \sigma_l < \sigma_r < \bar{\beta}$, $\beta_0 \in [\sigma_l, \sigma_r]$, $\sigma \in (\sigma_0, \bar{\beta}]$, $\mathscr{H}$ a compact set of positive definite matrices, $H_0 \in \mathscr{H}$, $\gamma_1 > 0$, $1 > \gamma_2 > 0$, and $0 < \mu_1 \leq \mu_2 < 1$.

Having $(x_i, \alpha_i, H_i, \sigma_i, \beta_i)$, obtain $(x_{i+1}, \alpha_{i+1}, H_{i+1}, \sigma_{i+1}, \beta_{i+1})$ as follows:

(1) If $\Delta(x_i, \sigma_i) \neq 0$ let $\binom{r_1(x_i)}{r_2(x_i)}$ be any element of $R(x_i, \sigma_i)$; otherwise let $r_1(x_i) \in R_0(x_i)$ and $r_2(x_i) = h(x_i)$.

(2) Let d_i be the solution to the convex program $Q(x_i, H_i, \sigma_i, \beta_i)$. If $d_i = 0$, stop; x_i is either a nonfeasible stationary point for ϕ, or x_i is a Kuhn–Tucker point for (1.1) (see Lemma 2.2).

(3) If $\nabla f(x_i)^T d_i + \alpha_i \Delta(x_i, \sigma_i) \leq -d_i^T H_i d_i$, set $\alpha_{i+1} := \alpha_i$; otherwise set

$$\alpha_{i+1} := max\left\{ \frac{\nabla f(x_i)^T d_i + d_i^T H_i d_i}{-\Delta(x_i, \sigma_i)}, 2\alpha_i \right\}.$$

(4) Set $x_{i+1} := x_i + \lambda_i d_i$ where λ_i is any scalar such that

$$P_{\alpha_{i+1}}(x_{i+1}) \leq P_{\alpha_{i+1}}(x_i) + \mu_1 \lambda_i [\nabla f(x_i)^T d_i + \alpha_{i+1} \Delta(x_i, \sigma_i)]$$

and either $\lambda_i \geqslant \gamma_1$ or there is a $\bar{\lambda}_i > 0$ such that $\lambda_i \geqslant \gamma_2 \bar{\lambda}_i > 0$ and

$$P_{\alpha_{i+1}}(x_i + \bar{\lambda}_i d_i) > P_{\alpha_{i+1}}(x_i) + \mu_2 \bar{\lambda}_i [\nabla f(x_i)^{\mathrm{T}} d_i + \alpha_{i+1} \Delta(x_i, \sigma_i)].$$

(5) Choose $H_{i+1} \in \mathscr{H}$, $\sigma_{i+1} \in [\sigma_l, \sigma_r]$, and $\beta_{i+1} \in (\sigma_{i+1}, \bar{\beta}]$.

Remarks. (1) The procedure for choosing the step length in step (4) of the algorithm was introduced by Calamai and Moré in [5]. This procedure can be viewed as a generalization of the Armijo method. Since, by Lemma 3.1 and the observations of the previous section, steps (1), (2) and (3) of the algorithm assure us that

$$\begin{aligned} P'_{\alpha_{i+1}}(x_i; d_i) &\leqslant \nabla f(x_i)^{\mathrm{T}} d_i + \alpha_i \Delta(x_i, \sigma_i) \\ &\leqslant -d_i^{\mathrm{T}} H_i d_i \\ &< 0, \end{aligned}$$

it is not difficult to verify (see [5]) that the criteria for specifying λ_i in step (4) are consistent.

(2) At the end of step (3) of the algorithm, one could choose to rescale the objective function f. That is, replace f by $\alpha_i^{-1} f$ and reset α_i to α_0.

(3) In step (5) one is allowed to adjust the parameters σ_i and β_i iteratively. Therefore it is possible to incorporate a trust region like strategy. However, our proof theory does not allow the radius of these trust regions to either decrease to zero or become unbounded.

5. The boundedness of the penalty parameter

Given the procedure for updating the penalty parameter described in the previous section, it is clear that the sequence of penalty parameters may well become unbounded, especially since the primary goal of the algorithm is to force the term $\Delta(x, \sigma)$ to zero. The unboundedness of the penalty parameters is clearly a great concern, since it may be the cause of serious numerical instability in the choice of step length. Moreover, it could ultimately lead to the breakdown of the convergence theory. Thus, it is essential for us to identify those situations in which the sequence of penalty parameters may become unbounded. For this analysis, we need only consider the quotient

$$\frac{\nabla f(x)^{\mathrm{T}} d + d^{\mathrm{T}} H d}{-\Delta(x, \sigma)} \tag{5.1}$$

where d solves $Q(x, H, \sigma, \beta)$, since if (5.1) remains bounded, then so do the penalty parameters. In the study of boundedness conditions for (5.1) we make use of the inequality

$$\frac{\nabla f(x)^{\mathrm{T}} d + d^{\mathrm{T}} H d}{-\Delta(x, \sigma)} \leqslant \frac{\left\| \begin{pmatrix} z \\ h(x) \end{pmatrix} - \begin{pmatrix} r_1(x) \\ r_2(x) \end{pmatrix} \right\|}{-\Delta(x, \sigma)} \operatorname{dist}_0(0 \mid M(x, H, \sigma, \beta)), \tag{5.2}$$

where $z \in R_0(x)$. Inequality (5.2) follows from (3.1). Our approach is to obtain conditions under which the right-hand side of (5.2) is bounded. Inequality (5.2) then assures us that these conditions induce the boundedness of (5.1). The following proposition is the first step in this development.

Proposition 5.1. *Let $x_0 \in \mathbb{R}^n$ and let $\sigma > 0$. If the quantity*

$$(-\Delta(x,\sigma))^{-1}\left\|\begin{pmatrix} z \\ h(x)\end{pmatrix} - \begin{pmatrix} r_1(x) \\ r_2(x)\end{pmatrix}\right\|, \tag{5.3}$$

where z is any element of $R_0(x)$, is locally unbounded at (x_0, σ), then x_0 is a stationary point for ϕ, i.e. $0 \in \partial\phi(x_0)$.

Proof. Since $\|\binom{z}{h(x)} - \binom{r_1(x)}{r_2(x)}\| \leq 2\|\binom{g(x)}{h(x)}\|$ for all $z \in R_0(x)$, and since g is continuous, the term (5.3) can be locally unbounded at x_0 only if there is a sequence $\{x_i, \sigma_i\}$ with $(x_i, \sigma_i) \to (x_0, \sigma)$ and $\Delta(x_i, \sigma_i) \to 0$. But then $\Delta(x_0, \sigma) = 0$ by Proposition 2.1. Consequently, $0 \in \partial\phi(x)$ by Lemma 2.1. □

The following example illustrates a situation in which (5.3) is locally unbounded.

Example. Let $h: \mathbb{R} \to \mathbb{R}^2$ be defined as g was in example (1) of Section 2, that is

$$h(x) := \begin{pmatrix} x^2+1 \\ x \end{pmatrix}$$

for all $x \in \mathbb{R}$, and let $m_1 = 0$. Next, set $\sigma = 10$, and equip $\mathbb{R}^2$ with the l_∞-norm. Then

$$R_0(x) = \left\{\begin{pmatrix} x^2+1 \\ x \end{pmatrix}\right\}$$

and, for $x \in [-1, 1]$,

$$R(x, 10) = \begin{cases} \left\{\dfrac{1-x^2}{1-2x}\begin{pmatrix} 1 \\ 1\end{pmatrix}\right\}, & \text{if } -1 \leq x < 0, \\ \left\{\begin{pmatrix} 1 \\ \theta\end{pmatrix} : \theta \in [-1, 1]\right\}, & \text{if } x = 0, \\ \left\{\dfrac{1-x^2}{1+2x}\begin{pmatrix} 1 \\ -1\end{pmatrix}\right\}, & \text{if } 0 < x \leq 1. \end{cases}$$

Hence, for $x \in [-1, 1]\backslash\{0\}$,

$$\left(\begin{pmatrix} x^2+1 \\ x\end{pmatrix} - r_2(x)\right) = \begin{cases} \dfrac{1}{2x-1}\left(\begin{pmatrix} 2x^3+2x^2+2x \\ x^2-x+1\end{pmatrix}\right), & \text{if } -1 \leq x < 0, \\ \dfrac{1}{2x+1}\left(\begin{pmatrix} 2x^3+2x^2+2x \\ x^2+x+1\end{pmatrix}\right), & \text{if } 0 < x \leq 1. \end{cases}$$

Therefore,

$$\lim_{x\to 0}\left\|\begin{pmatrix} x^2+1 \\ x \end{pmatrix} - r_2(x)\right\|(\Delta(x, 10))^{-1} = -\infty.$$

Also note that

$$0 \in \partial\phi(0) = \{\theta: \theta \in [-1, 1]\}.$$

Proposition 5.1 sheds some light on one of the factors in the right-hand side of (5.2). Results concerning the local boundedness of the other factor,

$$\text{dist}_0(0 | M(x, H, \sigma, \beta)), \tag{5.4}$$

are not as easily obtained. In this regard though, Gauvin [12] has shown that if σ is taken to be $+\infty$ and $\binom{r_1(x)}{r_2(x)} = 0$, then the set $M(x, H, \sigma, \beta)$ is compact if and only if g satisfies the Mangasarian–Fromowitz constraint qualification at x. Nguyen, Strodiot and Mifflin [19] have extended this result to cover more general types of constraint regions, subject to a natural modification of the Mangasarian–Fromowitz constraint qualification. Using these facts, one can derive results concerning the local boundedness of (5.3) in the case where $\sigma = \infty$ and $\binom{r_1(x)}{r_2(x)} = 0$. However, these results are not sufficiently general for our purposes, since we need to consider the case where $\binom{r_1(x)}{r_2(x)} \neq 0$ and $\sigma < \infty$. Nonetheless, the Mangasarian–Fromowitz constraint qualification is still the key to the analysis. Thus we will need to review its definition.

Definition 5.1 [18]. The Mangasarian–Fromowitz constraint qualification (MFCQ) is said to be satisfied at a point $x \in \mathbb{R}^n$, with respect to the underlying constraint system $g(x) \leqslant 0$, $h(x) = 0$, if

(1) there is a $z \in \mathbb{R}^n$ such that

$$\nabla g_i(x)^{\mathrm{T}} z < 0 \quad \text{for } i \in \{i: g_i(x) \geqslant 0, i = 1, \ldots, m_1\},$$
$$\nabla h_j(x)^{\mathrm{T}} z = 0 \quad \text{for } j = 1, 2, \ldots, m_2,$$

(2) the gradients $\{\nabla h_j(x): j = 1, \ldots, m_2\}$ are linearly independent.

Theorem 5.1. *Suppose that the MFCQ is satisfied at $x_0 \in \mathbb{R}^n$. Let $\sigma_l > 0$ and set $F := \{x: g(x) \leqslant 0, h(x) = 0\}$. Then there is a neighborhood $N(x_0)$ of x_0 such that*

(1) *the MFCQ is satisfied at every point in $N(x_0)$,*

(2) *if $x_0 \notin F$, then $0 \notin \partial\phi(x)$ for every $x \in N(x_0)$,*

(3) *if $x_0 \in F$, then $R(x, \sigma) = \{0\}$ for all $x \in N(x_0)$ and $\sigma \geqslant \sigma_l$, and so*

$$(-\Delta(x, \sigma))^{-1}\left\|\begin{pmatrix} z \\ h(x) \end{pmatrix} - \begin{pmatrix} r_1(x) \\ r_2(x) \end{pmatrix}\right\| = 1$$

for all $z \in R_0(x)$, $x \in N(x_0) \backslash F$, and $\sigma \geqslant \sigma_l$ and

(4) *if $x_0 \in F$, then*

$$\sup\{\text{dist}_0(0 | M(x, H, \sigma, \mathscr{B})): H \in \mathscr{H}, x \in N(x_0), \sigma \in [\sigma_l, \sigma_r], \beta \in (\sigma, \bar{\beta}]\} < +\infty$$

where $\mathcal{H} \subset \mathbb{R}^{n\times n}$ is any compact set of symmetric positive definite matrices and $0<\sigma_l<\sigma_r<\bar{\beta}$.

Proof. (1) This result is established in Robinson [22, Theorem 3].

(2) Let $N(x_0)$ be the neighborhood obtained in (1) with $N(x_0)\cap F=\emptyset$, and let $x\in N(x_0)$. From [4, Theorem 4.2] we know that

$$\partial\phi(x)=\begin{pmatrix} g'(x) \\ h'(x)\end{pmatrix}^{\mathrm{T}} \operatorname{argmax}\left\{ u^{\mathrm{T}}g(x)+v^{\mathrm{T}}h(x) \,\middle|\, \left\|\begin{matrix} u \\ v\end{matrix}\right\|_0 \leqslant 1 \text{ and } u\geqslant 0\right\}$$

$$=\begin{pmatrix} g'(x) \\ h'(x)\end{pmatrix}^{\mathrm{T}} \left\{ \begin{pmatrix} u \\ v\end{pmatrix}\in\mathbb{R}^{m_1}\times\mathbb{R}^{m_2} \,\middle|\, \begin{matrix} \left\|\begin{matrix} u \\ v\end{matrix}\right\|_0 \leqslant 1,\ u\geqslant 0 \\ u^{\mathrm{T}}g(x)+v^{\mathrm{T}}h(x)=\varphi(x)\end{matrix}\right\}.$$

Hence, if $x\notin F$ and $0\in\partial\phi(x)$, then there are vectors $\bar{u}\in\mathbb{R}_+^{m_1}$ and $\bar{v}\in\mathbb{R}^{m_2}$ such that

$$\bar{u}^{\mathrm{T}}g(x)+\bar{v}^{\mathrm{T}}h(x)=\operatorname{dist}\left[\begin{pmatrix} g(x) \\ h(x)\end{pmatrix}\,\middle|\, K\right]\neq 0,$$

$$g'(x)^{\mathrm{T}}\bar{u}+h'(x)^{\mathrm{T}}\bar{v}=0,$$

$$\bar{u}^{\mathrm{T}}g(x)+\bar{v}^{\mathrm{T}}h(x)=\max\left\{u^{\mathrm{T}}g(x)+v^{\mathrm{T}}h(x)\colon \left\|\begin{matrix} u \\ v\end{matrix}\right\|_0\leqslant 1,\ u\geqslant 0\right\}.$$

Hence $\binom{\bar{u}}{\bar{v}}\neq 0$ and $(u)_i=0$ if $g_i(x)<0$ for each $i=1,\ldots,m_1$. But the existence of such a pair $\binom{u}{v}$, in conjunction with the hypothesis that the MFCQ is satisfied at x, contradicts Motzkin's theorem of the alternative [27]. Hence $0\notin\partial\phi(x)$.

(3) It easily follows from (1) and the continuity of both g and g' that there is a neighborhood of x_0 on which the set

$$\{d\colon g(x)+g'(x)d\leqslant 0,\ h(x)+h'(x)d=0,\ \|d\|\leqslant\sigma_l\}$$

is nonempty for each x in this neighborhood (cf. Robinson [22, Theorem 3]). Hence $R(x,\sigma)=\{0\}$ on this neighborhood.

(4) Suppose to the contrary that the supremum is not finite. Then, from (2.6) and part (3) above, there exist sequences $\{x_i\}\subset\mathbb{R}^n$, $\{d_i\}\subset\mathbb{R}^n$, $\{u_i\}\subset\mathbb{R}_+^{m_1}$, $\{v_i\}\subset\mathbb{R}^{m_2}$, $\{\beta_i\}$, $\{z_i\}\subset\mathbb{R}^n$, and $\{H_i\}\subset\mathcal{H}$, such that $x_i\to x_0$, $\|\binom{u_i}{v_i}\|_0\uparrow\infty$, $\langle z_i,d_i\rangle=\beta_i\|z_i\|_0$ for all $i=1,2,\ldots,$ and

(a) $0=\nabla f(x_i)+H_id_i+g'(x_i)^{\mathrm{T}}u_i+h'(x_i)^{\mathrm{T}}v_i+z_i$,

(b) $g(x_i)+g'(x_i)d_i\leqslant 0$, $h(x_i)+h'(x_i)d_i=0$, $\|d_i\|\leqslant\beta_i$, and

(c) $(u_i)_j(g(x_i)+g'(x_i)d_i)_j=0$, for $j=1,\ldots,m_1$,

for all $i=1,2,\ldots.$ Moreover, by compactness, we can assume, with no loss of generality, that there exist $d_0\in\mathbb{R}^n$, $u_0\in\mathbb{R}_+^{m_1}$, $v_0\in\mathbb{R}^{m_2}$, β_0, and $H_0\in\mathcal{H}$ such that $H_i\to H_0$, $d_i\to d_0$, $\beta_i\to\beta_0>0$, and $\binom{u_i}{v_i}/\|\binom{u_i}{v_i}\|_0\to\binom{u_0}{v_0}$. Dividing (a) through by $\|\binom{u_i}{v_i}\|_0$ and taking the limit as $i\to\infty$ we see that $z_i/\|\binom{u_i}{v_i}\|_0$ necessarily converges to some limit, say z_0. Then u_0, v_0, and z_0 necessarily satisfy

(a)′ $0=g'(x_0)^{\mathrm{T}}u_0+h'(x_0)^{\mathrm{T}}v_0+z_0$, and

(b)′ $\langle z_0,d_0\rangle=\beta_0\|z_0\|_0$.

We now show that $z_0 = 0$. From (c) and (b), we have that

$$\text{(c)}' \quad u_0^{\mathrm{T}} g(x_0) + v_0^{\mathrm{T}} h(x_0) = -[u_0^{\mathrm{T}} g'(x_0) d_0 + v_0^{\mathrm{T}} h'(x_0) d_0].$$

Hence, if we multiply (a)′ through by d_0 and use (b)′ and (c)′, we find that

$$u_0^{\mathrm{T}} g(x_0) + v_0^{\mathrm{T}} h(x_0) = \beta_0 \|z_0\|_0.$$

But $h(x_0) = 0$, $g(x_0) \leqslant 0$ and $u_0 \geqslant 0$. Hence $0 \geqslant \beta_0 \|z_0\|_0$, and so $z_0 = 0$.

Now, since $z_0 = 0$, (a)′ implies that

$$\text{(c)}'' \quad g'(x_0)^{\mathrm{T}} u_0 + h'(x_0)^{\mathrm{T}} v_0 = 0.$$

Combining this with (c)′, we find that

$$u_0^{\mathrm{T}} g(x_0) + v_0^{\mathrm{T}} h(x_0) = 0,$$

and since $h(x_0) = 0$, this reduces to

$$u_0^{\mathrm{T}} g(x_0) = 0.$$

Now, since $u_0 \geqslant 0$ and $g(x_0) \leqslant 0$, it must be the case that $(u_0)_j g_j(x_0) = 0$ for $j = 1, \ldots, m_1$. But the existence of such a u_0 and v_0, with $\|{}^{u_0}_{v_0}\|_0 = 1$, in conjunction with (c)″ violates Motzkin's theorem of the alternative [27]. Hence the supremum in part (4) above is indeed bounded. □

By combining parts (3) and (4) of the above theorem, we obtain the following corollary.

Corollary 5.1. *Let $x_0 \in \mathbb{R}^n$ be such that $g(x_0) \leqslant 0$ and $h(x_0) = 0$, and the MFCQ is satisfied at x_0. Also let $0 < \sigma_l < \sigma_r < \bar{\beta}$ and let $\mathcal{H}$ be a nonempty compact set of $n \times n$ symmetric positive definite matrices. Then there is a neighborhood U of x_0 and a constant $\kappa \geqslant 0$ such that*

$$0 \leqslant -\Delta(x, \sigma)^{-1} \left\| \begin{pmatrix} z \\ h(x) \end{pmatrix} - \begin{pmatrix} r_1(x) \\ r_2(x) \end{pmatrix} \right\| \operatorname{dist}_0(0 \mid M(x, H, \sigma, \beta)) \leqslant \kappa$$

for all $(x, \sigma, \beta, H) \in U \times \Gamma(\sigma_l, \sigma_r, \bar{\beta}) \times \mathcal{H}$ where $\Gamma(\sigma_l, \sigma_r, \bar{\beta}) := \{(\sigma, \beta) \colon \sigma \in [\sigma_l, \sigma_r], \beta \in (\sigma, \bar{\beta}]\}$. □

In the case where $g(x) \leqslant 0$, $h(x) = 0$ and x satisfies the MFCQ, Corollary 5.1 and inequality (5.2) assure us that (5.1) remains locally bounded. In the absence of the MFCQ, it is well known that (5.4) can be locally unbounded. Thus, in a sense, Corollary 5.1 is the best one can expect concerning the local boundedness of the right-hand side of (5.2).

Points that are feasible for (1.1) and at which the MFCQ is not satisfied will be called *Fritz John points* in deference to the foundational work of Fritz John [17]. The significance of Fritz John points, with regard to first-order necessary conditions for (1.1), is illustrated in the following result.

Theorem 5.2. (Clarke [7, Theorem 6.1.1]). *Let x solve* (1.1). *Then x is either a Kuhn-Tucker point or a Fritz John point, or both.*

Remark. In the above theorem it is indeed possible that x is not a Kuhn-Tucker point. This is essentially the reason why one requires the MFCQ in establishing the boundedness of (5.4).

Given the results of this section, we can now state the goal of the convergence theory of the next section. The goal is to show that every cluster point of the sequence of iterates generated by the algorithm of Section 4 is either a Kuhn-Tucker point for (1.1), or a Fritz John point for (1.1), or a stationary point for ϕ that is not feasible for (1.1).

6. Convergence

Theorem 6.1. *Let $\{x_i\}$ be a sequence generated by the algorithm of Section* 4, *and suppose that the mappings ∇f, g', and h' are bounded on $\{x_i\}$ and uniformly continuous on $\overline{\text{co}}\{x_i\}$.*

(1) *If $\alpha_i \uparrow \infty$, then $\lim_{i\in S} \Delta(x_i, \sigma_i) = 0$ where S is the subsequence of indices $\{i: \alpha_i < \alpha_{i+1}\}$*

(2) *if $\sup_i \alpha_i =: \alpha < \infty$, then either*

(a) $\inf_i P_{\bar{\alpha}}(x_i) = -\infty$ *for all $\bar{\alpha} \in [0, \alpha]$, or*

(b) *$\{x_i\}$ is finitely terminating at $\bar{x}$, where $\bar{x}$ is either a nonfeasible stationary point of ϕ, or $\bar{x}$ is a Kuhn-Tucker point for* (1.1), *or*

(c) $\|d_i\| \to 0$, $\Delta(x_i, \sigma_i) \to 0$, *and* $\nabla_x L(x_i, u_i, v_i) \to 0$ *where* $L(x, u, v) = f(x) + u^{\mathrm{T}} g(x) + v^{\mathrm{T}} h(x)$ *and* $\binom{u_i}{v_i}$ *is any element of* $M(x_i, H_i, \sigma_i, \beta_i)$ *for each* $i = 0, 1, 2, \ldots$.

Remark. The hypotheses on ∇f, g', and h' allow us to avoid specifying that the iterates $\{x_i\}$ remain bounded.

Proof. (1) If $\alpha_i \uparrow \infty$, then by step (3) of the algorithm the sequence

$$\left\{\frac{\nabla f(x_i)^{\mathrm{T}} d_i + d_i^{\mathrm{T}} H_i d_i}{-\Delta(x_i, \sigma_i)}\right\}_{i\in S}$$

diverges to $+\infty$. Now, since all of the sequences $\{\nabla f(x_i)\}$, $\{d_i\}$, and $\{H_i\}$ are bounded, it must be the case that $\lim_{i\in S} \Delta(x_i, \sigma_i) = 0$.

(2) If the sequence $\{x_i\}$ terminates finitely then (b) follows from step (2) of the algorithm and Lemma 2.2. Thus we will assume that neither (a) nor (b) occur, and establish (c). By step (3) of the algorithm, there is no loss of generality in assuming that $\alpha_i = \alpha$ for all $i = 1, 2, \ldots$. Hence, since $P_\alpha(x_i)$ is bounded below, we have from step (4) that

$$[P_\alpha(x_{i+1}) - P_\alpha(x_i)] \to 0.$$

Thus, by steps (3) and (4),

$$\lambda_i d_i^{\mathrm{T}} H d_i \to 0.$$

If $d_i^{\mathrm{T}} H_i d_i \to 0$, then $\|d_i\| \to 0$ due to the compactness of $\mathcal{H}$. Also,

$$\Delta(x_i, \sigma_i) \to 0,$$

since, by Lemma 2.1,

$$\begin{aligned} 0 &\leq -\Delta(x_i, \sigma_i) \\ &\leq \phi(x_i) - \operatorname{dist}\left[\begin{pmatrix} g(x_i) + g'(x_i)d_i \\ h(x_i) + h'(x_i)d_i \end{pmatrix} \middle| K\right] \\ &\leq \operatorname{dist}\left(\begin{pmatrix} g'(x_i)d_i \\ h'(x_i)d_i \end{pmatrix} \middle| K\right) \\ &\leq \left|\begin{matrix} g'(x_i)d_i \\ h'(x_i)d_i \end{matrix}\right|, \end{aligned}$$

where both $\{g'(x_i)\}$ and $\{h'(x_i)\}$ are bounded. Moreover, by equation (2.6(c)), for any sequence $\{\binom{u_i}{v_i}\}$, with $\binom{u_i}{v_i} \in M(x_i, H_i, \sigma_i, \beta_i)$ for each $i = 0, 1, 2, \ldots$, there is a sequence $\{z_i\}$, with $z_i \in N(d_i | \beta_i B)$ for each $i = 0, 1, 2, \ldots$, such that

$$\nabla_x L(x_i, u_i, v_i) = -[H_i d_i + z_i].$$

Now, since $d_i \to 0$, eventually $N(d_i | \beta_i B) = \{0\}$. Therefore, $\nabla_x L(x_i, u_i, v_i) \to 0$. Thus, we need only consider the possibility that $d_i \nrightarrow 0$. In this case, there must exist a subsequence $\{x_i : i \in T\}$ such that

$$\lim_{i \in T} \lambda_i = 0,$$

with $\lambda_i < \gamma_1$ for all $i \in T$, and

$$\inf_T \{d_i^{\mathrm{T}} H_i d_i\} =: \zeta > 0.$$

Step (4) then yields the existence of a sequence $\{\bar{\lambda}_i : i \in T\}$ for which $\lambda_i \geq \gamma_2 \bar{\lambda}_i > 0$ and

$$\mu_2 \bar{\lambda}_i [\nabla f(x_i)^{\mathrm{T}} d_i + \alpha \Delta(x_i, \sigma_i)] < P_\alpha(x_i + \bar{\lambda}_i d_i) - P_\alpha(x_i) \tag{6.1}$$

for all $i \in T$. We assume, with no loss of generality, that $\bar{\lambda}_i < 1$ for all $i \in T$. Next, observe that

$$\begin{aligned} &\operatorname{dist}\left[\begin{pmatrix} g(x_i + \bar{\lambda}_i d_i) \\ h(x_i + \bar{\lambda}_i d_i) \end{pmatrix} \middle| K\right] - \operatorname{dist}\left[\begin{pmatrix} g(x_i) \\ h(x_i) \end{pmatrix} \middle| K\right] \\ &\quad \leq \left\| \begin{pmatrix} g(x_i + \bar{\lambda}_i d_i) \\ h(x_i + \bar{\lambda}_i d_i) \end{pmatrix} - \begin{pmatrix} g(x_i) + \bar{\lambda}_i g'(x_i) d_i \\ h(x_i) + \bar{\lambda}_i h'(x_i) d_i \end{pmatrix} \right\| \\ &\quad\quad + \operatorname{dist}\left[\begin{pmatrix} g(x_i) + \bar{\lambda}_i g'(x_i) d_i \\ h(x_i) + \bar{\lambda}_i h'(x_i) d_i \end{pmatrix} \middle| K\right] - \operatorname{dist}\left[\begin{pmatrix} g(x_i) \\ h(x_i) \end{pmatrix} \middle| K\right] \end{aligned}$$

$$\leqslant \bar{\lambda}_i \|d_i\| \omega_1(\bar{\lambda}_i d_i) + \bar{\lambda}_i \left[\operatorname{dist}\left[\begin{pmatrix} g(x_i) + g'(x_i) d_i \\ h(x_i) + h'(x_i) d_i \end{pmatrix} \middle| K \right] - \operatorname{dist}\left[\begin{pmatrix} g(x_i) \\ h(x_i) \end{pmatrix} \middle| K \right] \right]$$

$$\leqslant \bar{\lambda}_i \|d_i\| \omega_1(\bar{\lambda}_i d_i) + \bar{\lambda}_i \Delta(x_i, \sigma_i),$$

where $\omega_1(\bar{\lambda}_i d_i)$ is the modulus of continuity for $\binom{g'}{h'}$. Hence, from (6.1), we obtain

$$\mu_2 \bar{\lambda}_i [\nabla f(x_i)^{\mathrm{T}} d_i + \alpha \Delta(x_i, \sigma_i)]$$
$$\leqslant \bar{\lambda}_i [\nabla f(x_i)^{\mathrm{T}} d_i + \alpha \Delta(x_i, \sigma_i)] + \|f(x_i + \bar{\lambda}_i d_i) - (f(x_i) + \bar{\lambda}_i \nabla f(x_i)^{\mathrm{T}} d_i)\|$$
$$+ \bar{\lambda}_i \|d_i\| \alpha \omega_1(\bar{\lambda}_i d_i)$$
$$\leqslant \bar{\lambda}_i [\nabla f(x_i)^{\mathrm{T}} d_i + \alpha \Delta(x_i, \sigma_i)] + \bar{\lambda}_i \|d_i\| [\alpha \omega_1(\bar{\lambda}_i d_i) + \omega_2(\bar{\lambda}_i d_i)],$$

where $\omega_2(\bar{\lambda}_i d_i)$ is the modulus of continuity for ∇f. Therefore, in view of Step (3),

$$0 \leqslant (1 - \mu_2)[\nabla f(x_i)^{\mathrm{T}} d_i + \alpha \Delta(x_i, \sigma_i)] + \|d_i\| \omega(\bar{\lambda}_i d_i)$$
$$\leqslant (\mu_2 - 1)\zeta + \|d_i\| \omega(\bar{\lambda}_i d_i)$$

where $\omega := \alpha\omega_1 + \omega_2$. Taking the limit in i, for $i \in T$, we obtain the contradiction

$$0 \leqslant (\mu_2 - 1)\zeta < 0.$$

Hence it is necessarily the case that $\|d_i\| \to 0$. □

The above theorem provides a fairly general picture of the convergence properties of the algorithm. It is important to note that the procedure can fail even though the penalty parameters remain bounded. This problem is generic to methods that are dependent upon exact penalization. We will return to this issue again in the concluding section. Let us now proceed to investigate the nature of the cluster points of the algorithm when they exist.

Corollary 6.1. *Let the hypotheses of Theorem* 6.1 *hold, and let* $\{x_i\}$ *be a sequence generated by the algorithm of Section* 4. *If* $\alpha_i \uparrow \infty$, *then every cluster point of the subsequence* $\{x_i\colon \alpha_i < \alpha_{i+1}\}$ *is either a Fritz John point for* (1.1), *or a stationary point of* ϕ *that is not feasible for* (1.1).

Proof. By part (1) of Theorem 6.1, we know that the subsequence $\{\Delta(x_i, \sigma_i)\colon \alpha_i < \alpha_{i+1}\}$ converges to zero. Let $\bar{x}$ be a cluster point of the subsequence $\{i\colon \alpha_i < \alpha_{i+1}\}$ and let $T \subset \{i\colon \alpha_i < \alpha_{i+1}\}$ be such that $x_i \to^T \bar{x}$ and $\sigma_i \to^T \bar{\sigma}$ for some $\bar{\sigma} \in [\sigma_l, \sigma_r]$. Then, by Proposition 2.1, $\Delta(\bar{x}, \bar{\sigma}) = 0$. Hence, by Lemma 2.1, $\bar{x}$ is a stationary point of ϕ. If $\bar{x}$ is feasible, then, by step (3) of the algorithm, inequality (5.2), and Part (4) of Theorem 5.1, $\bar{x}$ cannot satisfy the MFCQ and so is a Fritz John point. □

Corollary 6.2. *Let the hypotheses of Theorem* 6.1 *hold, and let* $\{x_i\}$ *be a sequence generated by the algorithm of Section* 4. *If* $\sup_i \alpha_i < \infty$, *then every cluster point of* $\{x_i\}$ *is a stationary point of* ϕ. *Moreover, any cluster point that is also feasible for* (1.1) *is either a Kuhn–Tucker point for* (1.1) *or a Fritz John point for* (1.1).

Proof. Since $\sup_i \alpha_i < \infty$, the sequence $\{\alpha_i\}$ is eventually constant. Denote this constant value by α, and suppose that $\alpha_i = \alpha$ for all $i \geq i_0$. Then for all $i \geq i_0$ the sequence $\{P_\alpha(x_i)\}$ is decreasing. If the sequence $\{x_i\}$ has a cluster point $\bar{x}$, then $P_\alpha(x_i) \downarrow P_\alpha(\bar{x})$. Thus the sequence $\{P_\alpha(x_i)\}$ is bounded below. Hence, by part (2) of Theorem 6.1 we know that either the sequence $\{x_i\}$ is finitely terminating at $\bar{x}$ where $\bar{x}$ is either a Kuhn–Tucker point or a nonfeasible stationary point for ϕ, or $\Delta(x_i, \sigma_i) \to^{i \in T} 0$ and $\nabla_x L(x_i, u_i, v_i) \to^{i \in T} 0$ where $\sigma_i \to^T \bar{\sigma}$ with $\bar{\sigma} \in [\sigma_l, \sigma_r]$, $x_i \to^{i \in T} \bar{x}$ and $\{\binom{u_i}{v_i}\}$ is any sequence of multipliers such that $\binom{u_i}{v_i} \in M(x_i, H_i, \sigma_i, \beta_i)$ for each $i \in T$ for some subsequence T. If the sequence is finitely terminating we are done. Thus, we will assume that $\{x_i\}$ is infinite. By Proposition 2.1, $\Delta(\bar{x}, \bar{\sigma}) = 0$. Hence, by Lemma 2.1, $0 \in \partial\phi(\bar{x})$. If, moreover, $\bar{x}$ is feasible for (1.1) and the sequence $\{\text{dist}_0(0 \mid M(x_i, H_i, \sigma_i, \beta_i)) \colon i \in T\}$ is bounded, then, by passing to a further subsequence if necessary, one can choose $\binom{u_i}{v_i} \in M(x_i, H_i, \sigma_i, \beta_i)$ such that $\binom{u_i}{v_i} \to^{i \in T} \binom{u}{v}$ for some vectors $\binom{u}{v}$. Due to continuity, we have that $\nabla_x L(\bar{x}, u, v) = 0$, $u \geq 0$, and $u^T g(\bar{x}) = 0$. Hence $\bar{x}$ is a Kuhn–Tucker point for (1.1). On the other hand, if $\bar{x}$ is feasible for (1.1) and the sequence $\{\text{dist}_0(0 \mid M(x_i, H_i, \sigma_i, \beta_i)) \colon i \in T\}$ is unbounded, then, by part (4) of Theorem 5.1, $\bar{x}$ is a Fritz John point for (1.1). □

The next corollary follows immediately from Corollaries 6.1 and 6.2.

Corollary 6.3. *Let* $\{x_i\}$ *be a sequence generated by the algorithm of Section* 4. *If the sequence* $\{x_i\}$ *is bounded, then there is a cluster point that is either a Kuhn–Tucker point for* (1.1), *a Fritz John point for* (1.1), *or a stationary point for* ϕ *that is not feasible for* (1.1).

We conclude this section by discussing the convergence of the method when it is applied to the problem

$$\min_{x \in \mathbb{R}} f(x)$$

$$\text{subject to } 1 - e^x = 0, \; x = 0$$

where f is any continuously differentiable function on $\mathbb{R}$. The relevant information on this problem can be found in example (2) of Section 2. If $\mathbb{R}^2$ is equipped with the l_2-norm, then it is not difficult to show that $0 < -d_x < x$ for all $x > 0$ where d_x solves $Q(x, H, \sigma, \beta)$, regardless of the choice of $0 < \sigma < \beta$, and $H \in \mathcal{H}$. Hence, eventually, the sequence generated by the algorithm of Section 4 is either increasing with upper bound zero, or decreasing with lower bound zero. Consequently, the sequence is bounded and so Corollary 6.3 applies. Therefore, the sequence must converge to $\bar{x} = 0$ since $\bar{x} = 0$ is the only stationary point for this problem.

7. Computation with monotone norms

In this section we describe a few examples for which the computations required in the algorithm of Section 4 reduce to either linear or quadratic programs.

Definition 7.1. A norm $\|\cdot\|$ on $\mathbb{R}^m$ is said to be monotone with respect to the closed convex cone $\mathbb{R}^m_+$ if

$$\|x\| \leqslant \|y\| \quad \text{whenever } 0 \leqslant x \leqslant y.$$

Moreover, the norm $\|\cdot\|$ is said to be orthogonally monotone with respect to the cone $-K = \mathbb{R}^{m_1}_+ \times \{0\}_{\mathbb{R}^{m-m_1}}$ if $\left\|\binom{x}{z}\right\| \leqslant \left\|\binom{y}{z}\right\|$ whenever $x, y \in \mathbb{R}^{m_1}_+$ with $0 \leqslant x \leqslant y$ and $z \in \mathbb{R}^{m-m_1}$.

Lemma 7.1. *If $\|\cdot\|$ is a norm on $\mathbb{R}^m$ with $m = m_1 + m_2$ that is orthogonally monotone with respect to the cone $-K = \mathbb{R}^{m_1}_+ \times \{0\}_{\mathbb{R}^{m_2}}$, then*

$$\operatorname{dist}\left[\begin{pmatrix} y \\ z \end{pmatrix} \middle| K\right] = \left\|\begin{matrix} y_+ \\ z \end{matrix}\right\| \quad \textit{for all } \begin{pmatrix} y \\ z \end{pmatrix} \in \mathbb{R}^m.$$

Proof. Note that if $c \in \mathbb{R}^{m_1}_+$ then $0 \leqslant y_+ \leqslant (y+c)_+$ for every $y \in \mathbb{R}^{m_1}$. Hence

$$\begin{aligned}
\left\|\begin{matrix} y_+ \\ x \end{matrix}\right\| &\geqslant \operatorname{dist}\left[\begin{pmatrix} y \\ z \end{pmatrix} \middle| K\right] \geqslant \inf_{c \in \mathbb{R}^{m_1}_+} \left\|\begin{matrix} (y+c) \\ z \end{matrix}\right\| \\
&= \inf_{c \in \mathbb{R}^{m_1}_+} \left\|\begin{matrix} (y+c)_+ - (y+c)_- \\ z \end{matrix}\right\| \geqslant \inf_{c \in \mathbb{R}^{m_1}_+} \left\|\begin{matrix} (y+c)_+ \\ z \end{matrix}\right\| \\
&\geqslant \left\|\begin{matrix} y_+ \\ z \end{matrix}\right\|. \qquad \square
\end{aligned}$$

Thus for norms orthogonally monotone with respect to K we obtain the following identities for all $x \in \mathbb{R}^n$:

$$\begin{bmatrix} g(x)_+ \\ h(x) \end{bmatrix} \in R_0(x),$$

and

$$\|r(x)\| = \inf\left\{\left\|\begin{matrix} (g(x) + g'(x)d)_+ \\ h(x) + h'(x)d \end{matrix}\right\| : \|d\| \leqslant \sigma\right\}.$$

From these norms, the 1, 2 and ∞-norms will be of special interest, since for these norms the computation of a residual vector $r(x) \in R(x, \sigma)$ reduces to solving a linear or quadratic program. We briefly list a sampling of these programs.

Example 7.1. $(\mathbb{R}^n, \|\cdot\|_\infty)$, $(\mathbb{R}^m, \|\cdot\|_1)$.
Solve

$$\min_{(d,z_1,z_2)} \; e^{\mathrm{T}} z_1 + e^{\mathrm{T}} z_2$$

$$\begin{aligned} \text{subject to}\;\; & g(x)+g'(x)d \leqslant z_1, \\ & 0 \leqslant z_1, \\ & -z_2 \leqslant h(x)+h'(x)d \leqslant z_2, \\ & \sigma e \leqslant d \leqslant \sigma e. \end{aligned}$$

Then set $r(x) := \begin{bmatrix} (g(x)+g'(x)d^*)_+ \\ h(x)+h'(x)d^* \end{bmatrix}$ where d^* is optimal for the above program.

Example 7.2. $(\mathbb{R}^n, \|\cdot\|_1)$, $(\mathbb{R}^m, \|\cdot\|_2)$.
Solve

$$\min_{(d,z_1,z_2,y)} \; \tfrac{1}{2}(z_1^{\mathrm{T}} z_1 + z_2^{\mathrm{T}} z_2)$$

$$\begin{aligned} \text{subject to}\;\; & g(x)+g'(x)d \leqslant z_1 \\ & h(x)+h'(x)d = z_2 \\ & -y \leqslant d \leqslant y \\ & e^{\mathrm{T}} y \leqslant \sigma \end{aligned}$$

Then set $r(x) := \begin{bmatrix} (g(x)+g'(x)d^*)_+ \\ h(x)+h'(x)d^* \end{bmatrix}$ where d^* is optimal for the above program.

Example 7.3. $(\mathbb{R}^n, \|\cdot\|_\infty)$, $(\mathbb{R}^m, \|\cdot\|_\infty)$.
Solve

$$\min_{(d,\gamma)} \gamma$$

$$\begin{aligned} \text{subject to}\;\; & g(x)+g'(x)d \leqslant \gamma e, \\ & -\gamma e \leqslant h(x)+h'(x)d \leqslant \gamma e, \\ & -\sigma e \leqslant d \leqslant \sigma e. \end{aligned}$$

Then set $r(x) = \begin{bmatrix} (g(x)+g'(x)d^*)_+ \\ h(x)+h'(x)d^* \end{bmatrix}$ where d^* is optimal for the above program. Note that if $m_2 = 0$, then one must add the constraint $0 \leqslant \gamma$ to the above program.

8. Two examples

The two examples considered in this section have been included in order to indicate why the algorithm of this paper, or a similar method, should be included in a standard repertoire of mathematical programming techniques. The examples are such that the algorithm of Section 4 has no difficulty locating a solution for them. However, the SQP method of Wilson, Han and Powell fails to solve the first example, while the QL method of Fletcher fails to solve the second. In general, this type of comparison is of course quite unfair since the examples are specifically constructed to make the SQP and QL algorithms fail while the algorithm of Section 4 succeeds. However, the examples are instructive since they illustrate how the proposed method overcomes some of the drawbacks of the SQP and QL methods.

The example that we consider in conjunction with the SQP method has already been studied in example (2) of Section 2. Recall that all of the modifications to the SQP method considered by Powell [21] are defeated by this example regardless of the objective function f as long as the initial point x_0 is not chosen to be the solution $x=0$. On the other hand, it was shown at the end of Section 6 that the method of Section 4 converges monotonically to the solution $x=0$ regardless of the initial point x_0 and the choice of the C^1 objective function f.

Let us now consider the QL method. Recall that the QL method is a trust region algorithm for the global minimization of an exact penalty function for (1.1). The key element in the overall success of the method is the initial choice of the penalty parameter and a method for iteratively adjusting its value. For example, if one employs the l_1 penalty function (as in Fletcher [9]),

$$P_\alpha(x) := f(x) + \alpha[\|g_+(x)\|_1 + \|h(x)\|_1],$$

then locally the choice of the penalty parameter depends upon the following well known results.

Theorem 8.1. *Let $f:\mathbb{R}^n \to \mathbb{R}$, $g:\mathbb{R}^n \to \mathbb{R}^{m_1}$, and $h:\mathbb{R}^n \to \mathbb{R}^{m_2}$ be as given in* (1.1).

(1) *Let $(x, u, v) \in \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ be a Kuhn–Tucker triple for* (1.1), *then x is a stationary point for P_α for all $\alpha \geq \|\binom{u}{v}\|_\infty$.*

(2) *If $x \in \mathbb{R}^n$ is such that $g(x) \leq 0$ and $h(x) = 0$ and x is a stationary point for P_α for some $\alpha > 0$, then x is a Kuhn–Tucker point for* (1.1).

(3) *If* (1.1) *satisfies the second-order sufficiency condition at (x, u, v), then x is a strict local minimum of P_α whenever $\alpha > \|\binom{u}{v}\|_\infty$.* □

Remark. The proof of this result can be found in [6], [9], or [16].

Consequently, one should choose the penalty parameter to be greater than the l_∞-norm of one of the optimal vectors of Kuhn–Tucker multipliers. Of course, it may be possible that no such vector of Kuhn–Tucker multipliers exists, as is the case in the problem

$$\min\{x_1\colon x_2 \leq x_1^3,\ -x_1^3 \leq x_2\},$$

in which case the QL algorithm does not apply. But, even if a vector of Kuhn–Tucker multipliers exists, this choice of the penalty parameter may still be insufficient to assure the convergence of the QL method from points arbitrarily close to a solution point $\bar{x}$. In order to illustrate this possibility, we consider problems of the form

$$\begin{aligned}&\min\ -w(x)x\\&\text{subject to}\ \ x=0\end{aligned}\tag{8.1}$$

where $w:\mathbb{R}\to\mathbb{R}$ is C^3 and satisfies $w(0)\geqslant 0$, $w'(x)\geqslant 0$, $w''(x)\geqslant 0$, and $w'''(x)\geqslant 0$ for all $x\geqslant 0$. The solution to (8.1) is obviously $x=0$, and the associated Kuhn–Tucker multiplier is $w(0)$. Suppose $x_0>0$. If α_i is the value of the penalty parameter at the ith iteration, then in the first step of the ith iteration of the QL algorithm one solves the program

$$\begin{aligned}&\min\ -[(x_iw'(x_i)+w(x_i))d+\tfrac{1}{2}(x_iw''(x_i)+2w'(x_i))d^2]+\alpha_i\|x_i+d\|\\&\text{subject to}\ \ \|d\|\leqslant\delta_i.\end{aligned}\tag{8.2}$$

Hence, if

$$w(x_i)+w'(x_i)x_i>\alpha_i,$$

then the solution to (8.2) is $d_i=\delta_i$. Given the hypotheses on the function w, it is easy to show that the QL algorithm will accept this step since the ratio of actual to predicted reduction in the penalty function, q_i, is greater than or equal to unity. Consequently, given x_0, δ_0, and $\{\alpha_i\}$ it is possible to construct a function w so that $x_i=x_0+\sum_{j=0}^{i-1}\delta_j$ and $q_i\uparrow\infty$. For example, if $\delta_0=1$, we could set

$$w(x)=\alpha_0+\alpha_1+x+\sum_{k=1}^{\infty}w_k(x)$$

where for each $k=1,2,\ldots$

$$w_k(x)=\begin{cases}0 & \text{if } x\leqslant k,\\ \alpha_{k+1}(x-k)^4 & \text{otherwise.}\end{cases}$$

Next, suppose the algorithm of Section 4 is applied to (8.1) using the Armijo-like line search procedure

$$\begin{aligned}\lambda_i:=\ &\max\ \gamma^\nu\\&\text{subject to}\ \ \nu\in\{0,1,2,\ldots\},\\&P_{\alpha_{i+1}}(x_i+\gamma^\nu d_i)-P_{\alpha_{i+1}}(x_i)\leqslant\delta\gamma^\nu[\nabla f(x_i)^{\mathrm{T}}d_i+\alpha_{i+1}\Delta(x_i,\sigma_i)],\end{aligned}$$

where $\gamma\in(0,1)$ and $\delta\in(0,1)$ are chosen during the initialization stage. Then, if the initial point x_0 is such that $x_0>0$, it is straightforward to show that the procedure must terminate after a finite number of iterations at a solution to (8.1). In order to see this, first verify that

$$0\leqslant x_{i+1}\leqslant x_i$$

for all $i=0, 1, \ldots$. Hence $P_{\alpha_{i+1}}(x_i)$ is bounded below and so by Theorem 6.1 there is at least a subsequence $S\subset\{1, 2, \ldots\}$ for which $\Delta(x_i, \sigma_i)\to^S 0$. Consequently, after a finite number of iterations, one reaches a point x_{i_0} for which $-x_{i_0}=d_{i_0}$. Since $w'(x_i)\geqslant 0$, the procedure for updating the α_i's implies that $x_{i_0+1}=0$ and the algorithm terminates.

Though Fletcher's Sl_1QP method works satisfactorily for most well-posed problems, the above example shows that there are still rooms for improvement in how it adjusts the sizes of trust regions and updates penalty parameter in dealing with some very difficult situations. However, our algorithm seems very effective for those irregular cases and provides a reasonable alternative for handling such ill-posed problems. Because all these opinions are based on purely theoretical viewpoint, no conclusion can be made without further computational testing.

Acknowledgement

The authors gratefully acknowledge the detailed comments and suggestions of the referees. Their efforts have improved both the content and the presentation of the results in the paper. In particular, the rule that we employ to update the penalty parameter in step 3 of the algorithm in Section 4 was proposed by one of the referees. This suggestion greatly simplified the convergence theory of Section 6.

References

[1] M.C. Biggs, "Constrained minimization using recursive equality quadratic programming," in: F.A. Lootsma, ed., *Numerical Methods for Non-linear Optimization* (Academic Press, London, New York, 1972) pp. 411–428.

[2] J.V. Burke, "Descent methods for composite nondifferentiable optimization problems," *Mathematical Programming* 33 (1985) 260–279.

[3] J.V. Burke, "Methods for solving generalized inequalities with applications to nonlinear programming," P.h.D. Thesis, Department of Mathematics, University of Illinois at Urbana-Champaign (1983).

[4] J.V. Burke and S-P. Han, "A Gauss–Newton approach to solving generalized inequalities," *Mathematics of Operations Research* 4 (1986) 632–643.

[5] P.H. Calamai and J.J. Moré, "Projected gradient methods for linearly constrained problems," *Mathematical Programming* 39 (1987) 93–116.

[6] C. Charalambous, "A lower bound for the controlling parameter of the exact penalty function," *Mathematical Programming* 15 (1978) 278–290.

[7] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advance Texts (John Wiley & Sons, 1983).

[8] R. Fletcher, "A model algorithm for composite nondifferentiable optimization problems," *Mathematical Programming Studies* 17 (1982) 67–76.

[9] R. Fletcher, *Practical Methods for Optimization, Vol. II, Constrained Optimization* (Wiley, New York, 1981).

[10] R. Fletcher, "Second order corrections for nondifferentiable optimization," *Numerical Analysis Proceedings, Dundee, 1981,* Lecture Notes in Mathematics 912 (Springer-Verlag, 1981) pp. 85–114.

[11] U.M. Garcia-Palomares and O.L. Mangasarian, "Superlinearly convergent quasi-Newton algorithms for nonlinearly constrained optimization problems," *Mathematical Programming* 11 (1976) 1–13.

[12] J. Gauvin, "A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming," *Mathematical Programming* 12 (1977) 136-138.

[13] S-P. Han, "Least-squares solution of linear inequalities," Technical Summary Report 2141, Mathematical Research Center, University of Wisconsin (Madison, WI, 1980).

[14] S-P. Han, "A globally convergent method for nonlinear programming," *Journal Optimization Theory Applications* 22 (1977) 297-309.

[15] S-P. Han, "Superlinearly convergent variable metric algorithms for general nonlinear programming problems," *Mathematical Programming* 11 (1976) 263-282.

[16] S-P. Han and O.L. Mangasarian, "Exact penalty functions in nonlinear programming," *Mathematical Programming* 17 (1979) 251-269.

[17] F. John, "Extremum problems with inequalities as subsidiary conditions," *Studies and Essays Presented to R. Courant on his 60th Birthday* (Interscience, New York, NY 1948) pp. 187-204.

[18] O.L. Mangasarian, *Nonlinear Programming* (Robert E. Krieger Publishing, Inc.) (reprint edition 1979).

[19] V.H. Nguyen, J-J. Strodiot, and R. Mifflin, "On conditions to have bounded multipliers in locally Lipschitz programming," *Mathematical Programming* 18 (1980) 100-106.

[20] M.J.D. Powell, "Algorithms for nonlinear constraints that use Lagrangian functions," *Mathematical Programming* 14 (1978) 224-248.

[21] M.J.D. Powell, "A fast algorithm for nonlinearly constrained optimization calculations," *Proceedings of the 1977 Dundee Biennial Conference on Numerical Analysis* (Springer-Verlag, Berlin, 1977).

[22] S.M. Robinson, "Stability theory for systems of inequalities, part II: differential nonlinear systems," *SIAM Journal on Numerical Analysis* 4 (1976) 497-513.

[23] S.M. Robinson, "Perturbed Kuhn-Tucker points, and rates of convergence for a class of nonlinear programming algorithms," *Mathematical Programming* 7 (1974) 1-16.

[24] R.T. Rockafeller, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).

[25] K. Schittkowski, "The nonlinear programming method of Wilson, Han and Powell with an augmented Lagrangian type line search function," *Numerische Mathematik* 38 (1981) 83-114.

[26] K. Schittkowski, "Nonlinear programming methods with linear least squares subproblem," Lecture Notes in Economics and Mathematical Systems 199 (Springer-Verlag, Berlin-New York, 1982).

[27] J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions I* (Springer-Verlag, New York, NY, 1970).

[28] L.P. Vlasov, "Approximate properties of sets in normed linear spaces," *Russian Mathematical Surveys* 28 (1975) 1-62.

[29] R.B. Wilson, "A simplicial algorithm for concave programming," Ph.D. Dissertation, Graduate School of Business Administration, Harvard University (Boston, MA, 1963).