SIAM J. OPTIM.
Vol. 19, No. 2, pp. 846–862
© 2008 Society for Industrial and Applied Mathematics

CONVERGENCE RATE OF AN OPTIMIZATION ALGORITHM FOR MINIMIZING QUADRATIC FUNCTIONS WITH SEPARABLE CONVEX CONSTRAINTS*

RADEK KUČERA†

Abstract. A new active set algorithm for minimizing quadratic functions with separable convex constraints is proposed by combining the conjugate gradient method with the projected gradient. It generalizes recently developed algorithms of quadratic programming constrained by simple bounds. A linear convergence rate in terms of the Hessian spectral condition number is proven. Numerical experiments, including the frictional three-dimensional (3D) contact problems of linear elasticity, illustrate the computational performance.

Key words. quadratic function, separable convex constraints, active set, conjugate gradient method, projected gradient, convergence rate

AMS subject classifications. 65K05, 90C25

DOI. 10.1137/060670456

1. Introduction.

We shall be concerned with solving

$$\min_{x\in\Omega} f(x), \tag{1.1}$$

where $f(x) = \frac{1}{2}\,x^\top Ax - x^\top b$, $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite, $b \in \mathbb{R}^n$, $\Omega = \Omega_1 \times \cdots \times \Omega_m$, and $\Omega_i = \{\boldsymbol{x}_i \in \mathbb{R}^{n_i} : f_i(\boldsymbol{x}_i) \le 0\}$ are defined by continuously differentiable convex functions $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ so that $n_i \ge 1$, $\sum_{i=1}^m n_i = n$. Let us note that the feasible set Ω is *separable* in the sense that each part $\boldsymbol{x}_i$ of $x = (\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_m^\top)^\top$ is subject to one constraint $\boldsymbol{x}_i \in \Omega_i$.

This problem includes several independently investigated subproblems originating, for instance, in duality-based methods for the solution of contact problems of linear elasticity:

- If $n_i = 1$ and $f_i(x_i) \equiv l_i - x_i$ with l_i given, we obtain the *simple bound*

$$l_i \le x_i \tag{1.2}$$

  arising from two-dimensional contact problems [4].
- If $n_i = 2$, $\boldsymbol{x}_i = (x_{2i-1}, x_{2i})^\top$ and $f_i(\boldsymbol{x}_i) \equiv x_{2i-1}^2 + x_{2i}^2 - r_i^2$ with r_i given, we arrive at

$$x_{2i-1}^2 + x_{2i}^2 \le r_i^2 \tag{1.3}$$

  that can be interpreted as the *circular constraint*. A source of such constraints is an isotropic friction law for three-dimensional (3D) contact problems [8].
- If an anisotropic friction law is considered and if a finite element tearing and interconnecting (FETI) domain decomposition method is used for solving 3D

*Received by the editors September 22, 2006; accepted for publication (in revised form) April 3, 2008; published electronically August 1, 2008. This work was supported by grant 101/08/0574 of the Grant Agency of the Czech Republic and by the Research Project MSM6198910027 of the Czech Ministry of Education.

http://www.siam.org/journals/siopt/19-2/67045.html

†Department of Mathematics and Descriptive Geometry, VŠB-Technical University of Ostrava, 17. listopadu 15, CZ-70833 Ostrava-Poruba, Czech Republic (radek.kucera@vsb.cz).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

contact problems [11], then (1.3) is replaced by the *ellipsoidal constraint*,

$$(\boldsymbol{x}_i - \boldsymbol{s}_i)^\top \boldsymbol{F}_i (\boldsymbol{x}_i - \boldsymbol{s}_i) \le r_i^2,$$

with $\boldsymbol{F}_i \in \mathbb{R}^{2\times 2}$ positive definite and $\boldsymbol{s}_i = (s_{2i-1}, s_{2i})^\top$ given.

– Finally let us note that the unconstrained case may be described by $f_i(\boldsymbol{x}_i) = -1$.

Through the whole paper, we shall have in mind large-scale problems in which the Hessian matrix A is not formed explicitly. In this case, an iterative method is a suitable tool for solving (1.1) since it is the only action of A which is needed. A class of efficient algorithms appropriate for our research is based on the active set strategy that dates back at least to Polyak [13]. His algorithm solves quadratic programming problems constrained by simple bounds (1.2) using a restarted conjugate gradient method. After each start, a subset of variables is fixed at bounds (the active set) and the conjugate gradient method minimizes f with respect to remaining variables. The "inner" minimization is terminated if either the minimum is reached or an infeasible iteration is generated. In the first case, some of the fixed variables are released while, in the second case, new variables are added to the active set. In both cases, the value of f decreases so that the active set can never reappear, and therefore the algorithm converges in a finite number of steps.

As the Polyak algorithm suffers from several drawbacks, it was modified in order to exclude doubts about its efficiency; see discussions in [1, 5]. Here, we mention two improvements relevant for our more general constraints. Firstly, the exact solution of auxiliary "inner" problems can be replaced by an inexact one. We shall use a theoretically supported strategy of an adaptive precision control presented by Friedlander and Martínez [6] and by Dostál [2]. The basic idea is to control the "inner" precision by a ratio of the norms of a violation of the Karush–Kuhn–Tucker (KKT) conditions at fixed and free variables. Secondly, the qualitative progress has been achieved by Dostál and Schöberl [5] using the projected gradient for expanding the active set. It permits rapid changes in the active set without the necessity to perform computationally expensive steps (e.g., backtracking). Moreover, the algorithm has a linear convergence rate in terms of the spectral condition number of the Hessian matrix A.

The scheme of the later algorithm was successfully extended for solving 3D contact problems with an isotropic friction in [11]. As the constraints are circular (1.3), the algorithm does search not only for the active set corresponding to the solution, but also for the positions of the pairs $(x_{2i-1}, x_{2i})^\top$ lying on the boundaries of the circles $x_{2i-1}^2 + x_{2i}^2 = r_i^2$. It is easily seen that the finite terminating property cannot be expected in such cases, and therefore the convergence was proven in [10] by different arguments. On the other hand, it is relatively surprising that a linear convergence rate may be derived like for the simple bound case. This proof is the main goal of the paper.

Let us briefly outline the structure of the paper. After introducing notations in section 2, we prove that the KKT optimality conditions are equivalent to the zero projected gradient. Our algorithm for solving (1.1) is proposed in section 3 in a form suitable for theoretical analysis. Section 4 summarizes auxiliary statements on the projected gradient, while section 5 gives the main result of the paper concerning a linear convergence rate. A practical implementation of the algorithm for simple bounds (1.2) and circular constraints (1.3) is discussed in section 6, and finally, section 7 presents the results of numerical experiments.

2. Preliminaries and notations. We shall always assume that the feasible set Ω in (1.1) is nonempty. As f is the strictly convex function and Ω is the convex set,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

the existence of a unique solution to (1.1) is guaranteed. We shall denote it by x^*. It is well known that x^* is fully determined by the KKT conditions [12]. Before giving their appropriate form, we shall introduce notations.

Recall that a continuously differentiable function $F : \Omega \to \mathbb{R}$ is convex iff

$$F(y) - F(x) \geq (y - x)^\top \nabla F(x) \quad \forall x, y \in \Omega, \tag{2.1}$$

where ∇F denotes the gradient of F. The gradient of the objective function f at x shall be denoted by

$$g = g(x) = Ax - b.$$

Let $\mathcal{M}$ denote the set of indices so that

$$\mathcal{M} = \{1, \ldots, m\}.$$

We shall use the following convention: if $x \in \mathbb{R}^n$ is a vector, then $x_i \in \mathbb{R}$ is its ith entry, $1 \leq i \leq n$, and $\boldsymbol{x}_i \in \mathbb{R}^{n_i}$ is its ith segment, $i \in \mathcal{M}$, so that $x = (\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_m^\top)^\top$. We shall denote by $\|\boldsymbol{x}_i\|$ the Euclidean norm of $\boldsymbol{x}_i$.

In order to exclude pathological situations, we shall assume without loss of generality that f_i are not identically zero in Ω_i. In that case Ω_i has a nonempty interior, $\operatorname{int}\Omega_i = \{\boldsymbol{x}_i \in \mathbb{R}^{n_i} : f_i(\boldsymbol{x}_i) < 0\}$, and a possibly nonempty boundary $\partial\Omega_i = \{\boldsymbol{x}_i \in \mathbb{R}^{n_i} : f_i(\boldsymbol{x}_i) = 0\}$. The convexity of f_i implies that if $\partial\Omega_i \neq \emptyset$ and $\boldsymbol{x}_i \in \partial\Omega_i$, then $\nabla f_i(\boldsymbol{x}_i)$ is the outward normal vector to $\partial\Omega_i$ at $\boldsymbol{x}_i$; see Figure 2.1.a.

It is well known that the solution x^* to (1.1) is characterized by the existence of Lagrange multipliers λ_i^*, $i \in \mathcal{M}$, such that [12]

$$\boldsymbol{g}_i^* + \lambda_i^* \nabla f_i(\boldsymbol{x}_i^*) = \boldsymbol{0},\ f_i(\boldsymbol{x}_i^*) \leq 0,\ \lambda_i^* \geq 0,\ \lambda_i^* f_i(\boldsymbol{x}_i^*) = 0, \quad i \in \mathcal{M},$$

where $\boldsymbol{g}_i^*$ denotes the ith segment of $g^* = g(x^*)$. After eliminating λ_i^*, we obtain the following theorem.

THEOREM 2.1. *The vector $x^* \in \Omega$ is the solution to* (1.1) *iff for $i \in \mathcal{M}$:*

$$f_i(\boldsymbol{x}_i^*) < 0 \quad \textit{implies} \quad \boldsymbol{g}_i^* = \boldsymbol{0}, \tag{2.2}$$

$$f_i(\boldsymbol{x}_i^*) = 0 \quad \textit{implies} \quad \boldsymbol{g}_i^* + \frac{\|\boldsymbol{g}_i^*\|}{\|\nabla f_i(\boldsymbol{x}_i^*)\|} \nabla f_i(\boldsymbol{x}_i^*) = \boldsymbol{0}. \tag{2.3}$$

Conditions (2.2) and (2.3) are called the *inner KKT conditions* and the *boundary KKT conditions*, respectively.

As the feasible set Ω is separable, the projection $P_\Omega : \mathbb{R}^n \mapsto \Omega$ can be put together by the projections $P_{\Omega_i} : \mathbb{R}^{n_i} \mapsto \Omega_i$, $i \in \mathcal{M}$. Thus we define $P_\Omega(x)$ for any $x \in \mathbb{R}^n$ by

$$P_\Omega(x) = \begin{pmatrix} P_{\Omega_1}(\boldsymbol{x}_1) \\ \vdots \\ P_{\Omega_m}(\boldsymbol{x}_m) \end{pmatrix}, \tag{2.4}$$

where $P_{\Omega_i}(\boldsymbol{x}_i)$ are defined by

$$P_{\Omega_i}(\boldsymbol{x}_i) = \begin{cases} \boldsymbol{x}_i & \text{for } \boldsymbol{x}_i \in \Omega_i, \\ \boldsymbol{z}_i & \text{for } \boldsymbol{x}_i \notin \Omega_i, \end{cases}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

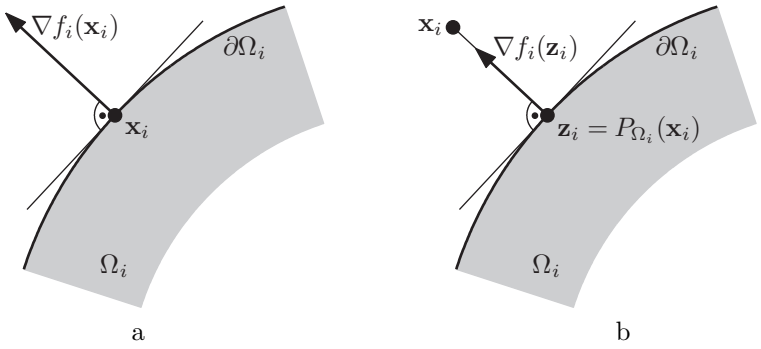

FIG. 2.1. (a) *The outward normal vector to* $\partial\Omega_i$; (b) *The projection to* $\partial\Omega_i$.

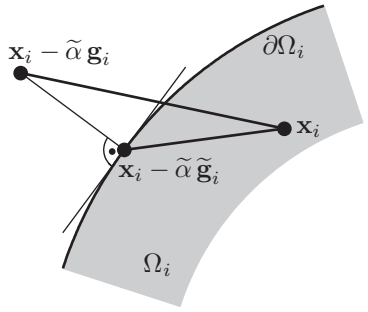

FIG. 2.2. *The projected gradient on* Ω_i.

with (see Figure 2.1.b)

$$f_i(\boldsymbol{z}_i) = 0, \tag{2.5}$$

$$\boldsymbol{z}_i + \frac{\|\boldsymbol{x}_i - \boldsymbol{z}_i\|}{\|\nabla f_i(\boldsymbol{z}_i)\|}\nabla f_i(\boldsymbol{z}_i) = \boldsymbol{x}_i. \tag{2.6}$$

Let us note that $\boldsymbol{z}_i$ is well defined by (2.5) and (2.6) because Ω_i is convex. The equality (2.6) yields orthogonality of P_{Ω_i} and, consequently, of P_Ω that is equivalent to the variational inequality

$$(x - P_\Omega(x))^\top (y - P_\Omega(x)) \le 0 \quad \forall x \in \mathbb{R}^n \ \forall y \in \Omega.$$

Let us define the *projected gradient* $\widetilde{g} = \widetilde{g}(x)$ at $x \in \Omega$ for fixed $\widetilde{\alpha} > 0$ by

$$\widetilde{g}(x) = \frac{1}{\widetilde{\alpha}}\left(x - P_\Omega\left(x - \widetilde{\alpha} g(x)\right)\right). \tag{2.7}$$

This definition enables us to describe the projection of the gradient step without P_Ω so that

$$P_\Omega(x - \widetilde{\alpha} g(x)) = x - \widetilde{\alpha}\widetilde{g}(x);$$

see Figure 2.2. In the next theorem, we prove that the zero project gradient represents an alternative optimality criterion to KKT conditions.

THEOREM 2.2. *The vector* $x^* \in \Omega$ *solves* (1.1) *iff* $\widetilde{g}(x^*) = 0$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof. The equality $\widetilde{g}(x^*) = 0$ is equivalent to

$$\boldsymbol{x}_i^* = P_{\Omega_i}(\boldsymbol{x}_i^* - \widetilde{\alpha}\boldsymbol{g}_i^*), \quad i \in \mathcal{M}. \tag{2.8}$$

We shall prove that the KKT conditions (2.2) and (2.3) are equivalently satisfied. We distinguish two cases. (i) Let $f_i(\boldsymbol{x}_i^*) < 0$. Then $\boldsymbol{x}_i^* \in \operatorname{int}\Omega_i$, which is equivalent by (2.8) to $\boldsymbol{x}_i^* - \widetilde{\alpha}\boldsymbol{g}_i^* \in \operatorname{int}\Omega_i$. Therefore

$$\boldsymbol{x}_i^* = P_{\Omega_i}(\boldsymbol{x}_i^* - \widetilde{\alpha}\boldsymbol{g}_i^*) = \boldsymbol{x}_i^* - \widetilde{\alpha}\boldsymbol{g}_i^*$$

so that $\boldsymbol{g}_i^* = \boldsymbol{0}$, and therefore (2.2) holds. (ii) Let $f_i(\boldsymbol{x}_i^*) = 0$. Using (2.6) with $\boldsymbol{x}_i$ replaced by $\boldsymbol{x}_i^* - \widetilde{\alpha}\boldsymbol{g}_i^*$ and with $\boldsymbol{z}_i$ replaced by $\boldsymbol{x}_i^*$, we obtain after simple manipulation

$$\boldsymbol{g}_i^* + \frac{\|\boldsymbol{g}_i^*\|}{\|\nabla f_i(\boldsymbol{x}_i^*)\|}\nabla f_i(\boldsymbol{x}_i^*) = \boldsymbol{0}$$

so that (2.3) holds. ∎

We shall decompose $\mathcal{M}$ at $x \in \Omega$ on the *free set* $\mathcal{F}(x)$ and the *active set* $\mathcal{A}(x)$ as

$$\mathcal{F}(x) = \{i \in \mathcal{M} : f_i(\boldsymbol{x}_i) < 0\},$$
$$\mathcal{A}(x) = \{i \in \mathcal{M} : f_i(\boldsymbol{x}_i) = 0\}.$$

Analogously, we can decompose $\widetilde{g}$ on the *projected free gradient* $\widetilde{\varphi} = \widetilde{\varphi}(x)$ and the *projected boundary gradient* $\widetilde{\beta} = \widetilde{\beta}(x)$ so that

$$\widetilde{\boldsymbol{\varphi}}_i = \widetilde{\boldsymbol{g}}_i \quad \text{for } i \in \mathcal{F}(x), \quad \widetilde{\boldsymbol{\varphi}}_i = \boldsymbol{0} \quad \text{for } i \in \mathcal{A}(x), \tag{2.9}$$

$$\widetilde{\boldsymbol{\beta}}_i = \boldsymbol{0} \quad \text{for } i \in \mathcal{F}(x), \quad \widetilde{\boldsymbol{\beta}}_i = \widetilde{\boldsymbol{g}}_i \quad \text{for } i \in \mathcal{A}(x). \tag{2.10}$$

Thus, the inner KKT conditions (2.2) are satisfied iff $\widetilde{\varphi}(x) = 0$, and the boundary KKT conditions (2.3) are satisfied iff $\widetilde{\beta}(x) = 0$. Moreover, we define the *free gradient* $\varphi = \varphi(x)$ so that

$$\widetilde{\boldsymbol{\varphi}}_i = \widetilde{\boldsymbol{g}}_i \quad \text{for } i \in \mathcal{F}(x), \quad \widetilde{\boldsymbol{\varphi}}_i = \boldsymbol{0} \quad \text{for } i \in \mathcal{A}(x). \tag{2.11}$$

Finally let us denote the A-energy norm of $x \in \mathbb{R}^n$ by $\|x\|_A$. Thus $\|x\|_A = (x^\top Ax)^{1/2}$, and $\|x\| = \|x\|_I = (x^\top x)^{1/2}$ is the Euclidean norm. The analogous notation will be used for the induced matrix norm so that

$$\kappa(A) = \|A\|\|A^{-1}\|$$

is the spectral condition number of A.

3. Algorithm. In this section we present our algorithm for solving (1.1) in a form convenient for the analysis, while technical details are postponed to section 6. The algorithm exploits a given constant $\Gamma > 0$ to decide on interrupting conjugate gradient iterations and a fixed steplength $\widetilde{\alpha} \in (0, \|A\|^{-1}]$ defining the projected gradient.

We shall combine three steps to generate a sequence of iterates $\{x^k\}$ that approximates the solution to (1.1):

– the *expansion step*, which may add indices to the active set, is defined by

$$x^{k+1} = x^k - \widetilde{\alpha}\widetilde{\varphi}(x^k); \tag{3.1}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

– the *proportioning step*, which may release indices from the active set, reads as

$$x^{k+1} = x^k - \widetilde{\alpha}\widetilde{\beta}(x^k); \tag{3.2}$$

– the *conjugate gradient step* is given by

$$x^{k+1} = x^k - \alpha_{cg}^k p^k, \quad \alpha_{cg}^k = \frac{g(x^k)^\top p^k}{(p^k)^\top A p^k}, \tag{3.3}$$

where the conjugate gradient directions p^k are constructed recurrently [7].

The conjugate gradient steps are used to carry out a minimization of the objective function f efficiently on the interior of the *face*

$$W_{\mathcal{A}(x^s)} = \{x \in \Omega : \boldsymbol{x}_i = \boldsymbol{x}_i^s \ \text{ for } \ i \in \mathcal{A}(x^s)\},$$

where x^s is determined by the expansion step or by the proportioning step. It requires that the parts of p^k corresponding to the indices of $\mathcal{A}(x^s)$ vanish, i.e., $\boldsymbol{p}_i^k = \mathbf{0}$ for $i \in \mathcal{A}(x^s)$. We can easily fulfill this requirement by adapting the classical recurrence generating p^k. We start (or restart) from $p^s = \varphi(x^s)$ and use

$$p^k = \varphi(x^k) - \gamma^k p^{k-1}, \quad \gamma^k = \frac{\varphi(x^k)^\top A p^{k-1}}{(p^{k-1})^\top A p^{k-1}}, \quad k > s. \tag{3.4}$$

The formulae (3.4) is used while the sequence of the conjugate gradient steps is unbroken. After changing the active set, we must always restart.

Later on, we shall need the following lemma.

LEMMA 3.1. *Let x^{k+1} be generated by the conjugate gradient step. Then*

$$f(x^{k+1}) \le f(x^k - \alpha\varphi(x^k)) \quad \forall \alpha \in \mathbb{R}.$$

Proof. It is a well-known property of the conjugate gradient method [7] that

$$f(x^{k+1}) = \min_{x \in x^s + \mathrm{Span}\{p^s, \dots, p^k\}} f(x). \tag{3.5}$$

As (3.4) implies $\varphi(x^k) \in \mathrm{Span}\{p^s, \dots, p^k\}$ and $x^k \in x^s + \mathrm{Span}\{p^s, \dots, p^{k-1}\}$, we obtain $x^k - \alpha\varphi(x^k) \in x^s + \mathrm{Span}\{p^s, \dots, p^k\}$. The lemma follows using (3.5). ∎

The last ingredient of our algorithm is the *releasing criterion*:

$$\widetilde{\beta}(x^k)^\top g(x^k) \le \Gamma^2 \widetilde{\varphi}(x^k)^\top g(x^k). \tag{3.6}$$

If this inequality holds, we call the iterate x^k *strictly proportional*. The criterion (3.6) is used to decide which of the steps will be performed.

ALGORITHM 3.1. *Let $x^0 \in \Omega$, $\Gamma > 0$, and $\widetilde{\alpha} \in (0, \|A\|^{-1}]$ be given. For $k \ge 0$ and x^k known, choose x^{k+1} by the following rules:*

(i) *If $\widetilde{g}(x^k) = 0$, set $x^{k+1} = x^k$.*

(ii) *If x^k is strictly proportional and $\widetilde{g}(x^k) \ne 0$, try to generate x^{k+1} by the conjugate gradient step. If $x^{k+1} \in \Omega$ and if it does not change the active set, then accept it, else generate x^{k+1} by the expansion step.*

(iii) *If x^k is not strictly proportional, then define x^{k+1} by the proportioning step.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

4. Properties of the projected gradient. In this section we summarize results necessary for the next analysis.

LEMMA 4.1. *It holds that*

$$\|\widetilde{\varphi}(x)\|^2 \le \widetilde{\varphi}(x)^\top g(x), \tag{4.1}$$

$$\|\widetilde{\beta}(x)\|^2 \le \widetilde{\beta}(x)^\top g(x), \tag{4.2}$$

$$\widetilde{\varphi}(x)^\top g(x) \le \varphi(x)^\top g(x). \tag{4.3}$$

Proof. The definition of the projected gradient (2.7) implies that

$$P_{\Omega_i}(\boldsymbol{x}_i - \widetilde{\alpha}\boldsymbol{g}_i) = \boldsymbol{x}_i - \widetilde{\alpha}\widetilde{\boldsymbol{g}}_i.$$

As P_{Ω_i} is orthogonal, we obtain

$$0 \ge (\boldsymbol{x}_i - \widetilde{\alpha}\boldsymbol{g}_i - P_{\Omega_i}(\boldsymbol{x}_i - \widetilde{\alpha}\boldsymbol{g}_i))^\top (\boldsymbol{x}_i - P_{\Omega_i}(\boldsymbol{x}_i - \widetilde{\alpha}\boldsymbol{g}_i)) = \widetilde{\alpha}^2(\widetilde{\boldsymbol{g}}_i - \boldsymbol{g}_i)^\top \widetilde{\boldsymbol{g}}_i$$

so that

$$\|\widetilde{\boldsymbol{g}}_i\|^2 \le \widetilde{\boldsymbol{g}}_i^\top \boldsymbol{g}_i. \tag{4.4}$$

Summing (4.4) over the indices of $\mathcal{F}(x)$ or $\mathcal{A}(x)$, we obtain (4.1) or (4.2), respectively. Using the well-known Cauchy inequality in (4.4), we get $\|\widetilde{\boldsymbol{g}}_i\| \le \|\boldsymbol{g}_i\|$, and therefore

$$\widetilde{\varphi}(x)^\top g(x) = \sum_{i\in\mathcal{F}(x)} \widetilde{\boldsymbol{g}}_i^\top \boldsymbol{g}_i \le \sum_{i\in\mathcal{F}(x)} \|\widetilde{\boldsymbol{g}}_i\|\|\boldsymbol{g}_i\| \le \sum_{i\in\mathcal{F}(x)} \boldsymbol{g}_i^\top \boldsymbol{g}_i = \varphi(x)^\top g(x). \qquad \square$$

LEMMA 4.2. *Let $d \in \mathbb{R}^n$, $d \ne 0$, and $x \in \Omega$. Then*

$$f(x) - f(x - \alpha d) \le \alpha d^\top g(x) \quad \forall \alpha \in \mathbb{R}. \tag{4.5}$$

Moreover, let $d^\top g(x) \ge 0$, and $\alpha_d = d^\top g(x)/d^\top A d$. Then

$$f(x) - f(x - \alpha d) \ge \frac{1}{2}\alpha d^\top g(x) \quad \forall \alpha \in [0, \alpha_d]. \tag{4.6}$$

Proof. The assertion (4.5) is equivalent to the convexity of f (compare with (2.1)). For $\alpha \in [0, \alpha_d]$, we derive

$$\begin{aligned} f(x - \alpha d) &= f(x) - \alpha d^\top g(x) + \frac{1}{2}\alpha^2 d^\top A d \\ &\le f(x) - \alpha d^\top g(x) + \frac{1}{2}\alpha\alpha_d d^\top A d \\ &= f(x) - \frac{1}{2}\alpha d^\top g(x). \qquad \square \end{aligned}$$

COROLLARY 4.3. *If d is replaced in (4.6) by φ, $\widetilde{\varphi}$, and $\widetilde{\beta}$, then the corresponding three inequalities hold for all $\alpha \in [0, \|A\|^{-1}]$.*

Proof. For $d = \varphi(x)$, we have

$$\alpha_d = \frac{\varphi(x)^\top g(x)}{\varphi(x)^\top A \varphi(x)} \ge \frac{\varphi(x)^\top \varphi(x)}{\|A\|\|\varphi(x)\|^2} = \|A\|^{-1}.$$

The same follows using Lemma 4.1 for $d = \widetilde{\varphi}(x)$ and $d = \widetilde{\beta}(x)$. $\square$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

LEMMA 4.4. *Let $x \in \Omega$. Then*

$$f(x) - f(x - \alpha\widetilde{g}(x)) \leq \alpha(\widetilde{\varphi}(x)^\top g(x) + \widetilde{\beta}(x)^\top g(x)) \quad \forall \alpha \in \mathbb{R}.$$

Proof. Using (4.5), we obtain

$$f(x) - f(x - \alpha\widetilde{g}(x)) \leq \alpha\widetilde{g}(x)^\top g(x) = \alpha \left(\sum_{i \in \mathcal{F}(x)} \widetilde{\boldsymbol{g}}_i^\top \boldsymbol{g}_i + \sum_{i \in \mathcal{A}(x)} \widetilde{\boldsymbol{g}}_i^\top \boldsymbol{g}_i \right),$$

and the definitions (2.9) and (2.10) complete the proof. ∎

LEMMA 4.5. *Let $x^* \in \Omega$ denote the solution to* (1.1)*, λ_1 denote the smallest eigenvalue of A, and $x \in \Omega$. Then*

$$f(x) - f(x - \widetilde{\alpha}\widetilde{g}(x)) \geq \widetilde{\alpha}\lambda_1(f(x) - f(x^*)) \quad \forall \widetilde{\alpha} \in [0, \|A\|^{-1}].$$

Proof. For $\widetilde{\alpha} \in (0, \|A\|^{-1}]$, we define

$$F(y) = \widetilde{\alpha} f(y) + \frac{1}{2}(y - x)^\top (I - \widetilde{\alpha}A)(y - x).$$

Notice that $I - \widetilde{\alpha}A$ is positive semidefinite. Therefore

$$F(y) \geq \widetilde{\alpha} f(y),$$

$$\nabla F(y) = \widetilde{\alpha} g(y) + (I - \widetilde{\alpha}A)(y - x) = y - x + \widetilde{\alpha} g(x), \tag{4.7}$$

and $F(x) = \widetilde{\alpha} f(x)$, $\nabla F(x) = \widetilde{\alpha} g(x)$. Let us note that the Hessian of F is the identity I. Let $y \in \Omega$ be arbitrary. The convexity of F, (4.7), and the orthogonality of P_Ω yields

$$\begin{aligned} F(y) - F(P_\Omega(x - \widetilde{\alpha} g(x))) &\geq (y - P_\Omega(x - \widetilde{\alpha} g(x)))^\top \nabla F(P_\Omega(x - \widetilde{\alpha} g(x))) \\ &= -(y - P_\Omega(x - \widetilde{\alpha} g(x)))^\top (x - \widetilde{\alpha} g(x) - P_\Omega(x - \widetilde{\alpha} g(x))) \\ &\geq 0 \end{aligned}$$

so that

$$F(y) \geq F(P_\Omega(x - \widetilde{\alpha} g(x))) \quad \forall y \in \Omega.$$

Using $\widetilde{\alpha}\lambda_1 \leq \|A\|^{-1}\lambda_1 \leq 1$ and denoting $d = x^* - x$, we derive

$$\begin{aligned} f(x - \widetilde{\alpha}\widetilde{g}(x)) &= f(P_\Omega(x - \widetilde{\alpha} g(x))) \\ &\leq \widetilde{\alpha}^{-1} F(P_\Omega(x - \widetilde{\alpha} g(x))) \;\leq\; \widetilde{\alpha}^{-1} \min_{y \in \Omega} F(y) \;\leq\; \widetilde{\alpha}^{-1} \min_{t \in [0,1]} F(x + td) \\ &\leq \widetilde{\alpha}^{-1} F(x + \widetilde{\alpha}\lambda_1 d) \;=\; f(x) + \widetilde{\alpha}\lambda_1 d^\top g(x) + \frac{1}{2}\widetilde{\alpha}\lambda_1^2 d^\top d \\ &\leq (1 - \widetilde{\alpha}\lambda_1) f(x) + \widetilde{\alpha}\lambda_1 f(x) + \widetilde{\alpha}\lambda_1 d^\top g(x) + \frac{1}{2}\widetilde{\alpha}\lambda_1 d^\top A d \\ &= (1 - \widetilde{\alpha}\lambda_1) f(x) + \widetilde{\alpha}\lambda_1 f(x^*). \end{aligned}$$ ∎

Lemma 4.4 and Lemma 4.5 immediately imply the following result.

COROLLARY 4.6. *Let $x^* \in \Omega$ denote the solution to* (1.1)*, λ_1 denote the smallest eigenvalue of A, and $x \in \Omega$. Then*

$$\lambda_1(f(x) - f(x^*)) \leq \widetilde{\varphi}(x)^\top g(x) + \widetilde{\beta}(x)^\top g(x)$$

for all $\widetilde{\alpha} \in (0, \|A\|^{-1}]$ defining $\widetilde{\varphi}$ and $\widetilde{\beta}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

5. Rate of convergence.

THEOREM 5.1. *Let $x^0 \in \Omega$, $\Gamma > 0$, and $\widetilde{\alpha} \in (0, \|A\|^{-1}]$ be given. Let $x^* \in \Omega$ denote the solution to (1.1), λ_1 denote the smallest eigenvalue of A, and $\widehat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}$. Let $\{x^k\}$ be the sequence generated by Algorithm 3.1. Then*

$$f(x^{k+1}) - f(x^*) \le \eta \left(f(x^k) - f(x^*) \right), \tag{5.1}$$

where

$$\eta = 1 - \frac{\widetilde{\alpha}\lambda_1}{2 + 2\widehat{\Gamma}^2} < 1. \tag{5.2}$$

The error in the A-energy norm is bounded by

$$\|x^k - x^*\|_A^2 \le 2\eta^k \left(f(x^0) - f(x^*) \right). \tag{5.3}$$

Proof. We shall estimate separately all three possible steps of Algorithm 3.1. As $\widetilde{\alpha} \in (0, \|A\|^{-1}]$, we can use the bound (4.6) for each of the steps due to Corollary 4.3.

Let us first assume that x^{k+1} is generated by the expansion step (3.1). Using (4.6), we obtain

$$f(x^{k+1}) = f(x^k - \widetilde{\alpha}\widetilde{\varphi}(x^k)) \le f(x^k) - \frac{1}{2}\,\widetilde{\alpha}\widetilde{\varphi}(x^k)^\top g(x^k). \tag{5.4}$$

If x^{k+1} is generated by the conjugate gradient step (3.3), we use Lemma 3.1, (4.6), and (4.3) so that

$$\begin{aligned} f(x^{k+1}) &\le f(x^k - \widetilde{\alpha}\varphi(x^k)) \\ &\le f(x^k) - \frac{1}{2}\,\widetilde{\alpha}\varphi(x^k)^\top g(x^k) \le f(x^k) - \frac{1}{2}\,\widetilde{\alpha}\widetilde{\varphi}(x^k)^\top g(x^k). \end{aligned} \tag{5.5}$$

Comparing (5.4) and (5.5), we may observe that we have the same estimates for both of the expansion and conjugate gradient steps. These steps are taken only when x^k is strictly proportional, i.e., when

$$\widetilde{\beta}(x^k)^\top g(x^k) \le \Gamma^2 \widetilde{\varphi}(x^k)^\top g(x^k).$$

After using Corollary 4.6, we get

$$\widetilde{\varphi}(x^k)^\top g(x^k) \ge \frac{\lambda_1}{1 + \Gamma^2}(f(x^k) - f(x^*)). \tag{5.6}$$

The estimates (5.4), (5.5) combined with (5.6) imply that

$$\begin{aligned} f(x^{k+1}) - f(x^*) &\le f(x^k) - f(x^*) - \frac{1}{2}\,\widetilde{\alpha}\widetilde{\varphi}(x^k)^\top g(x^k) \\ &\le \left(1 - \frac{\widetilde{\alpha}\lambda_1}{2 + 2\Gamma^2}\right)(f(x^k) - f(x^*)). \end{aligned} \tag{5.7}$$

Let us finally assume that x^{k+1} is generated by the proportioning step (3.2). Again using (4.6), we obtain

$$f(x^{k+1}) = f(x^k - \widetilde{\alpha}\widetilde{\beta}(x^k)) \le f(x^k) - \frac{1}{2}\,\widetilde{\alpha}\widetilde{\beta}(x^k)^\top g(x^k). \tag{5.8}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

As the proportioning step is taken when

$$\widetilde{\beta}(x^k)^\top g(x^k) > \Gamma^2 \widetilde{\varphi}(x^k)^\top g(x^k),$$

Corollary 4.6 yields

$$\widetilde{\beta}(x^k)^\top g(x^k) \geq \frac{\lambda_1}{1+\Gamma^{-2}}(f(x^k) - f(x^*)). \tag{5.9}$$

The estimate (5.8) combined with (5.9) implies that

$$f(x^{k+1}) - f(x^*) \leq \left(1 - \frac{\widetilde{\alpha}\lambda_1}{2+2\Gamma^{-2}}\right)(f(x^k) - f(x^*)). \tag{5.10}$$

Comparing the inequalities (5.10) and (5.7), and taking into account that by definition $\Gamma \leq \widehat{\Gamma}$ and $\Gamma^{-1} \leq \widehat{\Gamma}$, we can see that (5.1) holds.

In order to prove the error bound (5.3), we use (5.1) and the fact that the solution x^* to (1.1) is characterized by the variational inequality $(x - x^*)^\top g(x^*) \geq 0$ for all $x \in \Omega$. We obtain

$$\begin{aligned} \|x^k - x^*\|_A^2 &= 2(f(x^k) - f(x^*) - (x^k - x^*)^\top g(x^*)) \\ &\leq 2(f(x^k) - f(x^*)) \ \leq \ 2\eta^k(f(x^0) - f(x^*)). \qquad \square \end{aligned}$$

Theorem 5.1 yields the best estimate for $\Gamma = \widehat{\Gamma} = 1$ and $\widetilde{\alpha} = \|A\|^{-1}$ when

$$\eta = 1 - \frac{1}{4}\kappa(A)^{-1}.$$

We shall see that this result is in agreement with numerical experiments.

6. Implementation. We shall give the details of the implementation of Algorithm 3.1. Firstly, we describe a general scheme of the algorithm independently on the type of constraints. Then we show how to compute projections on the feasible set in our numerical experiments.

6.1. A general algorithmic scheme. The presented implementation differs from Algorithm 3.1 in that it exploits the current conjugate gradient direction to generate an intermediate iteration $x^{k+1/2}$ before the expansion step that generates x^{k+1} from $x^{k+1/2}$. Such modification does not require any additional matrix-vector multiplication, and the estimate (5.1) remains valid [5].

We describe the algorithm in an easily understandable variant of the Matlab language, in which we do not distinguish a generation of variables by indices unless it is convenient for further references.

It should be noted that the presented implementation is not influenced by the type of constraints defining the feasible set Ω. The relation to constraints is hidden in the components $\widetilde{\beta}$, $\widetilde{\varphi}$ of the projected gradient $\widetilde{g}$, and in the feasible steplength α_f. The projected gradient can be directly evaluated by its definition (2.7) using P_Ω.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

ALGORITHM 6.1. *Let $x^0 \in \Omega$, $\Gamma > 0$, $\widetilde{\alpha} \in (0, \|A\|^{-1}]$, and $\epsilon > 0$ be given.*

Set $k = 0$, $g = Ax^0 - b$, $p = \varphi(x^0)$. % *Initialization.*
while $\|\widetilde{g}(x^k)\| > \epsilon$
  if $\widetilde{\beta}(x^k)^\top g(x^k) \le \Gamma^2 \widetilde{\varphi}(x^k)^\top g(x^k)$
    $\alpha_{cg} = g^\top p / p^\top A p$ % *Conjugate gradient steplength.*
    $\alpha_f = \max\{\alpha : x^k - \alpha p \in \Omega\}$ % *Feasible steplength.*
    if $\alpha_{cg} < \alpha_f$ % *Conjugate gradient step.*
      $x^{k+1} = x^k - \alpha_{cg} p$, $g = g - \alpha_{cg} A p$, $\gamma = \varphi(x^{k+1})^\top A p / p^\top A p$, $p = \varphi(x^{k+1}) - \gamma p$
    else % *Expansion step.*
      $x^{k+1/2} = x^k - \alpha_f p$
      $x^{k+1} = x^{k+1/2} - \widetilde{\alpha}\widetilde{\varphi}(x^{k+1/2})$, $g = Ax^{k+1} - b$, $p = \varphi(x^{k+1})$
    end*if*
  else % *Proportioning step.*
    $x^{k+1} = x^k - \widetilde{\alpha}\widetilde{\beta}(x^k)$, $g = Ax^{k+1} - b$, $p = \varphi(x^{k+1})$
  end*if*
  $k = k + 1$
end*while*
$x = x^k$. % *Return step.*

6.2. Projections and the feasible steplength. In the previous sections, we have exploited the implicit definition of the projections given by (2.5) and (2.6). Now we shall show how to compute P_{Ω_i} for simple bounds (1.2) and circular constraints (1.3).

Let us consider the following variant of the problem (1.1):

$$\left\{ \begin{array}{ll} \text{minimize} & f(x), \\ \text{subject to} & x_i \ge l_i,\ i = 1, \dots, m_1, \\ & x_{m_1+2i-1}^2 + x_{m_1+2i}^2 \le r_i^2,\ i = 1, \dots, m_2, \end{array} \right. \tag{6.1}$$

where $m_1 + 2m_2 < n$ and l_i, r_i are given. The feasible set Ω in (6.1) is described by Ω_i, $i = 1, \dots, m$, $m = m_1 + m_2 + 1$, of three different types. Recall that each of Ω_i is defined by $\Omega_i = \{\boldsymbol{x}_i \in \mathbb{R}^{n_i} : f_i(\boldsymbol{x}_i) \le 0\}$ where $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ is continuously differentiable and convex. Let us denote the feasible steplength with respect to the ith constraint by

$$\alpha_{f,i} = \max\{\alpha : \boldsymbol{x}_i - \alpha \boldsymbol{p}_i \in \Omega_i\}.$$

(1) Let $n_i = 1$, and $f_i(x_i) \equiv l_i - x_i$. Then $x_i \in \Omega_i$, $i = 1, \dots, m_1$ represent the simple bounds in (6.1). The set Ω_i describes the half-line so that the projection is given by

$$P_{\Omega_i}(x_i) = \left\{ \begin{array}{ll} x_i & \text{for } l_i \le x_i, \\ l_i & \text{for } l_i > x_i \end{array} \right.$$

and the feasible steplength by

$$\alpha_{f,i} = \left\{ \begin{array}{ll} (x_i - l_i)/p_i & \text{for } p_i > 0, \\ \infty & \text{for } p_i \le 0. \end{array} \right.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

(2) Let $n_{m_1+i} = 2$, $f_{m_1+i}(\boldsymbol{x}_{m_1+i}) \equiv \boldsymbol{x}_{m_1+i}^\top \boldsymbol{x}_{m_1+i} - r_i^2$, and $\boldsymbol{x}_{m_1+i} = (x_{m_1+2i-1}, x_{m_1+2i})^\top$. Then $\boldsymbol{x}_{m_1+i} \in \Omega_{m_1+i}$, $i = 1, \ldots, m_2$ represent the circular constraints in (6.1). The set Ω_{m_1+i} describes the circle with the center at the origin of $\mathbb{R}^2$ and with the radius r_i so that the projection is given by

$$P_{\Omega_{m_1+i}}(\boldsymbol{x}_{m_1+i}) = \begin{cases} \boldsymbol{x}_{m_1+i} & \text{for } \|\boldsymbol{x}_{m_1+i}\| \le r_i, \\ \dfrac{r_i}{\|\boldsymbol{x}_{m_1+i}\|}\, \boldsymbol{x}_{m_1+i} & \text{for } \|\boldsymbol{x}_{m_1+i}\| > r_i \end{cases}$$

and the feasible steplength by

$$\alpha_{f,m_1+i} = \begin{cases} (\boldsymbol{x}_i^\top \boldsymbol{p}_i + \sqrt{(\boldsymbol{x}_i^\top \boldsymbol{p}_i)^2 - (\|\boldsymbol{x}_i\|^2 - r_i^2)\|\boldsymbol{p}_i\|^2}\,)/\|\boldsymbol{p}_i\|^2 & \text{for } \boldsymbol{p}_i \neq 0, \\ \infty & \text{for } \boldsymbol{p}_i = 0. \end{cases}$$

(3) Let $n_m = n - m_1 - 2m_2$, $\boldsymbol{x}_m = (x_{m_1+2m_2+1}, \ldots, x_n)^\top$, and $f_m(\boldsymbol{x}_m) = -1$. Then $\Omega_m = \mathbb{R}^{n_m}$, i.e., the components of $\boldsymbol{x}_m$ are unconstrained. For completeness, we define

$$P_{\Omega_m}(\boldsymbol{x}_m) = \boldsymbol{x}_m \quad \text{and} \quad \alpha_{f,m} = \infty.$$

Let us conclude that the projection P_Ω is put together by P_{Ω_i} as in (2.4) and the feasible steplength α_f is given by

$$\alpha_f = \min\{\alpha_{f,i} : i = 1, \ldots, m\}.$$

7. Numerical tests. We shall assess the performance of the algorithm by three examples. The first one is the benchmark of [10], in which only circular constraints occur. The second example represents a one-dimensional obstacle problem comprising both simple bounds and circular constraints. A more realistic third example shows the solution of frictional 3D contact problems of linear elasticity by means of a sequence of problems (6.1).

Let us note that either we shall use Algorithm 6.1 with $x^0 = 0$, $\Gamma = 1$, $\widetilde{\alpha} = \|A\|^{-1}$, and $\epsilon = 10^{-5}\|b\|$, or we shall comment different choices.

Example 7.1. Let us consider the problem (6.1) for $n = 12$, $m_1 = 0$, and $m_2 = 6$ with the five-diagonal matrix A:

$$\begin{aligned} &A = (a_{ij}),\ a_{ii} = 4,\ a_{ii\pm1} = a_{ii\pm2} = -1, \\ &b = Ay, \\ &r = (2, 1, 0.5, 2, 10^{-3}, 154)^\top, \end{aligned}$$

and $y = (2, 1, 0.5, 0, 0, 11, 10^{-5}, -1, \sqrt{2}, -0.1, 4.1 \cdot 10^{-4}, 143)^\top$. The solution $x^* \in \mathbb{R}^{12}$ has three active constraints so that $\mathcal{A}(x^*) = \{2, 3, 5\}$. Here, we shall denote Algorithm 6.1 by QPC and the algorithm of [10] by QPQ. In Table 7.1 we compare QPQ and QPC by the number of matrix-vector multiplications for various Γ. The algorithm QPC has better performance in all cases. Moreover, the experiments are in agreement with the conclusions of Theorem 5.1.

In order to explain the progress, we mention the ideas of QPQ. This algorithm is constructed by the KKT conditions that are given here in Theorem 2.1. Introducing the *turned boundary gradient* $\beta = \beta(x)$ at $x \in \Omega$ as

$$\boldsymbol{\beta}_i = \boldsymbol{0} \quad \text{for } i \in \mathcal{F}(x), \quad \boldsymbol{\beta}_i = \boldsymbol{g}_i + \frac{\|\boldsymbol{g}_i\|}{\|\nabla f_i(\boldsymbol{x}_i)\|} \nabla f_i(\boldsymbol{x}_i) \quad \text{for } i \in \mathcal{A}(x),$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

TABLE 7.1
Comparisons of the algorithms QPQ and QPC.

Γ	100	10	5	1	0.5	0.4	0.2	0.1	0.05	0.01	0.001
QPQ	43	41	39	32	37	40	48	41	38	36	36
QPC	30	24	21	19	20	20	22	22	26	37	50

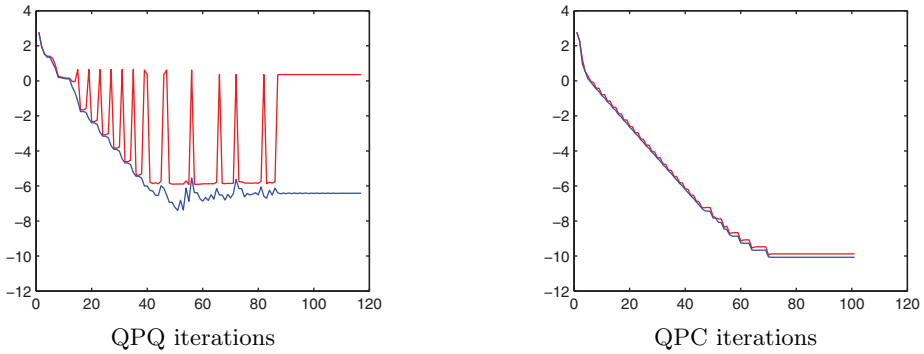

FIG. 7.1. *The iteration history of QPQ and QPC (logarithmic scale).*

we can define the *turned gradient* by $\nu = \nu(x) = \varphi(x) + \beta(x)$ where $\varphi(x)$ is the free gradient (2.9). The solution x^* to the problem (1.1) is then characterized by $\nu(x^*) = 0$. The proportioning step of QPC is replaced in QPQ by the step deactivation, in which β replaces $\tilde{\beta}$ (for a different steplength). It is easy to show that the deactivation step releases all indices from the current active set, for which the boundary KKT conditions (2.3) are not satisfied. After each release, the inner KKT conditions are violated usually on the same level. This observation is a practical consequence of the fact that the functions $\beta(x)$ and $\nu(x)$ are discontinuous (in contrast to $\tilde{\beta}$ and $\tilde{g}$). The typical situation is drawn in Figure 7.1 (left), where the norm of the projected gradient (solid) and the turned gradient (dotted) are depicted. In Figure 7.1 (right) we can see that oscillations arising in QPQ are eliminated in QPC. Comparing the stagnation levels, we can conclude that QPC is considerably more robust. Notice that the turned gradient did not curiously recognize the solution in QPQ due to round-off errors.

Example 7.2. In the second example, we solve

$$\text{minimize } \frac{1}{2}\int_0^1 \|\boldsymbol{x}'(t)\|^2\,dt - \int_0^1 \boldsymbol{x}(t)^\top \boldsymbol{f}(t)\,dt$$

subject to $\boldsymbol{x} = (X_1, X_2)^\top \in \mathcal{K}$, where

$$\mathcal{K} = \{\boldsymbol{x} \in (H_0^1(0,1))^2 :\ X_2(t) \geq l \text{ on } (0,0.5),\ \|\boldsymbol{x}(t)\| \leq r \text{ on } (0.5,1)\},$$

and $\boldsymbol{f}(t) = (36\pi^2 \sin 6\pi t, -4\pi^2 \sin 2\pi t)^\top$. This problem describes the loaded wire (see Figure 7.2) that is partially above the plan far off the distance l and partially inside the cylindrical tube of the radius r. A finite element discretization on a regular grid with n degrees of freedom leads to the problem (6.1) where $m_1 = m_2 = n/4$, $l_i = l$, and $r_i = r$. In tables below we summarize numbers of matrix-vector multiplications and information on active and free constraints as

$$n_{b,\mathcal{A}} : n_{b,\mathcal{F}} / n_{c,\mathcal{A}} : n_{c,\mathcal{F}},$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

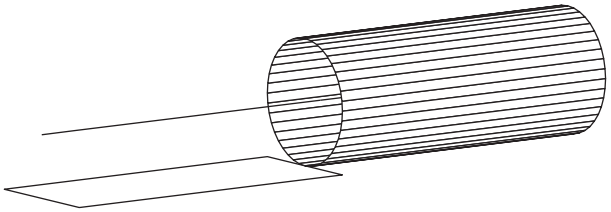

FIG. 7.2. *Geometry of the wire.*

TABLE 7.2
$r = 2$, only simple bounds are active in the solution.

n	$l = -1.5$	$l = -1$	$l = -0.8$	$l = -0.5$	$l = -0.1$	$l = 0$
	0:8/0:8	1:7/0:8	1:7/0:8	3:5/0:8	4:4/0:8	5:3/0:8
32	3	16	26	24	24	24
	0:16/0:16	1:15/0:16	1:15/0:16	4:12/0:16	6:10/0:16	9:7/0:16
64	3	25	57	57	72	62
	0:32/0:32	1:31/0:32	3:29/0:32	6:26/0:32	12:20//0:32	19:13/0:32
128	3	41	92	137	167	112
	0:64/0:64	1:63/0:64	4:60/0:64	10:54/0:64	23:41/0:64	37:27/0:64
256	3	68	262	307	369	269
	0:128/0:128	1:127/0:128	8:120/0:128	21:107/0:128	45:83/0:128	73:55/0:128
512	3	4	494	556	876	710
	0:256/0:256	1:255/0:256	16:240/0:256	40:216/0:256	89:167/0:256	146:110/0:256
1024	3	4	1305	1641	1966	1530

TABLE 7.3
$l = 0$, both simple bounds and circular constraints are active in the solution.

n	$r = 1.4$	$r = 1$	$r = 0.5$	$r = 0.3$	$r = 0.01$	$r = 0.001$
	5:3/2:6	6:2/4:4	7:1/5:3	7:1/5:3	8:0/8:0	8:0/8:0
32	89	80	54	42	17	19
	10:6/2:14	11:5/5:11	13:3/6:10	14:2/9:7	16:0/16:0	16:0/16:0
64	205	240	132	90	23	27
	20:12/4:28	22:10/5:27	26:6/10:22	29:3/16:16	32:0//31:1	32:0/32:0
128	608	620	324	239	96	41
	39:25/4:60	45:19/8:56	52:12/18:46	57:7/26:38	64:0/58:6	64:0/64:0
256	1677	1764	951	695	215	121
	77:51/4:124	89:39/12:116	104:24/33:95	114:14/49:79	127:1/111:17	128:0/126:2
512	4117	5248	3755	1960	646	309
	155:101/6:250	177:79/22:234	208:48/60:196	228:28/95:161	253:3/219:37	256:0/249:7
1024	12084	16851	10516	7065	2051	793

where $n_{b,\mathcal{A}}$, $n_{b,\mathcal{F}}$, $n_{c,\mathcal{A}}$, and $n_{c,\mathcal{F}}$ are numbers of active simple bounds, free simple bounds, active circular constraints, and free circular constraints, respectively.

(i) Let $r = 2$ so that no circular constraint is active in the solution; see Table 7.2. In this case, Algorithm 6.1 reduces to a variant of the algorithm of [5] with the finite terminating property. Let us note that the column labeled $l = -1.5$ corresponds to the unconstrained problem, in which our algorithm became the conjugate gradient method.

(ii) Let $l = 0$ so that both simple bounds and circular constraints may be active in the solution; see Table 7.3. It turns out that the algorithm is more efficient in situations when the constraints are tighter, i.e., when r and l are near zero so that the number of active constraints is considerably higher than the number of free constraints.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

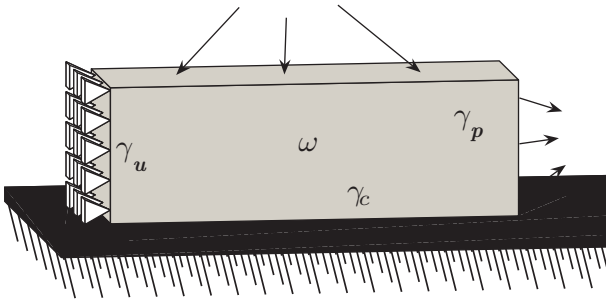

FIG. 7.3. *Geometry of the brick.*

Example 7.3. Let us consider the steel brick lying on a rigid foundation as it is shown in Figure 7.3. The brick occupies in the reference configuration the domain $\omega \subset \mathbb{R}^3$, whose boundary $\partial\omega$ is split into three nonempty disjoint parts γ_u, γ_p, and γ_c with different boundary conditions. The zero displacements are prescribed on γ_u, whereas the surface tractions act on γ_p. On γ_c, we consider the contact conditions, i.e., the nonpenetration and the effect of friction. The elastic behavior of the brick is described by Lamé equations that, after finite element discretization, lead to a symmetric positive definite stiffness matrix $K \in \mathbb{R}^{3n_c \times 3n_c}$ and to a load vector $f \in \mathbb{R}^{3n_c}$. Moreover, we introduce full rank matrices $N, T_1, T_2 \in \mathbb{R}^{m_c \times 3n_c}$ projecting displacements at contact nodes to normal and tangential directions, respectively, and we denote $B = \left(N^\top, T_1^\top, T_2^\top\right)^\top \in \mathbb{R}^{3m_c \times 3n_c}$. For more details about this model problem, we refer to [8]. Here, we shall use the dual formulation in terms of contact stresses.

We start with the contact problem with *Tresca friction* that reads as

$$\text{(7.1)} \qquad \begin{cases} \text{minimize} & \frac{1}{2}\lambda^\top Q\lambda - \lambda^\top h, \\ \text{subject to} & \lambda_{\nu,i} \geq 0,\ \lambda_{t_1,i}^2 + \lambda_{t_2,i}^2 \leq r_i^2,\ i = 1, \ldots, m_c, \\ & \lambda = (\lambda_\nu^\top, \lambda_{t_1}^\top, \lambda_{t_2}^\top)^\top,\ \lambda_\nu, \lambda_{t_1}, \lambda_{t_2} \in \mathbb{R}^{m_c}, \end{cases}$$

where $Q = BK^{-1}B^\top$, $h = BK^{-1}f$, and $r_i \geq 0$ are given slip bound values at contact nodes. Let us point out that λ_ν and $\lambda_{t_1}, \lambda_{t_2}$ represent normal and tangential contact stresses, respectively. It should be noted that the problem (7.1) can be solved by Algorithm 6.1 after rearranging the unknowns. In order to simplify notations, we denote $\mathbb{R}^{m_c}_+ = \{s \in \mathbb{R}^{m_c} : s_i \geq 0\}$.

The contact problem with *Coulomb friction* uses the friction law, in which the slip bound $r \in \mathbb{R}^{m_c}_+$ depends on the normal contact stress $\lambda_\nu \in \mathbb{R}^{m_c}_+$ by

$$r \equiv F\lambda_\nu,$$

where $F > 0$ is a coefficient of friction. As r is the input for (7.1) and λ_ν is the output, the problem with Tresca friction defines the mapping

$$\Psi : \mathbb{R}^{m_c}_+ \longmapsto \mathbb{R}^{m_c}_+ : r \longmapsto F\lambda_\nu.$$

It is easily seen that a fixed point of Ψ solves the problem with Coulomb friction, i.e., a point r such that $\Psi(r) = r$. Notice that Ψ is contractive for sufficiently small F,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

TABLE 7.4
Contact problem with Coulomb friction.

dof		$F = 0.3$				$F = 0.6$			
$3n_c$	$3m_c$	*Time*	*Iter*	n_Q	n_Q/n	*Time*	*Iter*	n_Q	n_Q/n
900	180	4	5	**535**	2.97	6	7	**801**	4.45
2646	378	24	5	**638**	1.68	35	6	**906**	2.40
5832	648	104	5	**758**	1.17	136	6	**1001**	1.54
10890	990	317	5	**814**	0.82	443	6	**1145**	1.16
18252	1404	789	5	**854**	0.61	1122	6	**1232**	0.88
28350	1890	1833	5	**947**	0.50	2222	6	**1169**	0.62

and then there is a unique fixed point. Moreover, successive approximations can be used for its computation:

$$r^0 \in \mathbb{R}^{m_c}_+ \text{ given;} \quad \text{for } k = 1, 2, \dots \text{ set } r^k = \Psi(r^{k-1}).$$

As the evaluation of Ψ requires us to solve (7.1), we can repeatedly apply Algorithm 6.1. In order to perform these computations efficiently, we initialize each iteration by results from the previous one. Finally we terminate if

$$\|r^k - r^{k-1}\|/\|r^k\| \le 10^{-4}.$$

In our numerical experiments, we consider the steel brick $\omega = (0,3)\times(0,1)\times(0,1)$ partitioned into $3N \times N \times N$ cubes by trilinear finite elements for $N = 4, 6, 8, 10, 12,$ and 14. The size of problems solved by Algorithm 6.1 is $n = 3m_c = 9N(N+1)$. In Table 7.4, we report CPU time in seconds (*time*), the number of successive approximations (*iter*), the total complexity by matrix-vector multiplications (n_Q), and the relative complexity (n_Q/n). The computations are carried out in Matlab 7 on Pentium(R)4, 3GHz, 512MB. The obtained results are promising; especially, n_Q is only mildly dependent on the finite element discretization so that the relative complexity considerably decreases for finer grids.

8. Comments and conclusions. We have analyzed a new active set algorithm for minimizing strictly convex quadratic functions with separable convex constraints. It generalizes a recently developed algorithm of quadratic programming constrained by simple bounds [5]. It should be noted that we did not need any requirement on nondegeneracy of the problem so that our algorithm is globally convergent for both the nondegenerate as well as the degenerate case.

The main goal is the proof of a linear convergence rate, which is in an optimal case described by the factor $\eta = 1 - \frac{1}{4}\kappa(A)^{-1}$ in terms of the condition number of the Hessian matrix A. Notice that η is not influenced by constraints. The key assumption is the restricted steplength $\widetilde{\alpha}$ defining the projected gradient $\widetilde{g}$, i.e., $\widetilde{\alpha} \le \|A\|^{-1}$, which means that the components $\widetilde{\varphi}$ and $\widetilde{\beta}$ of $\widetilde{g}$ are descent directions. We use them for adding/releasing indices to/from the active set in the step expansion/proportioning. Unfortunately, our analysis requires us to replace the conjugate gradient steplengths by $\widetilde{\alpha}$, which seems to be too restrictive, and the obtained convergence rate may be a bit pessimistic. That is the case for simple bound problems. If the constraints are conic (e.g., circular), the algorithm usually performs few valuable conjugate gradient steps combined with expansion steps, and then it alternates many proportioning steps with short conjugate gradient steps. The convergence rate is more realistic in such situations.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The algorithm presented here is an important ingredient in the numerical solution of 3D contact problems. It was shown in [3] that FETI domain decomposition methods are scalable for the frictionless contact. Our paper enables us to extend this result for frictional problems. It will be published in forthcoming papers.

Another class of algorithms relevant for our research (contact problems) is based on a specific semismooth Newton method that is identical again with an active set strategy [9]. The fundamental discrepancy is the fact that it allows infeasible iterates. In this case, the convergence rate is superlinear but requires a sufficiently accurate initial approximation. In general, the computational performance is comparable.

REFERENCES

[1] A. Conn, N. Gould, and P. Toint, *Testing a class of algorithms for solving minimization problems with simple bounds on the variables*, Math. Comp., 50 (1988), pp. 399–430.
[2] Z. Dostál, *Box constrained quadratic programming with proportioning and projections*, SIAM J. Optim., 7 (1997), pp. 871–887.
[3] Z. Dostál, *An optimal algorithm for a class of equality constrained quadratic programming problems with bounded spectrum*, Comput. Optim. Appl., 38 (2007), pp. 47–59.
[4] Z. Dostál, J. Haslinger, and R. Kučera, *Implementation of the fixed point method in contact problems with Coulomb friction based on a dual splitting type technique*, J. Comput. Appl. Math., 140 (2002), pp. 245–256.
[5] Z. Dostál and J. Schöberl, *Minimizing quadratic functions over non-negative cone with the rate of convergence and finite termination*, Comput. Optim. Appl., 30 (2005), pp. 23–44.
[6] A. Friedlander and M. Martínez, *On the maximization of a concave quadratic function with box constraints*, SIAM J. Optim., 4 (1994), pp. 117–192.
[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.
[8] J. Haslinger, R. Kučera, and Z. Dostál, *An algorithm for the numerical realization of 3D contact problems with Coulomb friction*, J. Comput. Appl. Math., 164-165 (2004), pp. 387–408.
[9] S. Hüeber, G. Stadler, and B. I. Wohlmuth, *A primal-dual active set algorithm for three-dimensional contact problems with Coulomb friction*, SIAM J. Sci. Comput., 30 (2008), pp. 572–596.
[10] R. Kučera, *Minimizing quadratic functions with separable quadratic constraints*, Optim. Methods Softw., 22 (2007), pp. 453–467.
[11] R. Kučera, J. Haslinger, and Z. Dostál, *A new FETI-based algorithm for solving 3D contact problems with Coulomb friction*, Lect. Notes Comput. Sci. Eng. 55, Springer, Berlin, 2007, pp. 645–652.
[12] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, 1999.
[13] B. T. Polyak, *The conjugate gradient method in extremal problems*, USSR Comput. Math. and Math. Phys., 9 (1969), pp. 94–112.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.