Optimization Methods and Software
Vol. 22, No. 3, June 2007, 453–467

Taylor & Francis
Taylor & Francis Group

Minimizing quadratic functions with separable quadratic constraints

R. KUČERA*

VŠB-Technical University, Ostrava, Czech Republic

(Received 31 October 2004; in final form 1 February 2006)

This article deals with minimizing quadratic functions with a special form of quadratic constraints that arise in 3D contact problems of linear elasticity with isotropic friction [Haslinger, J., Kučera, R. and Dostál, Z., 2004, An algorithm for the numerical realization of 3D contact problems with Coulomb friction. *Journal of Computational and Applied Mathematics*, **164/165**, 387–408.]. The proposed algorithm combines the active set method with the conjugate gradient method. Its general scheme is similar to the algorithms of Polyak's type that solve the quadratic programming problems with simple bounds. As the algorithm does not terminate in a finite number of steps, the convergence is proved. The implementation uses an adaptive precision control of the conjugate gradient loops. Numerical experiments demonstrate the computational efficiency of the method.

Keywords: Quadratic function; Separable quadratic constraints; Active set; Convergence; Conjugate gradient method; Adaptive precision control

2000 Mathematics Subject Classification: 65K05, 90C20

1. Introduction

We shall be concerned with solving

$$\min_{x\in\Omega} f(x) \tag{1}$$

where $f(x) = (1/2)\,x^\top Ax - x^\top b$, $\Omega = \{x \in \mathbb{R}^{2m}\colon\ x_i^2 + x_{2i}^2 \le g_i^2,\quad i = 1,\ldots,m\}$, $A \in \mathbb{R}^{2m\times 2m}$ is a symmetric positive definite matrix, $b \in \mathbb{R}^{2m}$, g_i are positive values and x_i denotes the ith entry of a vector $x \in \mathbb{R}^{2m}$. Let us point out that the quadratic constraints

$$x_i^2 + x_{2i}^2 \le g_i^2 \tag{2}$$

are separable with respect to the pairs $(x_i, x_{2i})^\top \in \mathbb{R}^2$ and can be interpreted so that the ith pair lies in the circle with the centre in the origin of $\mathbb{R}^2$ and with the radius g_i.

Problem (1) arises if we want to solve 3D contact problems of the linear elasticity with isotropic friction. In our previous paper [1], we have shown how to approximate the circles by

*Email: radek.kucera@vsb.cz

Optimization Methods and Software
ISSN 1055-6788 print/ISSN 1029-4937 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/10556780600609246

intersection of rotated squares so that problem (1) is approximated by a quadratic programming problem with simple bounds and equality constraints. This approach has the advantage that the existing fast algorithms can be applied immediately. However, fine approximations of the circles increase considerably the number of unknowns and lengthen the computational time. This paper is motivated by an endeavour to propose an algorithm, solving (1) directly so that the number of unknowns is not increased.

We are inspired by the efficient algorithms that combine the active set method [2] with the conjugate gradient (CG) method [3]. Algorithms of this type have been developed to minimizes the function f constrained by simple bounds (box constraints) [4–6] and date back at least to Polyak [7]. In Polyak's algorithm, the CG method minimizes the function f on the face of the domain defined by the current active set until either the minimum on the current face is reached or an infeasible iteration is generated. In the first case, indices of the constraints that violate the Karush–Kuhn–Tucker (KKT) conditions are released from the active set, whereas in the second case, the active set is expanded. In both cases, the value of f decreases so that the active sets can never reappear. As the number of possible active sets is finite, the algorithm necessarily finds the active set corresponding to the solution, and the solution itself, in a finite number of steps.

In principle, the scheme of Polyak's algorithm can be used to solve problem (1). The fundamental difference consists of the fact that now the algorithm does search not only for indices of active constraints but also for positions of corresponding pairs $(x_i, x_{2i})^\top$ lying on the boundaries of the circles $x_i^2 + x_{2i}^2 = g_i^2$. As the finite termination property can not be expected in such a case, the convergence must be proved by different arguments. This proof is the main goal of this article.

This article is organized as follows. In section 2, we introduce notations and discuss the KKT conditions. Section 3 is of cardinal importance for the proof of convergence as well as the implementation. In this section, we study the properties of a feasible descent direction that makes possible to release indices from the active set. The general scheme of the algorithm based on the active set method is given in section 4. We prove its convergence and discuss its variant with an adaptive precision control. The implementation using the CG method is described in section 5. Finally, section 6 presents the results of the numerical experiments.

2. Notations and preliminaries

Let us point out that a solution to (1) always exists and it is necessarily unique. We shall denote it by $\bar{x}$. It is well known that $\bar{x}$ is fully determined by the KKT conditions [8]. Before giving their appropriate form, we shall introduce notations.

We shall denote the Euclidean norm in $\mathbb{R}^{2m}$ by $\|\cdot\|$, and the same symbol stands for the induced matrix norm. Let $\mathcal{N}$ denote the set of all indices so that

$$\mathcal{N} = \{1, 2, \ldots, m\}.$$

We shall use the following convention throughout this article: if $x \in \mathbb{R}^{2m}$ is a vector, then x_i is its ith entry and $\boldsymbol{x}_i \in \mathbb{R}^2$ denotes the pair of its entries with indices that are bound in the ith quadratic constraint (2), i.e.

$$\boldsymbol{x}_i = (x_i, x_{2i})^\top, \quad i \in \mathcal{N}.$$

Quadratic constraint (2) then read as $\|\boldsymbol{x}_i\| \leq g_i$. Let us define the gradient at x by

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

As the quadratic constraints are separable, it can be shown that the KKT conditions are separable as well. To this end, we introduce the Lagrangian to (1) by

$$L(x, \mu) = \frac{1}{2} x^\top A x - x^\top b + \sum_{i \in \mathcal{N}} \mu_i (x_i^2 + x_{2i}^2 - g_i^2),$$

where μ_i are Lagrange multipliers. Then, the KKT conditions to (1) read as

$$\begin{aligned} \partial_{x_i} L(\bar{x}, \mu) &= \bar{r}_i + 2\mu_i \bar{x}_i = 0, \\ \partial_{x_{2i}} L(\bar{x}, \mu) &= \bar{r}_{2i} + 2\mu_i \bar{x}_{2i} = 0, \\ \partial_{\mu_i} L(\bar{x}, \mu) = \bar{x}_i^2 + \bar{x}_{2i}^2 - g_i^2 \le 0, &\quad \mu_i \ge 0, \quad \mu_i \partial_{\mu_i} L(\bar{x}, \mu) = 0 \end{aligned}$$

for $i \in \mathcal{N}$, where $\bar{r} = r(\bar{x})$. We can rewrite them more compactly by

$$\begin{aligned} &\bar{\boldsymbol{r}}_i + 2\mu_i \bar{\boldsymbol{x}}_i = 0, \\ &\|\bar{\boldsymbol{x}}_i\| \le g_i, \quad \mu_i \ge 0, \quad \mu_i (\|\bar{\boldsymbol{x}}_i\| - g_i) = 0 \end{aligned}$$

for $i \in \mathcal{N}$. After eliminating the Lagrange multipliers, the solution $\bar{x}$ to (1) is determined so that for $i \in \mathcal{N}$:

$$\|\bar{\boldsymbol{x}}_i\| < g_i \text{ implies } \bar{\boldsymbol{r}}_i = 0, \tag{3}$$

$$\|\bar{\boldsymbol{x}}_i\| = g_i \text{ implies } \bar{\boldsymbol{r}}_i + \frac{\|\bar{\boldsymbol{r}}_i\|}{g_i} \bar{\boldsymbol{x}}_i = 0. \tag{4}$$

Conditions (3) and (4) are called the inner KKT conditions and the boundary KKT conditions, respectively. They represent the equilibrium state of $\bar{x}$ with a simple geometrical interpretation (figure 1a). If the pair $\bar{\boldsymbol{x}}_i$ lies in the interior of the circle representing the ith quadratic constraint, then (3) implies that there is not any descent direction with respect to this pair (figure 1b). If the pair $\bar{\boldsymbol{x}}_i$ lies on the boundary of the circle, then (4) implies that $\bar{\boldsymbol{x}}_i$ and $\bar{\boldsymbol{r}}_i$ are opposite vectors. Therefore, any descent direction with respect to this pair is infeasible because with components of the gradient $\bar{\boldsymbol{r}}_i$ it necessarily forms an angle greater than $\pi/2$.

Figure 2a shows situations in which the KKT conditions are not satisfied. If the pair $\boldsymbol{x}_i$ lies in the interior of the circle, then $-\boldsymbol{r}_i$ is the descent direction with respect to this pair (figure 2b) If the pair $\boldsymbol{x}_i$ lies on the boundary of the circle, then the normal line to $\boldsymbol{r}_i$ going through $\boldsymbol{x}_i$ cuts off a segment of the circle (grey area). Any vector beginning at $\boldsymbol{x}_i$ and oriented into the segment is the feasible descent direction with respect to this pair. We shall justify this heuristic observation in the following section.

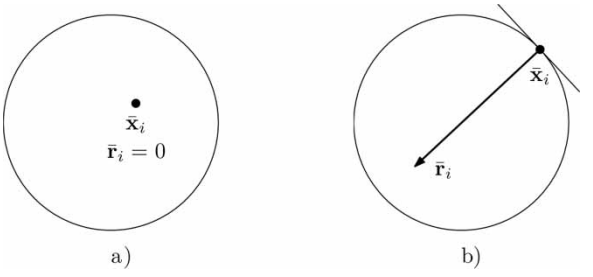

Figure 1. KKT conditions.

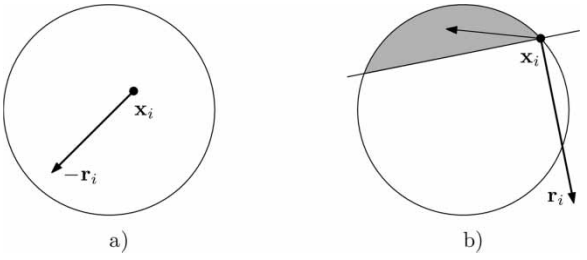

Figure 2. Situations in which KKT conditions are not satisfied.

In order to enable an alternative reference to the KKT conditions, we decompose $\mathcal{N}$ for given $x \in \Omega$ on the *active set* $\mathcal{A}(x)$ and the *free set* $\mathcal{F}(x)$ so that

$$\mathcal{A}(x) = \{i \in \mathcal{N}\colon \|\boldsymbol{x}_i\| = g_i\},$$
$$\mathcal{F}(x) = \{i \in \mathcal{N}\colon \|\boldsymbol{x}_i\| < g_i\}.$$

Moreover, we define the disequilibrium set $\mathcal{D}(x)$ as the subset of $\mathcal{A}(x)$ so that

$$\mathcal{D}(x) = \left\{i \in \mathcal{A}(x)\colon \boldsymbol{r}_i + \frac{\|\boldsymbol{r}_i\|}{g_i}\boldsymbol{x}_i \neq 0\right\}.$$

Let us denote the free gradient $\varphi = \varphi(x) \in \mathbb{R}^{2m}$ and the turned boundary gradient $\beta = \beta(x) \in \mathbb{R}^{2m}$ so that

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

and

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

respectively. Thus, KKT boundary conditions (4) are satisfied at $x \in \Omega$ iff $\beta(x) = 0$ and KKT conditions (3) and (4) are satisfied iff $\varphi(x) + \beta(x) = 0$. For a convenience, let us denote the turned gradient by $\nu(x) \in \mathbb{R}^{2m}$ so that $\nu(x) = \varphi(x) + \beta(x)$.

Finally, we introduce the product $\odot\colon \mathbb{R}^m \times \mathbb{R}^{2m} \to \mathbb{R}^{2m}$, $\alpha \odot \beta = \gamma$ so that the pairs $\boldsymbol{\gamma}_i$ are defined by

$$\boldsymbol{\gamma}_i = \alpha_i \boldsymbol{\beta}_i, \quad i \in \mathcal{N}.$$

3. Properties of the turned boundary gradient

In this section, we shall assume that boundary KKT conditions (4) are not satisfied at $x \in \Omega$ so that $\mathcal{D}(x)$ is not empty. We shall show that $-\beta$ is the feasible descent direction that releases indices of $\mathcal{D}(x)$ from the active set $\mathcal{A}(x)$. In other words, we shall prove that

$$x - \alpha \odot \beta \in \Omega, \tag{5}$$
$$f(x - \alpha \odot \beta) < f(x), \tag{6}$$
$$\mathcal{A}(x - \alpha \odot \beta) = \mathcal{A}(x) \setminus \mathcal{D}(x) \tag{7}$$

for $\alpha \in \mathbb{R}^m$ with sufficiently small (non-negative) entries α_i.

First of all, let us point out that only the pairs

$$\boldsymbol{\beta}_i = \boldsymbol{r}_i + \frac{\|\boldsymbol{r}_i\|}{g_i}\boldsymbol{x}_i \tag{8}$$

for $i \in \mathcal{D}(x)$ are non-vanishing. As $\boldsymbol{x}_i$ is the outer normal vector to the circle representing the ith quadratic constraint, it is easily seen in figure 3 that $-\boldsymbol{\beta}_i$ is oriented inside the circle. Therefore, $-\beta$ is feasible.

LEMMA 3.1 *Let us have $\alpha \in \mathbb{R}^m$ with $0 < \alpha_i < \widetilde{\alpha}_i$ for $i \in \mathcal{D}(x)$, where*

$$\widetilde{\alpha}_i = 2\,\frac{\boldsymbol{x}_i^\top \boldsymbol{\beta}_i}{\boldsymbol{\beta}_i^\top \boldsymbol{\beta}_i}. \tag{9}$$

Then (5) *and* (7) *hold.*

Proof It follows from the fact that $\widetilde{\alpha}_i$ is determined by $\|\boldsymbol{x}_i - \widetilde{\alpha}_i \boldsymbol{\beta}_i\| = g_i$. ■

Before proving (6), we give an auxiliary lemma.

LEMMA 3.2 *It holds*

$$\boldsymbol{\beta}_i^\top \boldsymbol{r}_i = \frac{1}{2}\|\boldsymbol{\beta}_i\|^2.$$

Proof Using (8), we obtain

$$\boldsymbol{\beta}_i^\top \boldsymbol{\beta}_i = 2\left(\boldsymbol{r}_i^\top \boldsymbol{r}_i + \frac{\|\boldsymbol{r}_i\|}{g_i}\boldsymbol{x}_i^\top \boldsymbol{r}_i\right) = 2\boldsymbol{\beta}_i^\top \boldsymbol{r}_i.$$

■

LEMMA 3.3 *Let us have $\alpha \in \mathbb{R}^m$ with $0 < \alpha_i \leq \eta \|A\|^{-1}$ for $i \in \mathcal{D}(x)$, where $0 < \eta < 1$. Then* (6) *holds.*

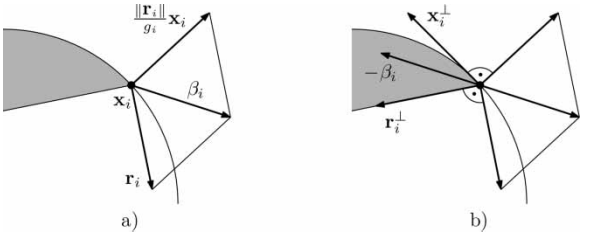

Figure 3. Pair $\boldsymbol{\beta}_i$ of the turned boundary gradient β.

Proof First, we prove the lower bound of the increment. Using Lemma 3.2, we obtain

$$\begin{aligned} f(x) - f(x - \alpha \odot \beta) &= (\alpha \odot \beta)^\top r - \frac{1}{2}(\alpha \odot \beta)^\top A(\alpha \odot \beta) \\ &\geq (\alpha \odot \beta)^\top r - \frac{1}{2}\|A\|\|\alpha \odot \beta\|^2 \\ &= \sum_{i \in \mathcal{D}(x)} \alpha_i \boldsymbol{\beta}_i^\top \boldsymbol{r}_i - \frac{1}{2}\|A\| \sum_{i \in \mathcal{D}(x)} \alpha_i^2 \|\boldsymbol{\beta}_i\|^2 \\ &\geq \sum_{i \in \mathcal{D}(x)} \alpha_i \boldsymbol{\beta}_i^\top \boldsymbol{r}_i - \frac{\eta}{2} \sum_{i \in \mathcal{D}(x)} \alpha_i \|\boldsymbol{\beta}_i\|^2 \\ &= \frac{1}{2} \sum_{i \in \mathcal{D}(x)} \alpha_i \|\boldsymbol{\beta}_i\|^2 - \frac{\eta}{2} \sum_{i \in \mathcal{D}(x)} \alpha_i \|\boldsymbol{\beta}_i\|^2 \\ &= \frac{1 - \eta}{2} \sum_{i \in \mathcal{D}(x)} \alpha_i \|\boldsymbol{\beta}_i\|^2. \end{aligned} \tag{10}$$

As the last term is positive, we have proved (6). ■

In order to satisfy (5), (6) and (7) simultaneously, we define $\alpha \in \mathbb{R}^m$ by

$$\alpha_i = \begin{cases} \min\{\eta \|A\|^{-1}, \delta \widetilde{\alpha}_i\} & \text{for } i \in \mathcal{D}(x), \\ 0 & \text{for } i \in \mathcal{N} \setminus \mathcal{D}(x), \end{cases} \tag{11}$$

where $0 < \eta < 1$, $0 < \delta < 1$ and $\widetilde{\alpha}_i$ are defined by (9).

COROLLARY 3.4 *Let us denote*

$$\mathcal{I}(x) = \{i \in \mathcal{D}(x) \colon \alpha_i = \eta \|A\|^{-1}\}.$$

One of the following three assertions holds for each $i \in \mathcal{A}(x)$*:*

$$(i) \quad \boldsymbol{\beta}_i = 0, \qquad \text{if } i \in \mathcal{A}(x) \setminus \mathcal{D}(x),$$

$$(ii) \quad f(x) - f(x - \alpha \odot \beta) \geq (1 - \eta)\delta \boldsymbol{x}_i^\top \boldsymbol{\beta}_i, \qquad \text{if } i \in \mathcal{D}(x) \setminus \mathcal{I}(x),$$

$$(iii) \quad f(x) - f(x - \alpha \odot \beta) \geq \frac{(1 - \eta)\eta}{2\|A\|} \|\boldsymbol{\beta}_i\|^2, \quad \text{if } i \in \mathcal{I}(x).$$

Proof The assertion (i) is obvious. In order to prove (ii) and (iii), we use (10), (11) and (9):

$$\begin{aligned} f(x) - f(x - \alpha \odot \beta) &\geq \frac{1 - \eta}{2} \sum_{i \in \mathcal{D}(x)} \alpha_i \|\boldsymbol{\beta}_i\|^2 \\ &= \frac{1 - \eta}{2} \left(\eta \|A\|^{-1} \sum_{i \in \mathcal{I}(x)} \|\boldsymbol{\beta}_i\|^2 + 2\delta \sum_{i \in \mathcal{D}(x) \setminus \mathcal{I}(x)} \boldsymbol{x}_i^\top \boldsymbol{\beta}_i \right). \end{aligned}$$

As all summation terms in the last expression are positive, we have proved (ii) and (iii). ■

LEMMA 3.5 *Let* $\|\boldsymbol{x}_i\| = g_i$. *If* $\boldsymbol{x}_i^\top \boldsymbol{\beta}_i = 0$, *then* $\boldsymbol{\beta}_i = 0$.

Proof Using (8), we obtain $0 = \boldsymbol{x}_i^\top \boldsymbol{\beta}_i = \boldsymbol{x}_i^\top \boldsymbol{r}_i + \|\boldsymbol{r}_i\| g_i$ implying $\boldsymbol{x}_i^\top \boldsymbol{r}_i = -\|\boldsymbol{r}_i\| g_i$. Substituting this in

$$\boldsymbol{\beta}_i^\top \boldsymbol{\beta}_i = 2\left(\|\boldsymbol{r}_i\|^2 + \frac{\|\boldsymbol{r}_i\|}{g_i}\boldsymbol{x}_i^\top \boldsymbol{r}_i\right) = 0$$

so that the lemma holds. ■

Remark 1 Let $i \in \mathcal{D}(x)$. Define the vectors orthogonal to $\boldsymbol{x}_i$, $\boldsymbol{r}_i$ by $\boldsymbol{x}_i^\perp$, $\boldsymbol{r}_i^\perp$, respectively, and assume that this definition is in agreement with figure 3b. It is easily seen from the figure that $-\boldsymbol{\beta}_i$ bisects the angle formed by $\boldsymbol{x}_i^\perp$ and $\boldsymbol{r}_i^\perp$. Therefore, $-\boldsymbol{\beta}_i$ is the ideal compromise between 'infeasibility' represented by $\boldsymbol{x}_i^\perp$ and 'stagnation' represented by $\boldsymbol{r}_i^\perp$. In this sense, the vector $-\beta$ is the *optimal* feasible descent direction that releases indices of $\mathcal{D}(x)$ from the active set $\mathcal{A}(x)$.

4. Algorithm and convergence

We shall explain the key idea of our algorithm for the one quadratic constraint (2), i.e. we shall assume for a moment that $m = 1$ so that Ω is the circle in $\mathbb{R}^2$. First, we compute the unconstrained minimum of f. If it is feasible, then we have the solution because the inner KKT condition (3) is satisfied. If the unconstrained minimum lies outside Ω, then we search for a solution on the boundary of Ω so that the boundary KKT condition (4) is satisfied. In this case, we use alternately the steps called expansion and deactivation, which generate iterations lying on the boundary of Ω and in the interior of Ω, respectively. The step expansion uses arbitrary (e.g. gradient) descent direction. The step deactivation is standardly performed by the turned boundary gradient β. As the values of f decrease in both cases, the iterations can never reappear and therefore the convergence can be proved.

The idea described here can be used simultaneously for all circles representing Ω in the general problem (1). In this case, we are not able to discover by a prior unconstrained minimization which of constraints are active in the solution. Therefore, we supplement the iterative process by the step minimization in which we compute the minimum of f with respect to components that are not active in the current iteration. To this end, we define the face at given $x \in \Omega$ by

$$W(x) = \{y \in \mathbb{R}^{2m} \colon \boldsymbol{y}_i = \boldsymbol{x}_i,\ i \in \mathcal{A}(x)\}.$$

The general scheme of our algorithm for solving problem (1) reads as follows:

ALGORITHM 4.1 *Let $x^0 \in \Omega$. If $\nu(x^k) \neq 0$, find $x^{k+1} \in \Omega$ by one of the following three steps*:
Minimization: If $\varphi(x^k) \neq 0$, then compute y so that

$$f(y) = \min\{f(x) \colon x \in W(x^k)\}. \tag{12}$$

If $y \in \Omega$, then set $x^{k+1} = y$ and so

$$f(x^{k+1}) < f(x^k) \quad \text{and} \quad \mathcal{A}(x^{k+1}) \supseteq \mathcal{A}(x^k).$$

Expansion: If $\varphi(x^k) \neq 0$ and $y \notin \Omega$, then find x^{k+1} so that

$$f(x^{k+1}) < f(x^k) \quad \text{and} \quad \mathcal{A}(x^{k+1}) \supset \mathcal{A}(x^k).$$

Deactivation: If $\varphi(x^k) = 0$ and $\beta(x^k) \neq 0$, then set

$$x^{k+1} = x^k - \alpha^k \odot \beta^k, \tag{13}$$

where $\beta^k = \beta(x^k)$ *is the turned boundary gradient and* $\alpha^k = \alpha$ *is defined by* (11) *and so*

$$f(x^{k+1}) < f(x^k) \quad \textit{and} \quad \mathcal{A}(x^{k+1}) \subset \mathcal{A}(x^k).$$

The convergence of the algorithm is enforced by properties of the turned boundary gradient in the step deactivation and by the fact that the algorithm generates a strictly decreasing sequence $\{f(x^k)\}$.

THEOREM 4.1 *Let* $x^0 \in \Omega$, $0 < \eta < 1$, $0 < \delta < 1$ *and let* $\{x^k\}$ *denote a sequence generated by Algorithm* 4.1. *Then*, $\{x^k\}$ *converges to the solution* $\bar{x}$ *of* (1).

Proof First of all, let us point out that Lemma 3.1 and Lemma 3.3 prove that the algorithm is well defined, i.e. iterations generated by (13) are feasible, decreasing and release indices from the active set.

The algorithm generates a finite or infinite sequence $\{x^k\}$. If $\{x^k\}$ is finite, then its last term solves problem (1) because it satisfies the KKT conditions. In the rest of the proof, we shall assume that $\{x^k\}$ is infinite.

Let $\mathcal{K}_{\mathrm{D}}$ denote the set of all indices such that $\varphi(x^k) = 0$ and $\beta(x^k) \neq 0$ for each $k \in \mathcal{K}_{\mathrm{D}}$, i.e. the next iteration x^{k+1} is generated by the step deactivation. Let us point out that $\mathcal{K}_{\mathrm{D}}$ is necessarily an infinite set. As Ω is compact and $x^k \in \Omega$, there is an accumulation point $\bar{x} \in \Omega$ of the sequence $\{x^k\colon k \in \mathcal{K}_{\mathrm{D}}\}$ and a subset $\mathcal{K}_{\mathrm{D}}^0$ of $\mathcal{K}_{\mathrm{D}}$ such that $\{x^k\colon k \in \mathcal{K}_{\mathrm{D}}^0\}$ converges to $\bar{x}$. Because f is continuous, it follows that $\{f(x^k)\colon k \in \mathcal{K}_{\mathrm{D}}^0\}$ converges to $f(\bar{x})$. We shall prove that $\bar{x}$ satisfies KKT conditions (3) and (4).

Let us first suppose that $\|\bar{\boldsymbol{x}}_i\| < g_i$ for fixed $i \in \mathcal{N}$. Then, $\|\boldsymbol{x}_i^k\| < g_i$ for sufficiently large $k \in \mathcal{K}_{\mathrm{D}}^0$ and because $\varphi(x^k) = 0$, we have $\boldsymbol{r}_i^k = 0$ for such k. Therefore, $\bar{\boldsymbol{r}}_i = 0$ so that KKT condition (3) is satisfied.

If $\|\bar{\boldsymbol{x}}_i\| = g_i$ for fixed $i \in \mathcal{N}$, we shall distinguish two situations. If there is an infinite subsequence of $\{x^k\colon k \in \mathcal{K}_{\mathrm{D}}^0\}$ such that $\|\boldsymbol{x}_i^k\| < g_i$, we can show as stated earlier that $\bar{\boldsymbol{r}}_i = 0$ so that KKT condition (4) is satisfied. If there is no such subsequence, then there is the infinite subset $\mathcal{K}_{\mathrm{D}}^1$ of $\mathcal{K}_{\mathrm{D}}^0$ such that $\|\boldsymbol{x}_i^k\| = g_i$ for each $k \in \mathcal{K}_{\mathrm{D}}^1$. Using Corollary 3.4, we obtain:

$$\begin{aligned}
&\boldsymbol{\beta}_i^k = 0, && \text{if } i \in \mathcal{A}(x^k) \setminus \mathcal{D}(x^k),\\
&f(x^k) - f(x^{k+1}) \geq (1-\eta)\delta(\boldsymbol{x}_i^k)^{\top}\boldsymbol{\beta}_i^k, && \text{if } i \in \mathcal{D}(x^k) \setminus \mathcal{I}(x^k),\\
&f(x^k) - f(x^{k+1}) \geq \frac{(1-\eta)\eta}{2\|A\|}\|\boldsymbol{\beta}_i^k\|^2, && \text{if } i \in \mathcal{I}(x^k)
\end{aligned}$$

for each $k \in \mathcal{K}_{\mathrm{D}}^1$. Let us decompose the set $\mathcal{K}_{\mathrm{D}}^1$ on $\mathcal{K}_{\mathrm{D}}^2$, $\mathcal{K}_{\mathrm{D}}^3$ and $\mathcal{K}_{\mathrm{D}}^4$ so that

$$\begin{aligned}
\mathcal{K}_{\mathrm{D}}^2 &= \{k \in \mathcal{K}_{\mathrm{D}}^1\colon i \in \mathcal{A}(x^k) \setminus \mathcal{D}(x^k)\},\\
\mathcal{K}_{\mathrm{D}}^3 &= \{k \in \mathcal{K}_{\mathrm{D}}^1\colon i \in \mathcal{D}(x^k) \setminus \mathcal{I}(x^k)\},\\
\mathcal{K}_{\mathrm{D}}^4 &= \{k \in \mathcal{K}_{\mathrm{D}}^1\colon i \in \mathcal{I}(x^k)\}.
\end{aligned}$$

At least one of the sets $\mathcal{K}_{\mathrm{D}}^2$, $\mathcal{K}_{\mathrm{D}}^3$ and $\mathcal{K}_{\mathrm{D}}^4$ is infinite. If $\mathcal{K}_{\mathrm{D}}^2$ is infinite, then obviously $\bar{\boldsymbol{\beta}}_i = 0$. If $\mathcal{K}_{\mathrm{D}}^3$ is infinite, then $\{(\boldsymbol{x}_i^k)^{\top}\boldsymbol{\beta}_i^k\colon k \in \mathcal{K}_{\mathrm{D}}^3\}$ converges to zero. We obtain $\bar{\boldsymbol{x}}_i^{\top}\bar{\boldsymbol{\beta}}_i = 0$, and using Lemma 3.5, we obtain $\bar{\boldsymbol{\beta}}_i = 0$. If $\mathcal{K}_{\mathrm{D}}^4$ is infinite, then $\{\|\boldsymbol{\beta}_i^k\|^2\colon k \in \mathcal{K}_{\mathrm{D}}^4\}$ converges to zero and therefore we obtain again $\bar{\boldsymbol{\beta}}_i = 0$. Summing up all the three cases, we conclude that KKT boundary condition (4) is satisfied.

We have proved that $\bar{x}$ is the solution to (1). Now we shall show that the whole sequence $\{x^k\}$ converges to $\bar{x}$. It is well known that the solution to (1) is characterized by the variational

inequality $(A\bar{x} - b)^\top (x - \bar{x}) \geq 0$ for all $x \in \Omega$. By means of this inequality, we obtain

$$f(x^k) - f(\bar{x}) = (A\bar{x} - b)^\top (x^k - \bar{x}) + \frac{1}{2}(x^k - \bar{x})A(x^k - \bar{x}) \geq \frac{1}{2}\lambda_{\min}\|x^k - \bar{x}\|^2, \quad (14)$$

where $\lambda_{\min}$ is the smallest eigenvalue of A. As the sequence $\{f(x^k)\}$ is lower bounded and decreasing, its only accumulation point is necessarily $f(\bar{x})$. Therefore (14) implies that the whole sequence $\{x^k\}$ converges to $\bar{x}$. ■

Motivated by the works of Dostál and Friedlander [4, 6], we shall use the *adaptive precision control* in Algorithm 4.1. We replace the criteria $\varphi(x^k) \neq 0$ by

$$\|\varphi(x^k)\| > \gamma\|\beta(x^k)\|, \quad (15)$$

where $\gamma > 0$. The convergence of the modified algorithm is proved by the following theorem.

THEOREM 4.2 *Let $x^0 \in \Omega$, $0 < \eta < 1$, $0 < \delta < 1$, $\gamma > 0$ and let $\{x^k\}$ denote a sequence generated by Algorithm* 4.1, *in which the criteria $\varphi(x^k) \neq 0$ is replaced by* (15). *Then, $\{x^k\}$ converges to the solution $\bar{x}$ of* (1).

Proof As the concept of the proof is the same with Theorem 4.1, we shall explain it briefly. Let $\mathcal{K}_{\mathrm{D}}$ denote again the set of all indices such that the next iteration is generated by the step deactivation, i.e. $\beta(x^k) \neq 0$ and $\|\varphi(x^k)\| \leq \gamma\|\beta(x^k)\|$ for each $k \in \mathcal{K}_{\mathrm{D}}$. Let us choose the subset $\mathcal{K}_{\mathrm{D}}^0$ of $\mathcal{K}_{\mathrm{D}}$ such that $\{x^k : k \in \mathcal{K}_{\mathrm{D}}^0\}$ converges to an accumulation point of $\{x^k : k \in \mathcal{K}_{\mathrm{D}}\}$, say $\bar{x}$. Let us point out that $\bar{x}$ necessarily satisfies

$$\|\varphi(\bar{x})\| \leq \gamma\|\beta(\bar{x})\|. \quad (16)$$

The arguments of the proof of Theorem 4.1 can be analogously used in order to prove $\bar{\boldsymbol{\beta}}_i = 0$ for $i \in \mathcal{A}(\bar{x})$, which is equivalent to $\|\beta(\bar{x})\| = 0$. Therefore, (16) implies $\|\varphi(\bar{x})\| = 0$ so that $\bar{x}$ satisfies KKT conditions (3) and (4). The rest of the proof is the same with Theorem 4.1. ■

5. Implementation

In this section, we describe the implementation of Algorithm 4.1 with the adaptive precision control (15) by the CG method. We use an easily understandable variant of the Matlab language, in which we do not distinguish generation of variables by indices unless it is convenient for further references. The algorithm is given subsequently, in which we do not distinguish generation of variables by indices unless it is convenient for further references. As the step deactivation is completely described by (13), we explain the steps minimization and expansion, which are performed by the CG loop.

First of all, we define the projection $P_\Omega : \mathbb{R}^{2m} \to \Omega$, $y = P_\Omega(x)$ so that for $i \in \mathcal{N}$,

$$\boldsymbol{y}_i = \boldsymbol{x}_i, \qquad \text{if } \|\boldsymbol{x}_i\| \leq g_i,$$
$$\boldsymbol{y}_i = \frac{g_i}{\|\boldsymbol{x}_i\|}\boldsymbol{x}_i, \quad \text{if } \|\boldsymbol{x}_i\| > g_i.$$

We apply this projection to any infeasible point generated by the algorithm.

Let us suppose that x^k satisfies (15). We start the CG loop by the initial approximation $y^{(0)} = x^k$ in order to compute the minimum of f on the face $W(x^k)$, i.e. we solve (12). Let

us point out that the constraints represented by the face can be neglected if we use another projection $P_{\mathcal{F}(x^k)}\colon \mathbb{R}^{2m} \to W(x^k) \setminus \{x^k\}$, $y = P_{\mathcal{F}(x^k)}(x)$, which is defined by

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{x}_i \quad \text{for } i \in \mathcal{F}(x^k), \\ \boldsymbol{y}_i &= 0 \quad \text{for } i \in \mathcal{N} \setminus \mathcal{F}(x^k). \end{aligned}$$

The CG loop is terminated by one of the two situations corresponding to the steps minimization or expansion.

(a) Provided that the lth CG approximation $y^{(l)} \in \Omega$ satisfies $\|\varphi(y^{(l)})\| \le \gamma \|\beta(y^{(l)})\|$, it is the solution to (12) with respect to the adaptive precision control and we set $x^{k+1} = y^{(l)}$.

(b) Provided that the lth CG approximation $y^{(l)} \notin \Omega$, we explore its projection $P_\Omega(y^{(l)})$. In advance, we check whether the value of f decreases. The CG loop is terminated if

$$f(P_\Omega(y^{(l)})) \ge f(P_\Omega(y^{(l-1)})). \tag{17}$$

In this case, we set

$$x^{k+1} = P_\Omega(y^{(l-1)} - \alpha p_{\mathcal{F}}^{(l-1)}), \tag{18}$$

where $p_{\mathcal{F}}^{(l-1)}$ is the current CG direction and α is a shortened steplength. If (17) does not hold, we test whether the adaptive precision control is reached. If

$$\|\varphi(P_\Omega(y^{(l)}))\| \le \gamma \|\beta(P_\Omega(y^{(l)}))\|,$$

then we set $x^{k+1} = P_\Omega(y^{(l)})$.

The shortened steplength in (18) must expand the active set. We can take $\alpha = 0$ provided that $y^{(l-1)} \notin \Omega$. If $y^{(l-1)} \in \Omega$, then we define α as the maximal feasible steplength. This procedure is described in the function SHORTENESS, in which we exploit the auxiliary variable *in_set*. This variable monitors whether the current approximation $y = y^{(l-1)}$ belongs to Ω. As we need to know whether the current CG approximation belongs to Ω, we monitor this information during the CG loop in the auxiliary variable in_set (= 0/1 that means $\notin / \in$).

ALGORITHM 5.1 *Let* $x^0 \in \Omega$, $0 < \eta < 1$, $0 < \delta < 1$, $\gamma > 0$ *and* $tol > 0$.

$k = 0$, $r^0 = Ax^0 - b$ {*Initialization*}
while $\|v\,(x^k)\| > tol$ {*Main loop*}
  if $\|\varphi\,(x^k)\| > \gamma\|\beta(x^k)\|$ {*Minimization and Expansion*}
    $y = x^k$, $in\_set = 1$ {*Initialization of CG method*}
    $\mathcal{F} = \mathcal{F}(x^k)$, $r_{\mathcal{F}} = P_{\mathcal{F}}(r^k)$, $p_{\mathcal{F}} = P_{\mathcal{F}}(r^k)$, $\alpha = 0$, $\alpha_{cg} = 0$
    {*CG loop*}
    while $\|v\,(P_\Omega(y))\| > tol$ and $\alpha = \alpha_{cg}$ and $\|\varphi\,(P_\Omega(y))\| > \gamma\|\beta(P_\Omega(y))\|$
      $\alpha_{cg} = r_{\mathcal{F}}^\top p_{\mathcal{F}} / p_{\mathcal{F}}^\top A p_{\mathcal{F}}$ {*CG steplength*}
      if $y - \alpha_{cg} p_{\mathcal{F}} \in \Omega$ {*Feasible CG iteration*}
        $\alpha = \alpha_{cg}$, $in\_set = 1$
      else {*Infeasible CG iteration*}
        if $f(P_\Omega(y - \alpha_{cg} p_{\mathcal{F}})) < f(P_\Omega(y))$
          $\alpha = \alpha_{cg}$, $in\_set = 0$
        else
          $\alpha =$ SHORTENESS $(y, p_{\mathcal{F}}, g, in\_set)$
        end if
      end if
      $y = y - \alpha p_{\mathcal{F}}$, $r_{\mathcal{F}} = r_{\mathcal{F}} - \alpha P_{\mathcal{F}}(Ap_{\mathcal{F}})$ {*New CG iteration*}
      $p_{\mathcal{F}} = r_{\mathcal{F}} - (r_{\mathcal{F}}^\top Ap_{\mathcal{F}} / p_{\mathcal{F}}^\top Ap_{\mathcal{F}}) p_{\mathcal{F}}$

end while
$x^{k+1} = P_\Omega(y),\ r^{k+1} = Ax^{k+1} - b,\ k = k+1$
end if
if $\|\varphi(x^k)\| \le \gamma \|\beta(x^k)\|$ {*Deactivation*}

$$\alpha_i^k = \begin{cases} \min\{\eta \|A\|^{-1},\ 2\delta\,(\mathbf{x}_i^k)^\top \boldsymbol{\beta}_i^k / (\boldsymbol{\beta}_i^k)^\top \boldsymbol{\beta}_i^k\} & \text{for } i \in \mathcal{D}(x^k) \\ 0 & \text{for } i \in \mathcal{N}\backslash\mathcal{D}(x^k) \end{cases}$$

$x^{k+1} = x^k - \alpha^k \odot \beta(x^k),\ r^{k+1} = r^k - A(\alpha^k \odot \beta(x^k)),\ k = k+1$
end if
end while

$\alpha =$ function SHORTENESS (y, p, g, in_set)
if $in\_set = 1$ {*Shortened steplength for* $y \in \Omega$}
$\alpha = \min\left\{\alpha_i : \alpha_i \ge 0,\ \|\boldsymbol{y}_i - \alpha_i \boldsymbol{p}_i\| = g_i,\ i \in \mathcal{N}\right\}$
else {*Shortened steplength for* $y \notin \Omega$}
$\alpha = 0$
end if

6. Numerical experiments

We assess the behaviour of Algorithm 5.1 by two examples. In the first example, we test experimentally the sensitivity of the algorithm on the values of γ, η and δ. The second example arises in 3D contact problems with isotropic friction. We shall demonstrate the high efficiency of the algorithm with respect to the methods used in ref. [1]. All computations are performed in Matlab 7 on Pentium 4, 3 GHz with 512 MB RAM.

Example 6.1 Let us consider problem (1) in which

$$A = \text{fivediag}(-1, -1, 4, -1, -1) \in \mathbb{R}^{12\times 12},$$
$$b = Ay,$$
$$g = (2, 1, 0.5, 2, 10^{-3}, 154)^\top$$

and $y = (2, 1, 0.5, 0, 0, 11, 10^{-5}, -1, \sqrt{2}, -0.1, 4.1 \times 10^{-4}, 143)^\top$. The solution $\bar{x} \in \mathbb{R}^{12}$ has three active constraints, $\mathcal{A}(\bar{x}) = \{2, 3, 5\}$. In tables 1–3, we shall denote k the total number of the iterations, k_{D} the number of the steps deactivation and k_{CG} the number of the CG steps, respectively. The computational costs are assessed by the number n_A of multiplications with the Hessian matrix A. We set $x^0 = 0$ and $\text{tol} = 10^{-6}\|b\|$.

Let $\eta = 0.5$ and $\delta = 0.5$. Table 1 summarizes experiments in which we change γ. The first row corresponds to the implementation of Algorithm 5.1 without the adaptive precision control so that criteria (15) is replaced by $\|\varphi(x^k)\| > \text{tol}$. The optimal values of γ are around a unity.

Let $\gamma = 1$ and $\delta = 0.5$. Table 2 summarizes experiments in which we change η. The optimal values of η are 0.8 and 0.9.

Let $\gamma = 1$ and $\eta = 0.8$. Table 3 summarizes experiments in which we change δ. The optimal values of δ are those greater than 0.4.

Table 1. Dependence on γ.

γ	k	k_{D}	k_{CG}	n_A
–	21	7	14	64
0.01	22	7	15	69
0.1	22	7	15	63
0.5	22	7	15	57
0.8	21	7	14	54
1	21	7	14	54
1.2	24	10	14	54
1.8	29	14	15	57
2	35	17	18	60
3	35	17	18	60

Table 2. Dependence on η.

η	k	k_{D}	k_{CG}	n_A
0.1	21	7	14	54
0.2	21	7	14	54
0.3	21	7	14	54
0.4	21	7	14	54
0.5	21	7	14	54
0.6	20	7	13	51
0.7	20	7	13	51
0.8	18	6	12	46
0.9	18	6	12	46
1	20	8	12	48

The algorithm is less sensitive on values of η and δ, which regulate the step deactivation. This property shows that the computational behaviour of the algorithm is determined above all by the truncated CG method used in the steps minimization and expansion.

Example 6.2 The discrete dual formulation of the 3D contact problem with (given) isotropic friction reads as follows:

$$\min \frac{1}{2}\lambda^\top Q\lambda - \lambda^\top h, \tag{19}$$

$$\text{s.t. } \lambda_i \geq 0, \quad \lambda_{2i}^2 + \lambda_{3i}^2 \leq g_i^2, \quad i = 1, \ldots, m, \tag{20}$$

where $Q \in \mathbb{R}^{3m\times 3m}$ is a symmetric positive definite matrix, $h \in \mathbb{R}^{3m}$ and $g_i \geq 0$. The unknowns λ_i represent a normal contact stress, whereas the unknowns $(\lambda_{2i}, \lambda_{3i})^\top$ represent

Table 3. Dependence on δ.

δ	k	k_{D}	k_{CG}	n_A
0.1	21	7	14	54
0.2	21	7	14	54
0.3	20	7	13	51
0.4	20	7	13	51
0.5	18	6	12	46
0.6	18	6	12	46
0.7	18	6	12	46
0.8	18	6	12	46
0.9	18	6	12	46

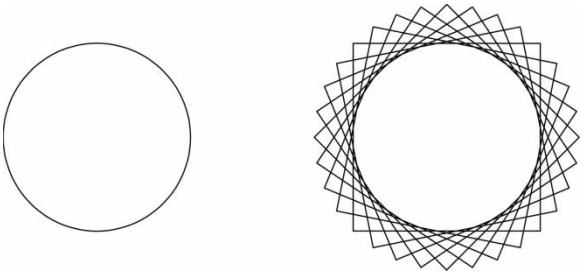

Figure 4. Approximation of the quadratic constraint.

a tangential contact stress. The method described in ref. [1] is based on approximations of the circles representing the quadratic constraints in (20) by intersections of rotated squares (figure 4). Hence, the problems (19) and (20) are approximated by the quadratic programming problem with simple bounds and linear equality constraints so that the efficient algorithm [9] based on the augmented Lagrangian can be used. We shall show that Algorithm 5.1 leads to much faster computation of the solution.

We split problems (19) and (20) by using the Gauss–Seidel-type method on independent minimizations with simple bounds and quadratic constraints, respectively. To this end, we introduce the natural decompositions of Q, h and λ as

$$Q = \begin{pmatrix} Q_{nn} & Q_{nt} \\ Q_{nt}^{\top} & Q_{tt} \end{pmatrix}, \quad h = \begin{pmatrix} h_n \\ h_t \end{pmatrix}, \quad \lambda = \begin{pmatrix} \lambda_n \\ \lambda_t \end{pmatrix}.$$

The Gauss–Seidel loop reads as follows:

Initialize $\lambda_t^{(0)}$, $k = 0$.
repeat

$$k = k + 1$$

$$\lambda_n^{(k)} = \operatorname{argmin} \left\{ \frac{1}{2} \lambda_n^{\top} Q_{nn} \lambda_n - \lambda_n^{\top} (h_n - Q_{nt} \lambda_t^{(k-1)}), \text{ s.t. } \lambda_{n,i} \geq 0 \right\} \tag{21}$$

$$\lambda_t^{(k)} = \operatorname{argmin} \left\{ \frac{1}{2} \lambda_t^{\top} Q_{tt} \lambda_t - \lambda_t^{\top} (h_t - Q_{nt}^{\top} \lambda_n^{(k)}), \text{ s.t. } \lambda_{t,i}^2 + \lambda_{t,2i}^2 \leq g_i^2 \right\} \tag{22}$$

until stopping criterion

We suppose that (21) is solved by the algorithm of ref. [5], whereas the solution to (22) is computed by Algorithm 5.1 in each iterative step. Consider the model contact problem of ref. [10] discretized by means of the finite element method so that n denotes the number of primal unknowns (displacements) and m denotes the number of dual unknowns (stresses). Table 4 compares computational times of the presented algorithm (Time) with computational times of the algorithm of ref. [1] for two (Time2) and four (Time4) rotated squares, respectively. Let us point out that the compared implementations are of the similar quality.

Table 5 shows the iteration history of the Gauss–Seidel loop for $n = 18759$ and $m = 1443$. n_A^{qpp} and n_A^{qpq} denote the number of multiplications with the Hessian matrix in (21) and (22), respectively. The stopping criterion reads as follows:

$$\text{Error} \equiv \frac{\|\lambda^{(k)} - \lambda^{(k-1)}\|}{\|\lambda^{(k)}\|} < 10^{-6}.$$

Table 4. The computational times in seconds.

n	m	Time	Time2	Time4
975	195	2	15	61
2,793	399	12	101	548
6,318	702	38	486	2,114
11,253	1,023	94	1,542	7,724
18,759	1,443	254	5,004	20,534

Table 5. The iteration history.

k	Error	n_A^{qpp}	n_A^{qpq}
1	1	39	43
2	1.7×10^{-1}	17	39
3	2.3×10^{-2}	12	31
4	2.3×10^{-3}	10	13
5	2.1×10^{-4}	7	10
6	1.2×10^{-5}	3	7
7	1.3×10^{-5}	3	4
8	0	2	4

Let us point out that the algorithm used in (21) has the finite termination property. Although our algorithm has not this property, the number of Hessian matrix multiplications n_A^{qpp} and n_A^{qpq} are comparable.

7. Conclusions and comments

We have presented a new active set-based algorithm for minimizing stricly convex quadratic functions with separable quadratic constraints. The algorithm is similar to the class of Polyak's algorithms that solve quadratic programming problems with simple inequality bounds in a finite number of iterations. Although our algorithm has not the finite termination property, numerical experiments show the comparable computational efficiency.

The algorithm has been proposed for solving 3D contact problems of linear elasticity with isotropic friction. Problems of this type are characterized by a large number of unknowns that make demands on computational efficiency as well as memory requirements. The presented algorithm is an important ingredient in the development of the FETI-based domain decomposition methods [11].

Acknowledgements

This work is supported by the National Program of Research Information Society under project 1ET400300415 and by the Grant Agency of the Czech Republic under grant 101/04/1145.

References

[1] Haslinger, J., Kučera, R. and Dostál, Z., 2004, An algorithm for the numerical realization of 3D contact problems with Coulomb friction. *Journal of Computational and Applied Mathematics*, **164/165**, 387–408.
[2] Fletcher, R., 1997, *Practical Methods of Optimization* (Chichester: John Wiley & Sons).
[3] Golub, G.H. and Van Loan, C.F., 1996, *Matrix Computation* (Baltimore: The Johns Hopkins University Press).
[4] Dostál, Z., 1997, Box constrained quadratic programming with proportioning and projections. *SIAM Journal on Optimization*, **7**, 871–887.

[5] Dostál, Z. and Schöberl, J., 2005, Minimizing quadratic functions over non-negative cone with the rate of convergence and finite termination. *Computational Optimization and Applications*, **30**, 23–44.
[6] Friedlander, A. and Martínez, M., 1994, On the maximization of a concave quadratic function with box constraints. *SIAM Journal on Optimization*, **4**, 117–192.
[7] Polyak, B.T., 1969, The conjugate gradient method in extremal problems. *USSR Computational Mathematics and Mathematical Physics*, **9**, 94–112.
[8] Bazaraa, M.S. and Shetty C.M., 1979, *Nonlinear Programming* (New York: John Wiley and Sons).
[9] Dostál, Z., Friedlander, A. and Santos, S.A., 2003, Augmented Lagrangian with adaptive precision control for quadratic programming with simple bounds and equality constraints. *SIAM Journal on Optimization*, **13**, 1120–1140.
[10] Kučera, R., Haslinger, J. and Dostál, Z., 2004, An algorithm for solving 3D contact problems with friction. In: T.E. Simos, Ch. Tsitouras (Eds) *International Conference on Numerical Analysis and Applied Mathematics*, (Berlin: WILEY-VCH), pp. 217–220.
[11] Kučera, R., Haslinger, J. and Dostál, Z., 2005, A new FETI based algorithm for solving 3D contact problems with Coulomb friction. *Lecture Notes in Computational Science and Engineering*, in press.