ON MINIMIZING QUADRATIC FUNCTIONS WITH SEPARABLE CONVEX CONSTRAINTS

R. Kučera
Dept. of Mathematics, VŠB–Technical University of Ostrava, Czech Republic
radek.kucera@vsb.cz

Abstract: A new active set algorithm for minimizing quadratic functions with separable convex constraints combines the conjugate gradient method with gradient projections. 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 proved. Numerical experiments including frictional 3D contact problems of linear elasticity illustrate the computational performance.

Keywords: *Quadratic function, Separable convex constraints, Active set, Conjugate gradient method, Projected gradient, Convergence rate*

1. Introduction

We shall be concerned with solving

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

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 $\Omega_i = \{\boldsymbol{x}_i \in \mathbb{R}^{n_i} : f_i(\boldsymbol{x}_i) \leq 0\}$ are defined by continuously differentiable convex functions $f_i : \mathbb{R}^{n_i} \mapsto \mathbb{R}$ so that $n_i \geq 1$, $\sum_{i=1}^m n_i = n$. 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*

$$l_i \leq x_i \tag{2}$$

arising from 2D contact problems [1].

– 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

$$x_{2i-1}^2 + x_{2i}^2 \leq r_i^2 \tag{3}$$

that can be interpreted as the *circular constraint*. A source of such constraints is a friction law for 3D contact problems [3].

2. Algorithm

Let $\mathcal{M} = \{1, \ldots, m\}$. The gradient of f at a point $x \in \Omega$ is $g := g(x) = Ax - b$ and the active-set is defined by $\mathcal{A} := \mathcal{A}(x) = \{i \in \mathcal{M} |\ f_i(\boldsymbol{x}_i) = 0\}$. Using the projection $P_\Omega : \mathbb{R}^n \mapsto \Omega$, we define the *projected gradient* for a fixed $\widetilde{\alpha} \geq 0$ as

$$\widetilde{g} := \widetilde{g}(x) = \frac{1}{\widetilde{\alpha}}(x - P(x - \widetilde{\alpha} g(x))).$$

The projected gradient characterizes the solution x^* of (1) by $\widetilde{g}(x^*) = 0$. Our algorithm is based on the fact that the non-zero components of $\widetilde{g}(x)$ at $x \neq x^*$ determine the descent directions changing appropriately the active-set. To this end, we introduce components of $\widetilde{g}(x)$ called the *projected free gradient* $\widetilde{\phi} := \widetilde{\phi}(x)$ and the *projected boundary gradient* $\widetilde{\beta} := \widetilde{\beta}(x)$, respectively, as follows:

$$\begin{aligned} \widetilde{\phi}_{\mathcal{A}} &= 0, & \widetilde{\phi}_{\mathcal{M}\setminus\mathcal{A}} &= \widetilde{g}_{\mathcal{M}\setminus\mathcal{A}}, \\ \widetilde{\beta}_{\mathcal{A}} &= \widetilde{g}_{\mathcal{A}}, & \widetilde{\beta}_{\mathcal{M}\setminus\mathcal{A}} &= 0. \end{aligned}$$

We combine three steps to generate a sequence $\{x^{(l)}\}$ that approximates the solution to (1):

- the *expansion step*: $x^{(l+1)} = x^{(l)} - \widetilde{\alpha}\widetilde{\phi}(x^{(l)})$,
- the *proportioning step*: $x^{(l+1)} = x^{(l)} - \widetilde{\alpha}\widetilde{\beta}(x^{(l)})$,
- the *conjugate gradient step*: $x^{(l+1)} = x^{(l)} - \alpha_{cg}^{(l)} p^{(l)}$, where $\alpha_{cg}^{(l)}$ and the conjugate gradient directions $p^{(l)}$ are computed recurrently; the recurrence starts from $x^{(s)}$ generated by the last expansion or the proportioning step and satisfies $\mathcal{A}(x^{(l+1)}) = \mathcal{A}(x^{(s)})$.

The expansion step may add indices while the proportioning step may release indices to/from the current active-set. The conjugate gradient steps are used to carry out efficiently the minimization of the objective f in the interior of the face $W(x^{(s)}) = \{x \in \Omega |\ x_{\mathcal{A}} := x_{\mathcal{A}}^{(s)}, \mathcal{A} = \mathcal{A}(x^{(s)})\}$. Moreover, the algorithm exploits a given constant $\gamma > 0$ and the *releasing criterion*

$$\widetilde{\beta}(x^{(l)})^\top g(x^{(l)}) \leq \gamma\ \widetilde{\phi}(x^{(l)})^\top g(x^{(l)}) \tag{4}$$

to decide which of the steps will be performed.

Algorithm QPC [4] Let $x^{(0)} \in \Omega$, $\gamma > 0$, $\widetilde{\alpha} \in (0, \|A\|^{-1}]$ and $\epsilon > 0$ be given. For $x^{(l)}$, $x^{(s)}$ known, $0 \leq s \leq l$, where $x^{(s)}$ is computed by the last step expansion or proportioning, choose $x^{(l+1)}$ by the following rules:

(i) If $\|\widetilde{g}(x^{(l)})\| \leq \epsilon$, return $x := x^{(l)}$.

(ii) If $x^{(l)}$ fulfils (4), try to generate $x^{(l+1)}$ by the conjugate gradient step. If $x^{(l+1)} \in \operatorname{Int} W(x^{(s)})$, accept it, otherwise generate $x^{(l+1)}$ by the expansion step.

(iii) If $x^{(l)}$ does not fulfil (4), generate $x^{(l+1)}$ by the proportioning step.

Contrary to simple bound problems analyzed in [2], the algorithm does not exhibit the finite terminating property while the same convergence rate is achieved.

Theorem 1 [5] *Let $x^* \in \Omega$ be the solution to (1), α_{min} denote the smallest eigenvalue of A and $\widehat{\gamma} = \max\{\gamma, \gamma^{-1}\}$. Let $\{x^{(l)}\}$ be the sequence generated by Algorithm 2.1 with $\epsilon = 0$. Then*

$$f(x^{(l+1)}) - f(x^*) \leq \eta \left(f(x^{(l)}) - f(x^*) \right),$$

where

$$\eta = 1 - \frac{\widetilde{\alpha}\alpha_{min}}{2 + 2\widehat{\gamma}} < 1.$$

The error in the A-energy norm is bounded by

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

Theorem 1 yields the optimal value of η for $\gamma = \hat{\gamma} = 1$ and $\widetilde{\alpha} = \|A\|^{-1}$, when

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

where $\kappa(A)$ is the spectral condition number of A.

3. Numerical experiments

We consider a steel brick supported by a rigid foundation that occupies the bounded domain. The boundary is split into three nonempty disjoint parts with different boundary conditions: zero displacements, surface tractions and contact conditions. The contact conditions are represented by the non-penetration, an effect of friction and the transmission of contact stresses; see [3].

The discrete contact problem with *Tresca friction* reduces to (1). As the more realistic *Coulomb friction* law leads to a sequence of Tresca friction problems we can repeatedly apply Algorithm QPC to solve it.

The table summarizes CPU time (in seconds), the number of successive approximations (*iter*) and the total complexity in terms of matrix-vector multiplications (n_A) for various primal and dual dofs $3n_c$ and $n = 3m_c$, respectively, and for two coefficients of friction F. The obtained results are promising, especially, n_A is only mildly dependent on the finite element discretization so that the relative complexity n_A/n considerably decreases for finer grids.

dofs		$F = 0.3$				$F = 0.6$			
$3n_c$	$3m_c$	*time*	*iter*	n_A	n_A/n	*time*	*iter*	n_A	n_A/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

Tab. 1. Contact problem with Coulomb friction.

Acknowledgements

Supported by the grant 101/08/0574 of the Grant Agency of the Czech Republic and by the Research Project MSM6198910027 of the Czech Ministry of Education.

References

[1] 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), 245-256.

[2] 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), 23-44.

[3] 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), 387-408.

[4] R. Kučera, Minimizing quadratic functions with separable quadratic constraints, *Optim. Meth. Soft.* **22** (2007), 453-467.

[5] R. Kučera, Convergence rate of an optimization algorithm for minimizing quadratic functions with separable convex constraints, *SIAM Optim.* **9** (2008), 846-862.