FACTA UNIVERSITATIS (NIŠ)
Ser. Math. Inform. 20 (2005), 129–135

ON AN OPTIMIZATION ALGORITHM FOR LC^1 UNCONSTRAINED OPTIMIZATION

Nada I. Djuranović-Miličić

Abstract. In this paper an algorithm for LC^1 unconstrained optimization problems is presented. The algorithm uses the second order Dini upper directional derivative. It is proved that the algorithm is well-defined, as well as the convergence of the sequence of points obtained by the algorithm to an optimal point. An estimate of the rate of convergence is given, too.

1. Introduction

We consider the following LC^1 problem of unconstrained optimization

$$\min\Big\{f(x) \mid x \in D \subset \mathbb{R}^n\Big\}, \tag{1.1}$$

where $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is a LC^1 function on the open convex set D, that means the objective function we want to minimize is continuously differentiable and its gradient ∇f is locally Lipschitzian, i.e.,

$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\| \quad \text{for } x, y \in D$$

for some $L > 0$.

We shall present an iterative algorithm, a modification of the algorithm from [3], for finding an optimal solution to problem (1.1) by generating the sequence of points $\{x_k\}$ of the following form:

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, \ldots, \quad d_k \neq 0, \tag{1.2}$$

where the step-size α_k and the directional vector d_k are defined by the particular algorithms.

Received April 28, 2004.
2000 *Mathematics Subject Classification.* Primary 65K05; Secondary 90C30.

2. Preliminaries

We shall give some preliminaries that will be used for the remainder of the paper.

Definition 2.1. ([3]) The second order Dini upper directional derivative of the function $f \in LC^1$ at x_k in the direction $d \in \mathbb{R}^n$ is defined to be

$$f''_D(x_k; d) = \limsup_{\lambda \downarrow 0} \frac{[\nabla f(x_k + \lambda d) - \nabla f(x_k)]^T d}{\lambda}.$$

If ∇f is directionally differentiable at x_k, we have

$$f''_D(x_k; d) = f''(x_k; d) = \lim_{\lambda \downarrow 0} \frac{[\nabla f(x_k + \lambda d) - \nabla f(x_k)]^T d}{\lambda} \quad \text{for all} \quad d \in \mathbb{R}^n.$$

Lemma 2.1. ([3]) *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a LC^1 function on D, where $D \subset \mathbb{R}^n$ is an open subset. If x is a solution of LC^1 optimization problem* (1.1)*, then:*

$$f'(x; d) = 0$$

and $f''_D(x; d) \geq 0$ for all $d \in \mathbb{R}^n$.

Lemma 2.2. ([3]) *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a LC^1 function on D, where $D \subset \mathbb{R}^n$ is an open subset. If x satisfies*

$$f'(x; d) = 0$$

and $f''_D(x; d) > 0$ for all $d \in \mathbb{R}^n \setminus \{0\}$, then x is a strict local minimizer of (1.1).

3. The Optimization Algorithm

At the k-th iteration the direction vector d_k in (1.2) presents a solution of the problem

$$\min\Big\{\Phi_k(d) \mid d \in \mathbb{R}^n\Big\}, \tag{3.1}$$

where $\Phi_k(d) = \nabla f(x_k)^T d + \frac{1}{2} f''_D(x_k; d)$, and the step-size α_k is a number satisfying

$$\alpha_k = q^{i(k)}, \quad 0 < q < 1,$$

where $i(k)$ is the smallest integer from $i = 0, 1, \ldots$ such that

$$x_{k+1} = x_k + q^{i(k)} d_k \in D$$

and

$$f(x_k + q^{i(k)} d_k) - f(x_k) \leq -\frac{1}{2} q^{i(k)} \sigma(f''_D(x_k; d_k)) \tag{3.2}$$

where $\sigma : [0, +\infty) \to [0, +\infty)$ is a continuous function satisfying $\delta_1 t \leq \sigma(t) \leq \delta_2 t$, $0 < \delta_1 < \delta_2 < 1$.

We suppose that there exist constants $c_2 \geq c_1 > 0$ such that

$$c_1 \|d\|^2 \leq f''_D(x; d) \leq c_2 \|d\|^2 \tag{3.3}$$

for every $d \in \mathbb{R}^n$. It follows from Lemma 3.1 in [3] that under the assumption (3.3) the optimal solution of the problem (3.1) exists and that the sequence $\{d_k\}$ is bounded.

Proposition 3.1. *If the function $f \in LC^1$ satisfies the condition* (3.3), *then:*

1) *the function f is uniformly and, hence, strictly convex, and, consequently;*

2) *the level set $L(x_0) = \{x \in D : f(x) \leq f(x_0)\}$ is a compact convex set;*

3) *there exists a unique point x^* such that $f(x^*) = \min_{x \in L(x_0)} f(x)$.*

Proof. 1) From the assumption (3.3) and the mean value theorem it follows that for all $x \in L(x_0)$ there exists $\theta \in (0, 1)$ such that

$$\begin{aligned} f(x) - f(x_0) &= \nabla f(x_0)^T (x - x_0) + \frac{1}{2} f''_D[x_0 + \theta(x - x_0); x - x_0] \\ &\geq \nabla f(x_0)^T (x - x_0) + \frac{1}{2} c_1 \|x - x_0\|^2 > \nabla f(x_0)^T (x - x_0) \end{aligned}$$

that is, f is uniformly and consequently strictly convex on $L(x_0)$.

2) From [2] it follows that the level set $L(x_0)$ is bounded. The set $L(x_0)$ is closed because of the continuity of the function f; hence, $L(x_0)$ is a compact set. $L(x_0)$ is also (see [4]) a convex set.

3) The existence of x^* follows from the continuity of the function f on the bounded set $L(x_0)$. From the definition of the level set it follows that

$$f(x^*) = \min_{x \in L(x_0)} f(x) = \min_{x \in D} f(x)$$

Since f is strictly convex it follows from [4] that x^* is a unique minimizer. □

Lemma 3.1. ([3]) *The following statements are equivalent:*

1. $d = 0$ *is a globally optimal solution of the problem* (3.1)*;*
2. 0 *is the optimum of the objective function of the problem* (3.1)*;*
3. *the corresponding* x_k *is a stationary point of the function* f*.*

Now we shall prove that there exists a finite $i(k)$, i.e. since d_k is defined by (3.1), that the algorithm is well-defined.

Proposition 3.2. *If* $d_k \neq 0$ *is a solution of* (3.1)*, then for any continuous function* $\sigma : [0, +\infty) \to [0, +\infty)$ *satisfying* $\delta_1 t \leq \sigma(t) \leq \delta_2 t,\ 0 < \delta_1 < \delta_2 < 1$ *there exists a finite* $i^*(k)$ *such that for all* $q^{i(k)} \in (0, q^{i^*(k)})$

$$f(x_k + q^{i(k)} d_k) - f(x_k) \leq -\frac{1}{2} q^{i(k)} \sigma(f''_D(x_k; d_k))$$

holds.

Proof. According to Lemma 9.3 from [1] and from the definition of d_k it follows that for $x_{k+1} = x_k + td_k,\ t > 0$ we have

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\leq t \nabla f(x_k)^T d_k + \frac{L}{2} t^2 \|d_k\|^2 \\ &\leq -t \frac{1}{2} f''_D(x_k; d_k) + \frac{L}{2} t^2 \|d_k\|^2 \\ &\leq -\frac{t}{2\delta_2} \sigma(f''_D(x_k; d_k)) + \frac{L}{2} t^2 \|d_k\|^2. \end{aligned} \tag{3.4}$$

If we choose $t = \sigma(f''_D(x_k; d_k))/(L\|d_k\|^2)$ and put in (3.4), we get

$$f(x_{k+1}) - f(x_k) \leq \frac{1}{2} \frac{\delta_2 - 1}{\delta_2} \frac{\sigma^2(f''_D(x_k; d_k))}{L\|d_k\|^2} = -\frac{K}{2} t \sigma(f''_D(x_k; d_k)),$$

since $(\delta_2 - 1)/\delta_2 = -K < 0$. Taking $q^{i^*(k)} = [Kt]/q$, i.e. $i^*(k) = \log_q([Kt]/q)$ we have that the claim of the theorem holds for all $q^{i(k)} \in (0, q^{i^*(k)})$. □

Convergence theorem. *Suppose that* $f \in LC^1$ *and that the assumption* (3.3) *holds. Then for any initial point* $x_0 \in D, x_k \to \bar{x}$*, as* $k \to +\infty$*, where* $\bar{x}$ *is a unique minimal point.*

Proof. If $d_k \neq 0$ is a solution of (3.1), it follows that $\Phi_k(d_k) \leq 0 = \Phi_k(0)$. Consequently, we have by the relation (3.3) that

$$\nabla f(x_k)^T d_k \leq -\frac{1}{2} f''_D(x_k; d_k) \leq -\frac{1}{2} c_1 \|d_k\|^2 < 0.$$

From the above inequality it follows that the vector d_k is a descent direction at x_k, i.e. from the relations (3.2) and (3.3) we get

$$\begin{aligned} f(x_k + q^{i(k)}d_k) - f(x_k) &\leq -\frac{1}{2}q^{i(k)}\sigma(f''_D(x_k; d_k)) \leq -\frac{1}{2}q^{i(k)}\delta_1 f''_D(x_k; d_k) \\ &\leq -\frac{1}{2}q^{i(k)}c_1\|d_k\|^2 \end{aligned}$$

for every $d_k \neq 0$. Hence the sequence $\{f(x_k)\}$ has the descent property, and, consequently, the sequence $\{x_k\} \subset L(x_0)$. Since $L(x_0)$ is by the Proposition 3.1 a compact convex set, it follows that the sequence $\{x_k\}$ is bounded. Therefore there exist accumulation points of the sequence $\{x_k\}$.

Since ∇f is by assumption continuous, then, if $\nabla f(x_k) \to 0$, $k \to +\infty$, it follows that every accumulation point $\bar{x}$ of the sequence $\{x_k\}$ satisfies $\nabla f(\bar{x}) = 0$. Since f is by the Proposition 3.1 strictly convex, there exists a unique point $\bar{x} \in L(x_0)$ such that $\nabla f(\bar{x}) = 0$. Hence, the sequence $\{x_k\}$ has a unique limit point $\bar{x}$ – and it is a global minimizer. Therefore we have to prove that $\nabla f(x_k) \to 0$, $k \to +\infty$.

There are two cases to consider:

a) The set of indices $\{i(k)\}$ for $k \in K_1$, is uniformly bounded above by a number I.

Because of the descent property, it follows that all points of accumulation have the same function value and

$$\begin{aligned} f(\bar{x}) - f(x_k) &= \sum_{k=0}^{+\infty} f(x_{k+1}) - f(x_k) \leq \sum_{k=0}^{+\infty} -\frac{1}{2}q^{i(k)}\sigma(f''_D(x_k; d_k)) \\ &\leq -\frac{1}{2}q^I\delta_1 \sum_{k\in K_1} f''_D(x_k; d_k) \leq -\frac{1}{2}q^I c_1 \sum_{k\in K_1} \|d_k\|^2. \end{aligned}$$

Since $f(\bar{x})$ is finite, it follows that $\|d_k\| \to 0 = \bar{d}$, $k \to +\infty$, $k \in K_1$.

By Lemma 2.1 it follows that $\bar{d} = 0$ is a globally optimal point of the problem (3.1) and, that the corresponding accumulation point $\bar{x}$ is a stationary point of the objective function f, i.e. $\nabla f(\bar{x}) = 0$. From the Proposition 3.1 it follows that $\bar{x}$ is a unique optimal point.

b) There is a subset $K_2 \subset K_1$ such that $\lim_{k\to+\infty} i(k) = +\infty$. By the definition of $i(k)$, we have for $k \in K_2$ that

$$f(x_k + q^{i(k)-1}d_k) - f(x_k) > -\frac{1}{2}q^{i(k)-1}\sigma(f''_D(x_k; d_k)). \tag{3.5}$$

Suppose that $\bar{x}$ is an arbitrary accumulation point of $\{x_k\}$, but not a stationary point of f. Then, from Lemma 2.1 it follows that the corresponding direction vector $d \neq 0$. Now, dividing both sides in the expression (3.5) by $q^{i(k)-1}$ and using $\lim_{k\to+\infty} q^{i(k)-1} = 0$, $k \in K_2$, we get

$$\nabla f(\bar{x})^T \bar{d} > -\frac{1}{2}\sigma(f''_D(\bar{x};\bar{d})) > -\frac{1}{2}\delta_2 f''_D(\bar{x};\bar{d}) > -\frac{1}{2} f''_D(\bar{x};\bar{d}).$$

But, from the property of the iterative function Φ_k, we have

$$\nabla f(\bar{x})^T \bar{d} \leq -\frac{1}{2} f''_D(\bar{x};\bar{d}).$$

Therefore, we get a contradiction. □

Convergence rate theorem. *Under the assumptions of the previous theorem we have that the following estimate holds for the sequence $\{x_k\}$ generated by the algorithm.*

$$f(x_n) - f(\bar{x}) \leq \mu_0 \left[1 + \mu_0 \frac{1}{\eta^2} \sum_{k=0}^{n-1} \frac{f(x_k) - f(x_{k+1})}{\|\nabla f(x_k)\|^2}\right]^{-1}, \quad n = 1, 2, \ldots$$

where $\mu_0 = f(x_0) - f(\bar{x})$, and $\mathrm{diam} L(x_0) = \eta < +\infty$ *since by Proposition* 3.1 *it follows that $L(x_0)$ is bounded.*

Proof. The proof directly follows from the Theorem 9.2, page 167, in [1]. □

4. Conclusion

In this paper we present a modification of the algorithm given in [3]. We prove that the algorithm is well-defined as well as the convergence of the algorithm by the original proofs. The estimate of the rate of convergence directly follows from [1].

Acknowledgements. This research was supported by the Serbian Ministry of Science and Environmental Protection, grant number #1389, through Institute of Mathematics, SANU.

REFERENCES

1. V.G. Karmanov: *Mathematical Programming.* Nauka, Moscow, 1975 (in Russian).
2. E. Polak: *Computational Methods in Optimization. A Unified Approach.* Mathematics in Science and Engineering, Vol. 77, Academic Press, New York – London, 1971 [Russ. trans.: Mir, Moscow, 1974].
3. Wen-yu Sun, R. J. B. de Sampaio and Jin-Yun Yuan: *Two algorithms for LC^1 unconstrained optimization.* J. Comput. Math. **18** (2000), 621–632.
4. V. Vujčić, M. Ašić and N. Miličić: *Mathematical Programming*, Institute of Mathematics, Belgrade, 1980 (in Serbian).

Department of Mathematics
Faculty of Technology and Metallurgy
University of Belgrade
Karnegijeva 4, 11000 Belgrade
Serbia and Montenegro

e-mail: `nmilicic@elab.tmf.bg.ac.yu`