On necessity of L-stationarity in Nonlinear Optimization with a Sparsity Constraint

Abbas Khademi, Majid Soleimani-damaneh*
School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran

Received: 18 June 2018 Accepted: 29 January 2018

Extended Abstract

Paper pages (387-392)

In this paper, we investigate a necessary optimality condition for a specific problem in nonlinear programming, called sparsity constrained problem. This model involves minimizing a continuously differentiable function over a sparsity constraint. We show that L-stationarity is necessary for optimality in sparsity constrained problems in general. This important property has been proved in the literature under Lipschitzness of the gradient mapping.

Preliminaries

Consider the problem

$$(\mathrm{P}): \quad \begin{cases} \min \mathrm{f}(x) \\ \mathrm{s.t.}\ \|x\|_0 \leq \mathrm{s} \end{cases}$$

in which f: $\mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function and s>0 is a given positive integer less than *n*. The zero norm of $x \in \mathbb{R}^n$ is defined as

$$\|x\|_0 \equiv \#\{i : x_i \neq 0\}.$$

Furthermore, the support of $x \in \mathbb{R}^n$ is defined by

$$I_1(x) := \{i : x_i \neq 0\}.$$

The complement of $I_1(x)$ is

$$I_o(x) := \{i : x_i = 0\}.$$

The set of feasible solutions of (P) is

$$C_s := \{x : \|x\|_0 \leq s\}.$$

The orthogonal projection operator corresponding to the nonempty closed set $D \subseteq \mathbb{R}^n$, denoted by $P_D(.)$, is defined as

$$P_D(y) := \mathrm{argmin}_{x \in D} \|y - x\|^2, \ y \in \mathbb{R}^n.$$

Necessary optimality condition

In the whole paper, we assume that the objective function of (P) in bounded blew over $\mathbb{R}^n$.

Definition. [1] The vector $x^* \in C_s$ is called a L-stationary point of (P) if

$$x^* \in P_{C_s}(x^* - \tfrac{1}{L}\nabla f(x^*)).$$

Assumption (*). The operator $\nabla f(.)$ is Lipschitz with modulus L_f on $\mathbb{R}^n$, i.e.,

$$\forall x, y \in \mathbb{R}^n : \quad \|\nabla f(x) - \nabla f(y)\| \leq L_f \|x - y\|.$$

In [1], it has been shown that under Assumption (*), each optimal solution of (P) is an L-stationary point with $L > L_f$.

Theorem 1. [1] Suppose that f is continuously differentiable, and Assumption (*) holds for $L > L_f$. If $x^*$is an optimal solution of (P), then

i) $x^*$is an L-stationary point.

ii) The set $P_{C_s}\left(x^* - \frac{1}{L}\nabla f(x^*)\right)$ is a singleton.

Theorem 2 proves the necessity of L-stationarity for optimality in Problem (P) without Lipschitz assumption. This is our main result.

Theorem 2. Assume that f is continuously differentiable. If $x^*$is an optimal solution of (P), then there exists some L>0 such that $x^*$is an L-stationary point.

Notice that L-stationary may not be sufficient for optimality. The following example clarifies the matter.

Example. Suppose that $f(x,y) = (x-10)^2 + (y-1)^2$ and s =1. Then $x^* = (0,1)$ is an L-stationary point for sufficiently large L>0. Indeed,

$$x^* - \frac{1}{L}\nabla f(x^*) = \begin{pmatrix}0\\1\end{pmatrix} - \frac{1}{L}\begin{pmatrix}-20\\0\end{pmatrix} = \begin{pmatrix}\frac{20}{L}\\1\end{pmatrix},$$

and so, for sufficiently large L>0,

$$P_{C_s}\left(x^* - \frac{1}{L}\nabla f(x^*)\right) = x^*.$$

On the other hand,

$$f(0,1) = 100 > 1 = f(10,0).$$

So, there exists some L>0 such that $x^* = (0,1)$ is an L-stationary point, while this vector is not optimal.

Keywords: Nonlinear programming, Sparsity constrained problems, L-stationarity, Optimality condition

*Corresponding author: soleimani@khayam.ut.ac.ir