Towards Performance Estimation Problems on Quadratic Functions

Nizar Bousselmi, Julien M. Hendrickx and François Glineur
ICTEAM, UCLouvain, 1348 Louvain-la-Neuve, Belgium.
Email: {nizar.bousselmi, julien.hendrickx, francois.glineur}@uclouvain.be

1 Introduction

The *Performance Estimation Problem* (PEP) methodology has been recently introduced [1, 2] to compute the exact worst-case performance of a first-order optimization method on a given class of function, e.g. L-smooth μ-strongly convex functions. In this work, we propose to solve PEP for the class of quadratic functions of the form $f(x) = \frac{1}{2}x^T Qx$ with $\mu I \preceq Q \preceq LI$ for given parameters μ and L, i.e. L-smooth μ-strongly convex homogeneous quadratic functions.

2 Problem statement

PEP can be formulated as semidefinite programs where the matrix variable is the $2N \times 2N$ Gram matrix $G = P^T P$ with $P = \begin{pmatrix} x_1 & \cdots & x_N & g_1 & \cdots & g_N \end{pmatrix}$. In the PEP context, the x_i's and g_i's are, respectively, the iterates and the gradient at the iterates, produced by the N iterations of the given method. However, in this work, we will consider for simplicity that the x_i's and g_i's are just given.

We define a Gram matrix associated to a quadratic function.

Definition 1. A symmetric matrix $G \in \mathbb{S}^{2N}$ is a (μ, L, N)-*quadratic-Gram matrix* if and only if there exist a dimension $d \in \mathbb{N}$, a symmetric matrix $Q \in \mathbb{S}^d$ with $\mu I \preceq Q \preceq LI$ and a sequence $x_i \in \mathbb{R}^d$ for $i = 1, \ldots, N$ such that $G = P^T P$ with

$$P = \begin{pmatrix} x_1 & \cdots & x_N & \overbrace{Qx_1}^{g_1} & \cdots & \overbrace{Qx_N}^{g_N} \end{pmatrix} \in \mathbb{R}^{d \times 2N}. \tag{1}$$

The set of all (μ, L, N)-quadratic-Gram matrices is denoted $\mathscr{G}_{\mu,L,N}$. It can be shown that any conic combination of (μ, L, N)-quadratic-Gram matrices is also a (μ, L, N)-quadratic-Gram matrix, thus, the set $\mathscr{G}_{\mu,L,N}$ is a convex cone.

Theorem 1. The set of all (μ, L, N)-quadratic-Gram matrices $\mathscr{G}_{\mu,L,N}$ is a convex cone.

Since the set $\mathscr{G}_{\mu,L,N}$ is convex, we seek an explicit convex description of it in order to add the constraints to PEP.

3 Case with $N = 1$ point

First, we look at the case $N = 1$ where we only have one point x_1 and its gradient g_1. It can be shown that the set of $(\mu, L, 1)$-quadratic-Gram matrices of one point is exactly described by three convex inequalities.

Theorem 2. Given a symmetric matrix $G = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix}$, the conditions

$$g_{11} \geq 0 \tag{2}$$

$$g_{12}^2 \leq g_{11} g_{22} \tag{3}$$

$$g_{22} \leq -\mu L g_{11} + (\mu + L) g_{12} \tag{4}$$

are necessary and sufficient conditions for $G \in \mathscr{G}_{\mu,L,1}$.

4 Case with N points

Now, we consider the general case $N \geq 1$. First of all, we can write a quadratic-Gram matrix under the following form

$$G = \begin{pmatrix} X^T X & X^T QX \\ X^T QX & X^T Q^2 X \end{pmatrix} \tag{5}$$

where $X = \begin{pmatrix} x_1 & \cdots & x_N \end{pmatrix} \in \mathbb{R}^{d \times N}$. In addition to the global symmetry and positive semidefiniteness of G, hence of diagonal blocks $X^T X$ and $X^T Q^2 X$, we observe that off diagonal block $X^T QX$ is also symmetric and positive semidefinite. Actually, it is possible to find several necessary characterizations of a quadratic-Gram matrix (see for example Theorem 3 in [3] in a different context). However, it appears to be much less straightforward to check or prove the sufficiency, which is the question we are investigating.

5 Conclusion and perspectives

Theorem 1 ensures that the set of quadratic-Gram matrices is convex. Therefore it is likely that there exists a way to describe the set with an explicit list of convex constraints. Once these constraints are identified and we are able to solve PEP on quadratic functions, it will be possible to measure and quantify the gap between the worst-case performance of a given first-order optimization method on general L-smooth μ-strongly convex functions and on quadratic functions.

N. Bousselmi is supported by the French Community of Belgium through a FRIA fellowship (F.R.S.-FNRS).

References

[1] A.B. Taylor, J.M. Hendrickx, F. Glineur *Smooth strongly convex interpolation and exact worst-case performance of first-order methods*. Mathematical Programming, 2017, vol. 161, no 1, p. 307-345.

[2] Y. Drori, M. Teboulle *Performance of first-order methods for smooth convex minimization: a novel approach*. Mathematical Programming, 2014, vol. 145, no 1, p. 451-482.

[3] S. Colla, J.M. Hendrickx *Automated Worst-Case Performance Analysis of Decentralized Gradient Descent*. 2021