A Geometric Structure of Acceleration and Its Role in Making Gradients Small Fast

Jongmin LEE \(^1\), Chanwoo PARK \(^2\) and Ernest K. RYU \(^1\)

1) Department of Mathematical Sciences, Seoul National University, Seoul 08826, KOREA
2) Department of Statistics, Seoul National University, Seoul 08826, KOREA

Corresponding Author: Ernest K. RYU, ernestryu@snu.ac.kr

ABSTRACT

The many accelerated first-order methods since Nesterov’s seminal 1983 work have been presented without a common underlying structure. In this work, we identify a geometric structure satisfied by a wide range of first-order accelerated methods. Using this geometric insight, we present several novel generalizations of accelerated methods. Most interesting among them is a method that can reduce the squared gradient norm with \(O(1/K^4)\) rate in the prox-grad setup, faster than the \(O(1/K^3)\) rates of Nesterov’s FGM or Kim and Fessler’s FPGM-m.