

A Differential Equation for Modeling Nesterov's Accelerated Gradient Method: Theory and Insights

W. Su, S. Boyd, and E. Candes

Journal of Machine Learning Research, 17(153):1-43, September 2016.

Shorter version appeared in *Proceedings Neural and Information Processing Systems*, December 2014.

- [JMLR paper](#)
- [NIPS paper](#)

We derive a second-order ordinary differential equation (ODE), which is the limit of Nesterov's accelerated gradient method. This ODE exhibits approximate equivalence to Nesterov's scheme and thus can serve as a tool for analysis. We show that the continuous time ODE allows for a better understanding of Nesterov's scheme. As a by-product, we obtain a family of schemes with similar convergence rates. The ODE interpretation also suggests restarting Nesterov's scheme, leading to an algorithm which can be rigorously proven to converge at a linear rate whenever the objective is strongly convex.