

SEMINARIO

OPTIMIZACIÓN Y EQUILIBRIO

EXPOSITOR

Juan Peypouquet

UFSM, Chile

TITULO

Recent advances on the acceleration of first-order methods in convex optimization

Abstract:

Let $f : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper lower-semicontinuous convex function defined on a Hilbert space H (think of $\mathbb{R}^N$), and let (x^k) be a sequence in H generated by means of a “typical” first-order method, and intended to minimize f . If $\operatorname{argmin}(f) \neq \emptyset$, then (x^k) will converge weakly, as $k \rightarrow +\infty$, to a minimizer of f , with a worst-case theoretical convergence rate of

$$f(x^k) - \min(f) = O(1/k).$$

In the 1980’s Y. Nesterov came up with a revolutionary - yet remarkably simple! - idea to modify the computation of the iterates with essentially the same computational cost, in order to improve this theoretical convergence rate down to $O(1/k^2)$. For three decades, researchers tried unsuccessfully to prove the (weak) convergence of the resulting sequence. Despite this drawback, Nesterov’s scheme is current common practice, especially in image processing and statistics, and systematically converges faster than the unaccelerated counterpart.

In this talk, we revisit Nesterov’s acceleration scheme, interpreting it as a time discretization of a second-order differential equation. It turns out that this analogy brings a deeper insight into this procedure. In particular, it is possible to prove that an “infinitesimal” variant of the accelerated method yields sequences which are not only convergent, but exhibit a convergence rate of

$f(x^k) - \min(f) = o(1/k)$. The (weak) convergence of the original method of Nesterov remains a puzzling open question.

Miércoles 06 de Abril a las 16:30 hrs, Sala de Seminario John Von Neumann CMM, séptimo piso, torre norte, Beauchef 851.