Abstracts

Convergence Rates of First and Higher Order Dynamics for Solving Linear Inverse Problems

Otmar Scherzer
(joint work with Radu Boţ & Guozhi Dong & Peter Elbau)

We consider the problem of solving a linear inverse problem, formulated as solving an operator equation

$$Lx = y, \tag{1}$$

where $L : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator between (infinite dimensional) real Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$. If the range of L is not closed, Equation 1 is ill-posed and regularization has to be employed for a stable solution.

In this talk we consider dynamical regularization methods: That is, we approximate the minimum norm solution $x^\dagger$ of Equation 1 by the solution of a dynamical system at an appropriate time. An established example of such a dynamical regularization method is Showalter's method [6], which consists in calculating the solution of the Cauchy problem

$$\begin{aligned} &\xi'(t) = -L^*L\xi(t) + L^*y \text{ for all } t \in (0,\infty),\\ &\xi(0) = 0. \end{aligned} \tag{2}$$

More recently, second order dynamical systems have been investigated for optimizing **general** convex functionals, see [7, 4, 1, 2]. One motivation for these dynamic equations has been to consider them as time continuous limits of Nesterov's algorithm [5] to explain its fast convergence.

We consider dynamical systems for solving linear ill-posed operator equations by focusing on the particular convex functionals

$$\mathcal{J}(x) = \frac{1}{2}\|Lx - y\|^2. \tag{3}$$

Specifically, we develop a regularization theory to analyse N-th order dynamical method of the form

$$\begin{aligned} \xi^{(N)}(t) + \sum_{k=1}^{N-1} a_k(t)\xi^{(k)}(t) &= -L^*L\xi(t) + L^*y \quad &&\text{for all } t \in (0,\infty),\\ \xi^{(k)}(0) &= 0 &&\text{for all } k = 0,\dots,N-1, \end{aligned} \tag{4}$$

where $N \in \mathbb{N}$ and $a_k : (0,\infty) \to \mathbb{R}, k = 1,\dots,N-1$, are continuous functions. When $N = 1$ this is Showalter's method. When $N = 2$, and a_1 is constant and positive the method is called *heavy ball dynamical method* (HBD). And for $N = 2$ and $a_1 = \frac{b}{t}$, $b > 0$, the analogous method as considered in [7, 4, 1, 2], we call it the *vanishing viscosity flow* (VVF).

In [3] we proved the following convergence rates result and compared it with the literature, see Table 1.

Method ($t \to \infty$)	$\mathcal{J}(\xi(t)) - \min \mathcal{J}$	$\|\xi(t) - x^\dagger\|^2$
VVD (convex)	$o(t^{-2})$ $(b > 3)$, $\mathcal{O}(t^{-2})$ $(b = 3)$, $\mathcal{O}(t^{-2b/3})$ $(b < 3)$	–
VVD (IP)	$\boldsymbol{o(t^{-2})}$ $\boldsymbol{(b > 2)}$, $\boldsymbol{o(t^{-b})}$ $\boldsymbol{(b \leq 2)}$	$\boldsymbol{o(1)}$
Showalter	$o(t^{-1})$	$o(1)$
HBD	$o(t^{-1})$	$o(1)$

Table 1. Convergence Rates without Source Conditions. Convergence for VVD from [7, 4, 1, 2] for minimizing general convex functionals $\mathcal{J}$, and results from [3] (IP) for the convex functionals from Equation 3. Analogous results for Showalter and HBD. Results in bold face are from [3].

Method ($t \to \infty$)	$\mathcal{J}(\xi(t)) - \min \mathcal{J}$	$\|\xi(t) - x^\dagger\|^2$
VVD (IP) [Max Rate]	$\boldsymbol{\mathcal{O}(t^{-2\mu-2})}$ $[\boldsymbol{\mathcal{O}(t^{-b+\epsilon})}]$	$\boldsymbol{\mathcal{O}(t^{-2\mu})}$ $[\boldsymbol{\mathcal{O}(t^{-b+\epsilon})}]$
Showalter	$\mathcal{O}(t^{-\mu-1})$	$\mathcal{O}(t^{-\mu})$
HBD	$\mathcal{O}(t^{-\mu-1})$	$\mathcal{O}(t^{-\mu})$

Table 2. Convergence rates with source conditions. In the case of VVD, the parameters are restricted to $0 < \mu < \frac{b}{2} - 1$ (and thus $b > 2$) for $\mathcal{J}(\xi(t)) - \min \mathcal{J}$ and $0 < \mu < \frac{b}{2}$ for $\|\xi(t) - x^\dagger\|^2$, which leads to the given maximal rates (for arbitrarily small $\epsilon > 0$). For general convex problems, source conditions, Equation 5, are not known to provide improved convergence rates.

Note that the results from convex analysis (VVD (convex) in Table 1) prove convergence rates of the residuum and not of the solution. The results of (IP) show an improvement of the convergence rates for general convex problems. Actually these results support a conjecture of [7]: "However, from a different perspective, this example suggests that $\mathcal{O}(t^{-b})$ convergence rate can be expected ..."

Moreover, we showed that under classical source conditions even convergence rates for $\|\xi(t) - x^\dagger\|^2$ can be proven. In this case we require source conditions like: There exists some $w \in \mathcal{X}$ such that

$$x^\dagger = (L^* L)^{\frac{\mu}{2}} w. \tag{5}$$

We summarize some results from [3] on regularized dynamical systems when the solution $x^\dagger$ satisfies some source condition in Table 2.

Acknowledgments

- RB acknowledges support from the Austrian Science Fund (FWF) within the project I2419-N32 (Employing Recent Outcomes in Proximal Theory Outside the Comfort Zone).
- GD is supported by a MATHEON Research Center project CH12 funded by the Einstein Center for Mathematics (ECMath) Berlin.
- PE and OS are supported by the Austrian Science Fund (FWF), with SFB F68, project F6804-N36 (Quantitative Coupled Physics Imaging) and project F6807-N36 (Tomography with Uncertainties).
- OS acknowledges support from the Austrian Science Fund (FWF) within the national research network Geometry and Simulation, project S11704 (Variational Methods for Imaging on Manifolds) and I3661-N27 (Novel Error Measures and Source Conditions of Regularization Methods for Inverse Problems).

References

[1] H. Attouch, Z. Chbani, J. Peypouquet, and P. Redont. Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity. *Mathematical Programming. A Publication of the Mathematical Optimization Society*, 168(1–2):123–175, 2018.

[2] H. Attouch, Z. Chbani, and H. Riahi. Combining fast inertial dynamics for convex optimization with Tikhonov regularization. *Journal of Mathematical Analysis and Applications*, 457(2):1065–1094, 2018.

[3] R. Boţ, G. Dong, P. Elbau, and O. Scherzer. Convergence rates of first and higher order dynamics for solving linear ill-posed problems. Preprint on ArXiv 1812.09343, University of Vienna, Austria, 2018.

[4] R. I. Boţ and E. R. Csetnek. Second order forward-backward dynamical systems for monotone inclusion problems. *SIAM Journal on Control and Optimization*, 54(3):1423–1443, 2016.

[5] Yu. E. Nesterov. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Doklady Akademii Nauk SSSR*, 269(3):543–547, 1983.

[6] D. Showalter. Representation and computation of the pseudoinverse. *Proceedings of the American Mathematical Society*, 18:584–586, 1967.

[7] W. Su, S. Boyd, and E. Candès. A differential equation for modeling nesterov's accelerated gradient method: Theory and insights. *Journal of Machine Learning Research (JMLR)*, 17(153):1–43, 2016.

Reporter: