Iterative regularization of nonsmooth ill-posed problems

Christian Clason Vu Huu Nhu

Fakultät für Mathematik, Universität Duisburg-Essen

Inverse and Ill-Posed Problems: Theory and Numerics
Novosibirsk, October 9, 2020

Motivation

Inverse problem: find

- unknown parameter $u^\dagger$
  e.g., heat source, diffusion constant, thermal conductivity, heat capacity, latent heat density, …

given

- measurement y
- model $S : u \mapsto y$, e.g., solution of PDE

$\rightsquigarrow$ solve

$$S(u) = y$$

Problem: measurement $y = y^\delta$ noisy, range of S not closed

$\rightsquigarrow$ ill-posed, needs regularization

Regularization

Solve approximate, stable problem:

1. Tikhonov regularization $\rightsquigarrow$ optimal control

2. iterative regularization, e.g., Landweber iteration

$$u_{n+1} = u_n + w_n S'(u_n)^*(y^\delta - S(u_n)) \qquad n = 1, \ldots, N$$

- $S'(u)$ Fréchet derivative, $S'(u)^*$ adjoint
- stopping index $N = N(\delta) < \infty$ (regularization parameter)
- regularization: $N(\delta) \to \infty, \quad u_{N(\delta)} \to u^\dagger \quad$ for $\delta \to 0$

Here: S solution operator for non-smooth PDE

$\rightsquigarrow$ not Fréchet differentiable

Non-smooth equation

$$-\Delta y + \max\{0, y\} = u$$

solution operator $S : u \mapsto y \;\; (:= S(u))$

- well-posed (in suitable – standard – spaces)
- Lipschitz continuous
- completely continuous ($\rightsquigarrow$ ill-posed)
- not Fréchet differentiable unless $|\{x : y(x) = 0\}| = 0$
- model for membrane under water, plasma MHD equilibrium
- can be extended to arbitrary $f(y)$ piecewise differentiable
- simplified model for sharp phase transition (Stefan problem)

Non-smooth equation

$$-\Delta y + \max\{0, y\} = u$$

solution operator $S : u \mapsto y \ \ (:= S(u))$

- not Fréchet differentiable unless $|\{x : y(x) = 0\}| = 0$
- but: directionally differentiable

directional derivative $S'(u; h) =: \eta$ solves

$$-\Delta\eta + \mathbb{1}_{\{y=0\}} \max\{0, \eta\} + \mathbb{1}_{\{y>0\}}\eta = h$$

not linear in $h \rightsquigarrow$ not useful for algorithm

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

Bouligand subdifferential

$$\partial_B S(u) := \left\{ G \text{ linear} \;\middle|\; \begin{array}{l} \text{there is } \{u_n\} \text{ Gâteaux differentiable with} \\ u_n \to u \text{ and } S'(u_n; h) \to G\,h \text{ for all } h \end{array} \right\}$$

special choice: $G_u : h \mapsto \eta$ solving (for $y = S(u)$)

$$-\Delta\eta + \mathbb{1}_{\{y>0\}}\eta = h$$

- $G_u \in \partial_B S(u)$ Bouligand derivative [Christof/Meyer/Walter/C.]
- $u \mapsto G_u$ uniformly bounded (in right spaces)
- linear $\rightsquigarrow$ use for Landweber in place of $S'(u)$

Bouliguand–Landweber iteration

$$u_{n+1}^\delta = u_n^\delta + w_n G_{u_n^\delta}^* \left(y^\delta - S(u_n^\delta)\right) \qquad n = 0, 1, \dots, N(\delta)$$

- $S : u \mapsto y$ non-smooth
- y^δ with $\|y^\delta - y^\dagger\| \leqslant \delta, \quad y^\dagger = S(u^\dagger)$ (assume unique)
- $u_0^\delta = u_0$ starting point
- w_n step sizes
- stopping index $N(\delta)$ by Morozov discrepancy principle:

$$\|y^\delta - S(u_{N(\delta)}^\delta)\| \leqslant \tau\delta < \|y^\delta - S(u_n^\delta)\| \qquad 0 \leqslant n < N(\delta)$$

(modified Landweber iteration [Scherzer '95])

Well-posedness

UNIVERSITÄT DUISBURG ESSEN Open-Minded

Assume:

1. $\{G_u\}$ uniformly bounded
2. generalized tangential cone condition (GTCC)

   $$\|S(u') - S(u) - G_u(u' - u)\| \leqslant \mu \|S(u') - S(u)\| \quad \text{for all } u, u' \in B_\rho(u^\dagger)$$

   non-smooth PDE: satisfied for $1 > \mu > C(|\{x : y^\dagger(x) = 0\}|)$
3. $u_0 \in B_\rho(u^\dagger)$

Then (under conditions on μ, τ, w_n):

- $u_n^\delta \in B_\rho(u^\dagger)$ for all $n \leqslant N(\delta)$
- $\delta > 0$: $N(\delta) < \infty$ and $\|u_n^\delta - u^\dagger\| < \|u_{n-1}^\delta - u^\dagger\|$ for $n \leqslant N(\delta)$
- $\delta = 0$: $N(\delta) = \infty$ and $u_n^0 \to u^\dagger$ for $n \to \infty$

Regularization

Goal: show that $u^\delta_{N(\delta)} \to u^\dagger$ for $\delta \to 0$

Standard proof: combine

1. monotonicity: $\|u^\delta_n - u^\dagger\| < \|u^\delta_{n-1} - u^\dagger\| \quad$ for $n \leqslant N(\delta)$
2. stability: $u^\delta_n \to u^0_n \quad$ for all $n = 1, \ldots$

Problem:

- stability requires continuity of $u \mapsto G_u$
- $u \mapsto G_u$ not continuous for non-smooth S
- $\rightsquigarrow$ use asymptotic stability

Asymptotic stability

Definition

Iterative method generating $\{u_n^\delta\}_{n\leqslant N(\delta)}$ asymptotically stable for $\delta \to 0$ if subsequence $\{\delta_k\}$ exists with

- for all $0 \leqslant n \leqslant \overline{N} := \lim_{k\to\infty} N(\delta_k) \in \mathbb{N} \cup \{\infty\}$

$$u_n^{\delta_k} \to \tilde{u}_n \quad \text{as } k \to \infty$$

  for some $\tilde{u}_n \in \overline{B}_U(u^\dagger, \rho)$

- if $\overline{N} = \infty$,

$$\tilde{u}_n \to u^\dagger \quad \text{as } n \to \infty$$

- $\tilde{u}_n$ generated by perturbed noise-free iteration
- perturbation needs to vanish for $n \to \infty$

Asymptotic stability

Bouligand–Landweber iteration

$$u_{n+1}^{\delta} = u_n^{\delta} + w_n G_{u_n^{\delta}}^{*} \left(y^{\delta} - S(u_n^{\delta}) \right), \qquad n = 0, 1, 2, \ldots, N(\delta)$$

- asymptotically stable for non-smooth PDE
  (proof uses GTCC and compact embedding of $\mathcal{R}(G_u^*)$)
- $\rightsquigarrow$ regularization (under conditions on μ, τ, w_n):

$$u_{N(\delta)}^{\delta} \to u^{\dagger} \quad \text{for} \quad \delta \to 0$$

but: $N(\delta) = \mathcal{O}(\delta^{-2})$

Levenberg–Marquardt iteration

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

$$u_{n+1}^\delta = \operatorname{argmin}_{u\in D(S)} \left\|S'(u_n^\delta)(u-u_n^\delta) - y^\delta - S(u_n^\delta)\right\|^2 + \alpha_n \left\|u-u_n^\delta\right\|^2$$
$$= u_n^\delta + \left(\alpha_n I + S'(u_n^\delta)^* S'(u_n^\delta)\right)^{-1} S'(u_n^\delta)^* \left(y^\delta - S(u_n^\delta)\right)$$

- $\alpha_n = \alpha_0 r^n, r < 1$ [Kaltenbacher et al. '08]
- stopping by Morozov discrepancy principle [Q. Jin '10]
- TCC + transfer operator property

$$S'(u_2) = Q(u_1,u_2)S'(u_1) \qquad Q \text{ linear, near identity}$$

- $\rightsquigarrow$ stable, convergent regularization
- $N(\delta) = \mathcal{O}\left(1 + |\log\delta|\right)$

Bouliganad–Levenberg–Marquardt iteration

$$u_{n+1}^\delta = \operatorname{argmin}_{u\in D(S)} \|G_{u_n^\delta}(u-u_n^\delta) - y^\delta - S(u_n^\delta)\|^2 + \alpha_n\|u-u_n^\delta\|^2$$
$$= u_n^\delta + \left(\alpha_n I + G_{u_n^\delta}^* G_{u_n^\delta}\right)^{-1} G_{u_n^\delta}^* \left(y^\delta - S(u_n^\delta)\right)$$

- $\alpha_n = \alpha_0 r^n, r < 1$
- stopping by discrepancy principle
- GTCC + transfer operator property (holds for non-smooth PDE)

$$G_{u_2} = Q(u_1, u_2) G_{u_1} \qquad Q \text{ linear, near identity}$$

- $\rightsquigarrow$ asymptotically stable, convergent regularization
- $N(\delta) = \mathcal{O}\left(1 + |\log \delta|\right)$

Numerical example

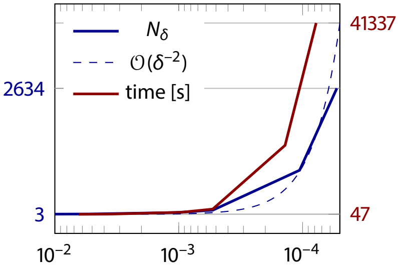
UNIVERSITÄT DUISBURG ESSEN

Open-Minded

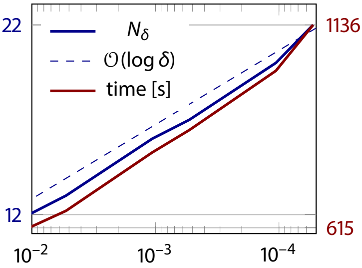
$$-\Delta y + \max\{0, y\} = u$$

- finite element discretization
- semismooth Newton method for solution (evaluation of S)
- constructed exact solution $u^\dagger$ with $|\{x : y^\dagger(x) = 0\}| > 0$
- random Gaussian noise: $\|y^\delta - y^\dagger\| = \delta$
- $\mu = 0.1, \quad \tau = 1.4, \quad \rho = 5, \quad w_n = \frac{2-2\mu}{\bar{L}^2}, \quad \bar{L} = 5 \times 10^{-2}$
- compare two starting points:
  1. $u_0 \equiv 0$
  2. $\bar{u}_0$ satisfying $u^\dagger - u_0 \in \mathcal{R}(G^*_{u^\dagger})$ (generalized source condition)

Numerical example: results with u_0

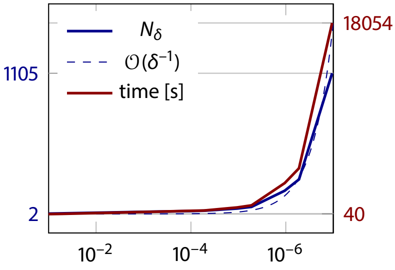


(a) Bouligand–Landweber

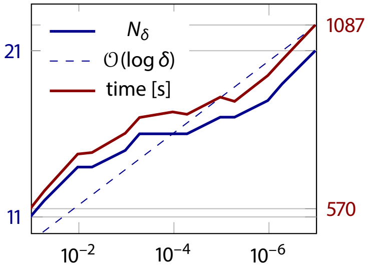


(b) Bouligand–Levenberg–Marquardt

Numerical example: results with $\bar{u}_0$



(a) Bouligand–Landweber



(b) Bouligand–Levenberg–Marquardt

Alternatives

Alternative approaches:

1. Nesterov acceleration of Bouligand–Landweber
2. Bouligand–Newton method

   $\rightsquigarrow$ Newton step ill-posed:

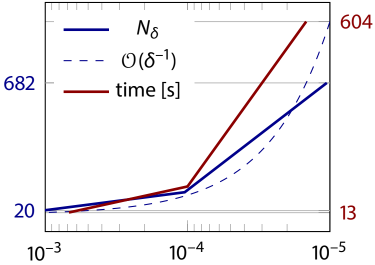
   1. iterative regularization of Newton step
   2. Tikhonov regularization of Newton step

Nesterov–Bouligand–Landweber

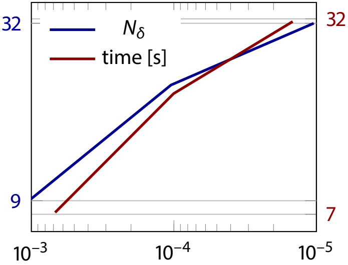
$$u_{n+1}^\delta = \hat{u}_n^\delta + w_n G_{\hat{u}_n^\delta}^* \left(y^\delta - S(\hat{u}_n^\delta)\right)$$
$$\hat{u}_{n+1}^\delta = u_{n+1}^\delta + \tfrac{n-1}{n+2}(u_{n+1}^\delta - u_n^\delta)$$

- Nesterov acceleration of gradient descent [Neubauer '17, Hubmer/Ramlau '18]
- stopping by discrepancy principle
- smooth case: $N(\delta) = \mathcal{O}(\delta^{-1})$
- but asymptotic stability unclear

Nesterov–Bouligand–Landweber



(a) $u_0 = 0$



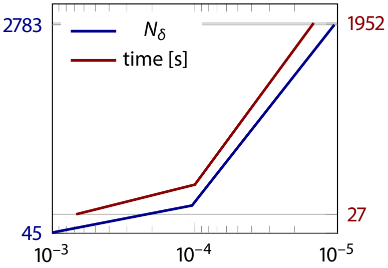
(b) $u_0 = \bar{u}_0$

Nesterov–Bouligand–Newton

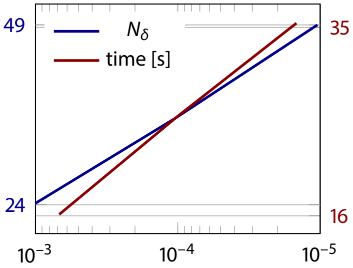
$$
\begin{aligned}
s_{k+1} &= \hat{s}_k + G^*_{u_n^\delta}\left(y^\delta - S(u_n^\delta) - G_{u_n^\delta}\hat{s}_k\right)\\
\hat{s}_{k+1} &= s_{k+1} + \tfrac{k-1}{k+2}(s_{k+1} - s_k)\\
&\dots\\
u_{n+1}^\delta &= u_n^\delta + s_{K_n}
\end{aligned}
$$

- Bouligand–Newton iteration $G_{u_n^\delta} s = y^\delta - S(u_n^\delta), \quad u_{n+1}^\delta = u_n^\delta + s$
- regularized solution by Nesterov-accelerated gradient method
- outer iteration: $N(\delta)$ stopped by discrepancy principle
- inner iteration: K_n stopped by linearized discrepancy principle (inner residual $<$ μ outer residual)
- count total number of Nesterov steps $N_\delta := \sum_n K_n$

Nesterov–Bouligand–Newton



(a) $u_0 = 0$



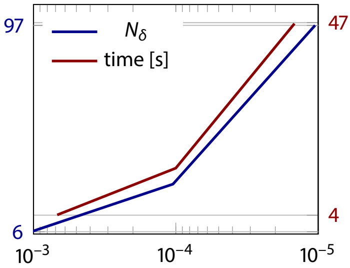
(b) $u_0 = \bar{u}_0$

CG–Bouligand–Newton

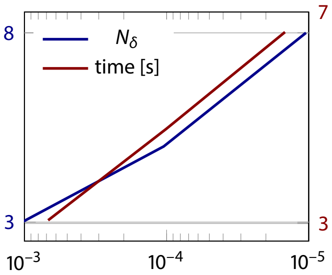
$$
\begin{aligned}
s_{k+1} &= \alpha_k \hat{s}_k \\
\hat{s}_{k+1} &= \beta_k \hat{s}_k + y^\delta - S(u_n^\delta) - G_{u_n^\delta} s_{k+1} \\
&\dots \\
u_{n+1}^\delta &= u_n^\delta + s_{K_n}
\end{aligned}
$$

- Bouligand–Newton iteration $G_{u_n^\delta} s = y^\delta - S(u_n^\delta), \quad u_{n+1}^\delta = u_n^\delta + s$
- regularized solution by conjugate gradient method (optimal two-point method for s.p.d. linear systems)
- outer iteration: $N(\delta)$ stopped by discrepancy principle
- inner iteration: K_n stopped by linearized discrepancy principle (inner residual $<$ μ outer residual)
- count total number of CG steps $N_\delta := \sum_n K_n$

CG–Bouligand–Newton



(a) $u_0 = 0$



(b) $u_0 = \bar{u}_0$

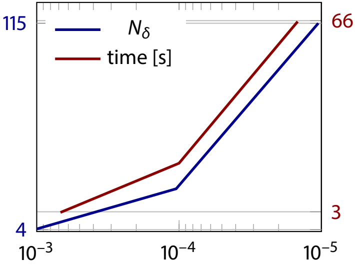
Tikhonov–Bouligand–Newton

$$s_{K_n} \approx \operatorname{argmin}_s \frac{1}{2}\|G_{u_n^\delta} s + S(u_n^\delta) - y^\delta\|^2 + \frac{\alpha_n}{2}\|s + u_n^\delta\|^2$$
$$u_{n+1}^\delta = u_n^\delta + s_{K_n}$$

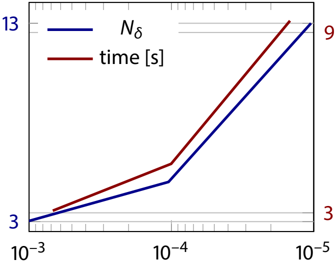
- Bouligand–Newton iteration $G_{u_n^\delta} s = y^\delta - S(u_n^\delta), \quad u_{n+1}^\delta = u_n^\delta + s$
- regularized solution by Tikhonov regularization

  $\rightsquigarrow$ Iteratively Regularized Bouligand–Gauß–Newton Method [Kaltenbacher et al. '97, '98]
- outer iteration: stopped by discrepancy principle
- inner iteration: CG, stopped by linearized discrepancy principle
- choice $\alpha_n(\mu, \tau, \rho)$ (residual norm)
- count total number of CG steps $N_\delta := \sum_n K_n$

Tikhonov–Bouligand–Newton



(a) $u_0 = 0$



(b) $u_0 = \bar{u}_0$

Conclusion

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

Summary

- iterative regularization using Bouligand derivatives
- inverse source problems for non-smooth PDEs
- convergence under asymptotic stability

Outlook

- convergence rates under source condition
- Tikhonov–Bouligand–Newton method
- other non-smooth equations, variational inequalities
- coefficient inverse problems

Preprints, Python/Julia codes:

http://www.uni-due.de/mathematik/agclason/clason_pubs.php