Iterative regularization for an inverse problem in plasma physics

Mini-course/workshop on the application of computational mathematics to plasma physics

Rommel R. Real

[1]Mathematical Sciences Institute, The Australian National University

June 25, 2019

Outline

Plasma physics example

Consider an open bounded subset $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2\}$, with a Lipschitz boundary $\partial\Omega$, and consider the nonsmooth semilinear equation

$$\begin{cases} -\Delta y + y^+ = u & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

with $u \in L^2(\Omega)$ and $y^+(x) := \max(y(x), 0)$ for almost every $x \in \Omega$.

Applications:

1. deflection of a stretched thin membrane partially covered by water [Kikuchi et.al.,'84];
2. free boundary problems for a confined plasma [Kikuchi et.al.,'84;Rappaz,'84; Temam,'75]

Consider $-\Delta v = \lambda f(v)$ in Ω, with $v = -\alpha$ on $\partial\Omega$, with $f(u) = u^+$. Then let $y = v/\alpha$.

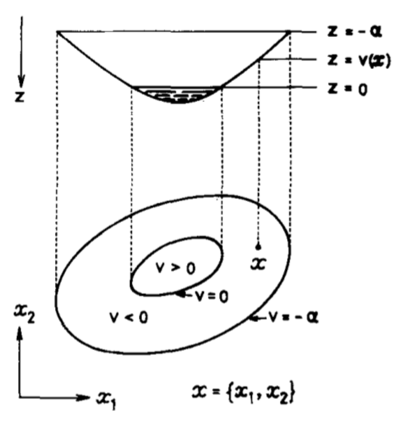

FIGURE: A stretched thin membrane partially covered with water [Kikuchi et.al.,'84]

Direct v inverse

J.B. Keller gave the following description (*J.B. Keller. Inverse problems. Amer. Math. Monthly, 83:107-118,1976*):

"We call two problems inverses of one another if the formulation of each involves all or part of the solution of the other. Often, for historical reasons, one of the two problems has been studied extensively for some time, while the other is newer and not so well understood."

In real world application, we may use the following criterion:

- Direct problems: given the causes and determine the result
- Inverse problems: from the observed effect to determine the causes

A more effective way to distinguish the two: ill-posedness

Definition 2.1

A problem is called **well-posed** in the sense of Hadarmard if

- the problem has a solution; and
- it has a unique solution; and
- the solution depends continuously on the data.

If one of the requirements is violated, the problem is called **ill-posed**.

If two problems are inverse to each other, we will call the well-posed one as the direct problem, and the ill-posed one as inverse problem.

Framework

Let X, Y be Hilbert spaces with inner products $(\cdot,\cdot)$ and norms $\|\cdot\|$, resp; $F : \mathcal{D}(F) \to Y$, with $\mathcal{D}(F) \subset X$. Consider the nonlinear operator equation

$$F(x) = y. \tag{3.1}$$

Assume (3.1) has a solution x_* (which may not be unique), we have noisy data y^δ with

$$\left\|y^\delta - y\right\| \le \delta.$$

Consider ill-posed problems of the form (3.1).
To obtain reasonable approximations to x_*, we employ regularization in solving (3.1).

Regularization

Regularization: recover the solution to an ill-posed problem by solving a series of approximating well-posed problem
Landweber method: starting with an initial guess x_0, solve for

$$x_{k+1}^{\delta} = x_k^{\delta} - F'(x_k^{\delta})^*(F(x_k^{\delta}) - y^{\delta}) \tag{3.2}$$

and terminate up to an appropriate index (to be determined later).
Notice this is just the gradient method for minimizing $\left\|F(x) - y^{\delta}\right\|^2/2$!

Differentiability

Definition 4.1

Let X, Y be Banach spaces. A function $F : X \to Y$ is said to be *Gâteaux differentiable* at $x \in X$ if there exists a bounded linear operator $dF(x)$ such that $\forall v \in X$,

$$\lim_{\alpha \to 0} \frac{F(x+\alpha v) - F(x)}{\alpha} = dF(x)v.$$

We call $dF(x)$ the *Gâteaux derivative* of F at x. If for some fixed v the limit

$$\delta_v F(x) := \frac{d}{d\alpha}\bigg|_{\alpha=0} F(x+\alpha v) = \lim_{\alpha \to 0} \frac{F(x+\alpha v) - F(x)}{\alpha}$$

exists, we say F has a *directional derivative* at x in the direction v.

$$\delta_v F(x) := \frac{d}{d\alpha}\bigg|_{\alpha=0} F(x+\alpha v) = \lim_{\alpha\to 0} \frac{F(x+\alpha v) - F(x)}{\alpha}$$

If the limit (in the sense of the Gâteaux derivative) exists *uniformly* in v on the unit sphere of X, we say F is *Fréchet differentiable* at x and $dF(x)$ is the *Fréchet derivative* of F at x. Equivalently, if we set $h = \alpha v$ then $\alpha \to 0$ iff $h \to 0$. Thus F is *Fréchet differentiable* at x if

$$F(x+h) - F(x) - dF(x)h = o(\|h\|)$$

and we call $dF(x)$ the derivative of F at x.

N.B. The limit in the Fréchet case only depends on the norm of h.

Assumptions

1. There is $\rho > 0$ such that x_* is a solution of (3.1) in $\mathcal{B}_{\rho/2}(x_0)$;
2. The Fréchet derivative $F'(\cdot)$ of F is locally uniformly bounded in a neighborhood of x_0, i.e., $\|F'(x)\| \leq 1$ for $x \in \mathcal{B}_\rho(x_0)$;
3. Tangential cone condition: for $x, \tilde{x} \in \mathcal{B}_\rho(x_0)$ and $\eta < 1/2$,

$$\|F(x) - F(\tilde{x}) - F'(x)(x - \tilde{x})\| \leq \eta \, \|F(x) - F(\tilde{x})\| \,.$$

Convergence result

Discrepancy principle: terminate the Landweber iteration (3.2) after $k_* = k_*(\delta)$ steps when

$$\left\| y^\delta - F(x_{k_*}^\delta) \right\| \le \tau\delta < \left\| y^\delta - F(x_k^\delta) \right\|, \quad 0 \le k < k_*, \tag{4.1}$$

where τ is a positive number depending on η.

Convergence results [Hanke, Neubauer, and Scherzer, '95]:

- The sequence of errors $\left\{ \left\| x_* - x_k^\delta \right\| \right\}$ is monotonically decreasing for $0 \le k < k_*$;
- Via discrepancy principle,

$$x_{k_*}^\delta \to x_*, \quad \delta \to 0.$$

MODIFIED LANDWEBER ITERATION

The Fréchet derivative of F in the Landweber iteration (3.2) can be replaced by another linear operator $G_{x_k^\delta}$ that is sufficiently close to $F'(x_k^\delta)$, leading to the so-called **modified Landweber method** [Scherzer, '95]:

$$x_{k+1}^\delta = x_k^\delta - w_k G_{x_k^\delta}^*(F(x_k^\delta) - y^\delta), \quad k \geq 0$$

for the starting point $x_0^\delta = x_0$ and step sizes $w_k > 0$.
Convergence via discrepancy principle (4.1) also holds.

What actual choices of $G_{x_k^\delta}$ are available?

An application in plasma physics

Consider an open bounded subset $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2\}$, with a Lipschitz boundary $\partial\Omega$, and consider the nonsmooth semilinear equation

$$\begin{cases} -\Delta y + y^+ = u & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1}$$

with $u \in L^2(\Omega)$ and $y^+(x) := \max(y(x), 0)$ for almost every $x \in \Omega$.

Direct problem: for each $u \in L^2(\Omega)$, a unique solution y_u in $H^1_0(\Omega) \cap C(\Omega)$ exists.

Exploring the forward operator

Define the forward operator $F : L^2(\Omega) \to H_0^1(\Omega) \cap C(\Omega)$ where y_u solves (5.1) for $u \in L^2(\Omega)$.

Proposition 5.1 (Christof et.al.,'18)

F is globally Lipschitz continuous as a function from $L^2(\Omega)$ to $H_0^1(\Omega) \cap C(\Omega)$.

Lemma 5.1 (Clason and Nhu, '19)

$F : L^2(\Omega) \to L^2(\Omega)$ is completely continuous, i.e. $u_n \rightharpoonup u$ implies $F(u_n) \to F(u)$.

Notation: $\{v = 0\} : \{x \in \Omega \; : \; v(x) = 0\}$.

Proposition 5.2 (Christof et.al.,'18)

For any $u \in L^2(\Omega)$ and $h \in L^2(\Omega)$, $F : L^2(\Omega) \to H_0^1(\Omega)$ is directionally differentiable, has a directional derivative $F'(u; h)$ in the direction $h \in L^2(\Omega)$ given by $\eta \in H_0^1(\Omega)$ which solves

$$\begin{cases} -\Delta\eta + \mathbb{1}_{\{y_u=0\}}\eta^+ + \mathbb{1}_{\{y_u>0\}}\eta = h & \text{in } \Omega, \\ \eta = 0 & \text{on } \partial\Omega. \end{cases}$$

Moreover, computing directional derivative of F could be difficult.

A non-Gâteaux differentiability result

Notation: λ^d: d-dimensional Lebesgue measure.

However, F in general is not Gâteaux differentiable.

Proposition 5.3 (Clason and Nhu, '19)

Let $u \in L^2(\Omega)$; $F : L^2(\Omega) \to L^2(\Omega)$ is Gâteaux differentiable in u if and only if $\lambda^d(\{y_u = 0\}) = 0$.

The Bouligand subdifferential

We introduce the set of Gâteaux points of F as

$$D := \left\{ v \in L^2(\Omega) \; : \; F : L^2(\Omega) \to H_0^1(\Omega) \text{ is Gâteaux differentiable in } v \right\}$$

Define the (strong-strong) **Bouligand subdifferential** at $u \in L^2(\Omega)$:

$$\begin{aligned} \partial_B F(u) := \{ G_u \in \mathbb{L}(L^2(\Omega), H_0^1(\Omega)) \; : \; \exists \{u_n\}_{n\mathbb{N}} \subset D \text{ such that } \\ u_n \to u \text{ in } L^2(\Omega) \\ \text{and } F'(u_n; h) \to G_u h \text{ in } H_0^1(\Omega) \text{ for all } h \in L^2(\Omega) \}, \end{aligned}$$

For all $u \in L^2(\Omega)$, $F'(u) \in \partial_B F(u)$ when F is Gateaux differentiable in u [Christof et.al.,'18].

Uniform boundedness: there exists constants L and $\hat{L}$ satisfying

$$\|G_u\|_{\mathbb{L}(L^2(\Omega),L^2(\Omega))} \leq L, \qquad \|G_u\|_{\mathbb{L}(L^2(\Omega),H_0^1(\Omega))} \leq \hat{L}$$

PROPOSITION 5.4 (CHRISTOF ET.AL.,'18)

Given $u \in L^2(\Omega)$, let $G_u : L^2(\Omega) \to H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ be the solution operator that maps $h \in L^2(\Omega)$ to the unique solution $\eta \in H_0^1(\Omega)$ of

$$\begin{cases} -\Delta\eta + \mathbb{1}_{\{y_u>0\}}\eta = h & \text{in } \Omega, \\ \eta = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.2}$$

where $y_u := F(u)$. Then $G_u \in \partial_B F(u)$.

Disretisation via standard finite element

[Clason and Nhu, '19] Set $\Omega \in \mathbb{R}^2$, $\mathcal{T}_h$ the triangulation of Ω, V_h the finite-dimensional subspace of $H_0^1(\Omega)$ consists of functions v_h s.t. $v_h|_T \in P_1(T)$, for all $T \in \mathcal{T}_h$.
$\{\phi_j\}_{j=1}^n$: basis of V_h corresponding to the set of nodes $\mathcal{N}_h := \{p_1, \dots, p_n\}$; $\phi_j(p_i) = \delta_{ji}$,the Kronecker delta.
Weak form of (5.1):

$$\int_\Omega \nabla y_h \cdot \nabla v_h dx + \frac{1}{3} \sum_{T \in \mathcal{T}_h} |T| \sum_{p_i \in \mathcal{N} \cap \bar{T}} \max(y_h(p_i, 0) v_h(p_i)$$

$$= \int_\Omega u_h v_h dx \qquad \forall v_h \in V_h,$$

$y_h, v_h \in V_h$ the FE approximation of y and u, resp.

Numerical examples

We consider $\Omega = (0,1)^2 \subset \mathbb{R}^2$, use a uniform triangulation with 128×128 vertices. Discretisation of the nonsmooth semilinear PDE and the Bouligand subdifferential were done using standard continuous piecewise linear finite elements.

Obtain noisy data with $\delta := \left\| y_h^\dagger - y_h^\delta \right\|_{L^2(\Omega)} = 1.04927 \times 10^{-4}$

We used a pre-defined constant w_k in the iterative method

$$u_{k+1} = u_k - w_k G_{u_k}^*(F(u_k) - y^\delta), \quad k \geq 0$$

under the discrepancy principle.

Numerical examples

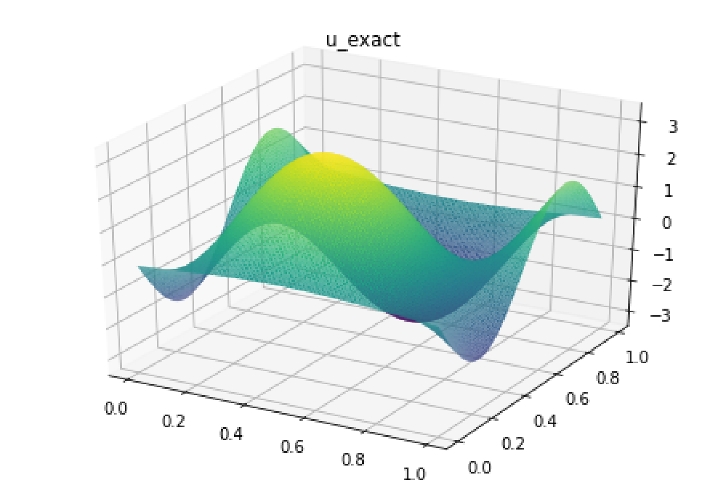

Figure: Exact solution $u^\dagger$

Numerical examples

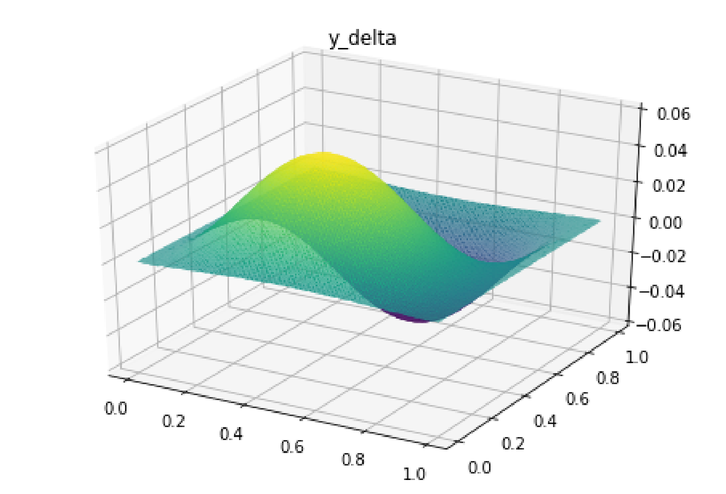

Figure: Noisy data y^δ

Numerical examples

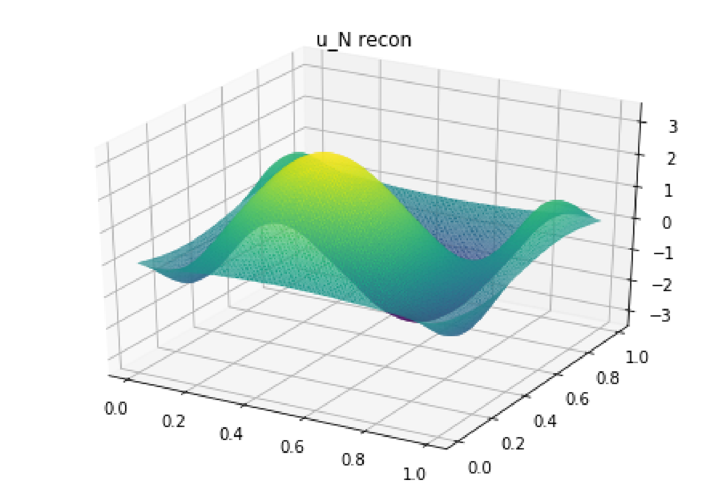

Figure: Reconstruction u_N via DP, $\tau = 1.4$, $N_\delta = 1103$.

What else to do?

- Considering discontinuous source terms (piecewise constant, sparse).
- Using convergent iterative methods with variable step sizes
- Explore more mathematically interesting inverse problems in plasma physics!

References

- Christof, C., Clason, C., Meyer, C., and S. Walter, *Optimal control of a non-smooth semilinear elliptic equation*, Math. Control Relat. Fields, 8(2018), No. 1, 247-276.
- C. Clason and V. H. Nhu, *Bouligand-Landweber iteration for a non-smooth ill-posed problem*, 2018 (arXiv:1803.02290)
- M. Hanke, A. Neubauer and O. Scherzer, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), 21–37.
- Kikuchi, F., Nakazato, K and Ushijima, T., *Finite element approximation of a nonlinear eigenvalue problem related to MHD equilibria*, Jpn. J. Appl. Math., 1(1984), 364–403.
- Rappaz, J., *Approximation of a nondifferentiable nonlinear problem related to MHD equilibria*, Numer. Math., 45 (1984), 117–133.
- Scherzer, O., *Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems*, J. Math. Anal. Appl., 194(1995), 911-933.
- Temam, R., *A non-linear eigenvalue problem: the shape at equilibrium of a confined plasma*, Arch. Ration. Mech. Anal., 60(1975), 51—73.