



Topics in Numerical Analysis II Computational Inverse Problems

Lecturer: Bangti Jin (btjin@math.cuhk.edu.hk)

Chinese University of Hong Kong

October 5, 2022



Outline

1 Tikhonov regularization - Nonlinear case



Nonlinear inverse problems

$$F(x) = y,$$

with $F : X \rightarrow Y$ being nonlinear operator (possibly compact) between Hilbert spaces X and Y

- (i) F is continuous
- (ii) F is **weakly sequentially closed**, i.e., $(x_n) \subset D(F)$, weak convergence of x_n to x^* and weak convergence of $F(x_n)$ to $F(x^*)$ in Y imply $x^* \in X$ and $F(x^*) \in Y$

- electrical impedance tomography
- diffuse optical tomography
- inverse scattering
- ...



Tikhonov regularization

$$J_{\alpha}(x) = \|F(x) - y\|^2 + \alpha\|x\|^2$$

and approximation x_{α}

$$x_{\alpha} \in \arg \min_{x \in D(F)} J_{\alpha}(x)$$

- existence of a global minimizer, not unique
- stability of the approximation (on subsequence)
- consistency of the approximation as $\delta \rightarrow 0$ (subseq.)
- convergence rate



convergence rate analysis Engl-Kunisch-Neubauer Inverse Problems 1989

Let $D(F)$ be **convex**, $y^\delta \in Y$ with $\|y^\delta - y^\dagger\| \leq \delta$, and $x^\dagger$ be an x_0 -minimum norm solution. Moreover, the following condition holds

- (i) F is **Frechet differentiable**.
- (ii) $\exists L > 0$ s.t. $\|F'(x^\dagger) - F'(z)\| \leq L\|x^\dagger - z\|$ for all $z \in D(F)$
- (iii) $\exists w \in Y$ s.t. $x^\dagger - x_0 = F'(x^\dagger)^* w$ and $L\|w\| < 1$.

Then with $\alpha \sim O(\delta)$, we have

$$\|x_\alpha^\delta - x^\dagger\| = O(\delta^{\frac{1}{2}}).$$

the same result holds for discrepancy principle



Model problem

Most inverse problems for PDEs are nonlinear in nature
model problem:

$$-\Delta u + qu = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega,$$

with given $q \in L^\infty(\Omega)$, $q \geq 0$, and $f \in L^2(\Omega)$

- The forward is well-posed: there exists a unique solution $u \in H_0^1(\Omega)$ (by Lax-Milgram theorem)
- Inverse problem: given $g \approx u^\delta$ in Ω , find q
- The operator $F(q) : q \mapsto u(q)$ is weakly continuous from $L^2(\Omega)$ to $L^2(\Omega)$.



Is the forward operator differentiable on the set $\mathcal{A}$?

- the set $\mathcal{A}$ does not have an interior point with respect to $L^2(\Omega)$!
- (Fréchet) differentiable at $q \in \mathcal{A}$ in the sense

$$\lim_{q+h \in \mathcal{A}: \|h\|_{L^2(\Omega)} \rightarrow 0} \frac{\|F(q+h) - F(q) - F'(q)h\|}{\|h\|_{L^2(\Omega)}} = 0$$

with $F'(q)$ being a bounded linear operator w.r.t. $L^2(\Omega)$

- Gâteaux differentiable at $q \in \mathcal{A}$ in the sense

$$\lim_{t \rightarrow 0, q+th \in \mathcal{A}} \frac{\|F(q+th) - F(q) - tF'(q)h\|}{t} = 0$$



candidate for the linearized operator

$$\int_{\Omega} \nabla u(q) \cdot \nabla v dx + \int_{\Omega} q u(q) v dx = \int_{\Omega} f v dx$$

$$\int_{\Omega} \nabla u(q+h) \cdot \nabla v dx + \int_{\Omega} (q+h) u(q+h) v dx = \int_{\Omega} f v dx$$

$\Rightarrow$ (with $\delta u = u(q+h) - u(q)$)

$$\int_{\Omega} \nabla \delta u \cdot \nabla v dx + \int_{\Omega} q \delta u \cdot \nabla v dx + \int_{\Omega} h u(q+h) v dx = 0$$

since $u(q+h) \approx u(q)$ for small $h \Rightarrow$ (with $\bar{u} = u'(q)h$)

$$\int_{\Omega} \nabla \bar{u} \cdot \nabla v dx + \int_{\Omega} q \bar{u} v dx = - \int_{\Omega} h u v dx \quad \forall v \in H_0^1(\Omega)$$

sensitivity problem: $\bar{u}$ is a linear operator in h



boundedness of the operator $\bar{u}$

$$\begin{aligned} \int_{\Omega} \nabla \bar{u} \cdot \nabla \bar{u} dx + \int_{\Omega} q \nabla \bar{u} \bar{u} dx &= - \int_{\Omega} h u \bar{u} \\ &\leq \|h\|_{L^2(\Omega)} \|u\|_{L^4(\Omega)} \|\bar{u}\|_{L^4(\Omega)} \quad (\text{generalized Hölder inequality}) \\ &\leq c \|h\|_{L^2(\Omega)} \|u\|_{H^1(\Omega)} \|\bar{u}\|_{H^1(\Omega)} \quad (\text{Sobolev embedding theorem}) \end{aligned}$$

$$+ \|u\|_{H^1(\Omega)} \leq c \Rightarrow$$

$$\|\bar{u}\|_{H^1(\Omega)} \leq c \|h\|_{L^2(\Omega)}$$



establish the approximation property

$$\int_{\Omega} \nabla u(q) \cdot \nabla v dx + \int_{\Omega} qu(q)v dx = \int_{\Omega} f v dx$$

$$\int_{\Omega} \nabla u(q+h) \cdot \nabla v dx + \int_{\Omega} (q+h)u(q+h)v dx = \int_{\Omega} f v dx$$

$$\int_{\Omega} \nabla u'(q)h \cdot \nabla v dx + \int_{\Omega} hu(q)hv dx = \int_{\Omega} f v dx$$

$$\delta u = u(q+h) - u(q) - u'(q)h \Rightarrow$$

$$\begin{aligned} 0 &= \int_{\Omega} \nabla \delta u \cdot \nabla v dx + \int_{\Omega} ((q+h)u(q+h) - qu(q) - hu(q))v dx \\ &= \int_{\Omega} \nabla \delta u \cdot \nabla v dx + \int_{\Omega} q\delta u v dx + \int_{\Omega} (hu'(q)h + h(u(q+h) - u(q)))v dx \end{aligned}$$



letting $v = \delta u$, we deduce

$$\begin{aligned}\|\delta u\|_{H^1(\Omega)}^2 &\leq c \left| \int_{\Omega} (hu'(q)h + h(u(q+h) - u(q))\delta u) dx \right| \\ &\leq c \|h\|_{L^2(\Omega)} (\|u'(q)h\|_{L^4(\Omega)} + \|u(q+h) - u(q)\|_{L^4(\Omega)}) \|\delta u\|_{L^4(\Omega)} \\ &\leq c \|h\|_{L^2(\Omega)} (\|u'(q)h\|_{H^1(\Omega)} + \|u(q+h) - u(q)\|_{H^1(\Omega)}) \|\delta u\|_{H^1(\Omega)} \\ &\leq c \|h\|_{L^2(\Omega)} (\|h\|_{L^2(\Omega)} + \|u(q+h) - u(q)\|_{H^1(\Omega)}) \|\delta u\|_{H^1(\Omega)}\end{aligned}$$

similar one can estimate

$$\|u(q+h) - u(q)\|_{H^1(\Omega)} \leq c \|h\|_{L^2(\Omega)} \quad \Rightarrow \quad \|\delta u\|_{H^1(\Omega)} \leq c \|h\|_{L^2(\Omega)}^2$$

$$\lim_{q+h \in \mathcal{A}: \|h\|_{L^2(\Omega)} \rightarrow 0} \frac{\|F(q+h) - F(q) - F'(q)h\|_{H^1(\Omega)}}{\|h\|_{L^2(\Omega)}} = 0$$

i.e., the forward operator is differentiable ...



convergence rates Engl-Kunish-Neubauer 1989

With $\alpha \sim \delta$, and $q^\dagger = F'(q^\dagger)^* w$ and $L\|w\| < 1$, there holds

$$\|q^\dagger - q_\alpha^\delta\| \leq c\delta^{1/2}$$



Let $A(q) : H^2(\Omega) \cap H_0^1(\Omega) \rightarrow L^2(\Omega)$ by

$$A(q)v = -v_{xx} + qv$$

Then $F(q) = A^{-1}(q)f$ is given by

$$F'(q)h = -A(q)^{-1}(hu)$$



What is the condition $F'(q^\dagger)^* w = q^\dagger - q_0$?

- $q^\dagger - q_0 = F'(q^\dagger)^* w = -u(q^\dagger)A(q^\dagger)^{-1}w$, with $w \in L^2(\Omega)$
- such an element exists if

$$\frac{q_0 - q^\dagger}{u(q^\dagger)} \in H^2(\Omega) \cap H_0^1(\Omega)$$

and it is given by

$$w = A(q^\dagger) \frac{q_0 - q^\dagger}{u(q^\dagger)}$$

$q_0 - q^\dagger$: should vanish near the boundary sufficiently quickly,
and should be smooth



in practice, one needs to discretize the problem ...

- finite difference methods
- **finite element methods**
- boundary element methods
- meshfree methods, particle methods
- ...

finite element methods

- versatile for general domains ...
- handle variable coefficient easily ...
- (relatively easy) theoretical justification ...



crash course on finite element methods

- $\mathcal{T}_h$: shape regular quasi-uniform triangulation of Ω , h is the mesh size, with $\Omega = \cup T$
- piecewise linear finite element space

$$V_h = \{v_h \in C_0(\bar{\Omega}) : v_h|_T \in P_1(T) \quad \forall T \in \mathcal{T}_h\}$$

- nodal basis $\{\varphi_i\}_{i=1}^N$ for V_h , i.e.,

$$\varphi_i(P_j) = \delta_{ij} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$



finite element problem: find $u_h \in V_h$ s.t.

$$\int_{\Omega} \nabla u_h \cdot \nabla v_h dx + \int_{\Omega} q_h u_h v_h dx = \int_{\Omega} f v_h dx \quad \forall v_h \in V_h$$

where $q_h \in \mathcal{A}_h = \mathcal{A} \cap V_h$

- the FEM problem has a unique solution (Lax-Milgram theorem)

With the nodal basis φ_j , the FEM is given by: for $j = 1, \dots, N$

$$\int_{\Omega} \nabla \sum_{i=1}^N u_i \varphi_i(x) \cdot \nabla \varphi_j(x) dx + \int_{\Omega} q_h(x) \sum_{i=1}^N u_i \varphi_i(x) \varphi_j(x) dx = \int_{\Omega} f \varphi_j(x) dx$$



resulting linear system

$$\mathbf{A}\mathbf{u} = \mathbf{f},$$

■ $\mathbf{u} = [u_1, \dots, u_N]^t \in \mathbb{R}^N$

■ $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ with

$$a_{ij} = \int_{\Omega} \nabla \varphi_i \cdot \nabla \varphi_j dx + \int_{\Omega} q_h \varphi_i \varphi_j dx$$

■ $\mathbf{f} = [f_1, \dots, f_N] \in \mathbb{R}^N$

All the integrals are only needed on the support, and can be evaluated analytically or using quadrature rules !!



discrete Tikhonov regularization

$$\min_{q_h \in \mathcal{A}_h} \{J_{\alpha, h}(q_h) = \frac{1}{2} \|F_h(q_h) - g\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|q_h\|_{L^2(\Omega)}^2\}$$

with $F_h(q_h)$ the FEM solution

- the discrete problem has a solution $q_h^* \in \mathcal{A}_h$
- what is the proper choice of h ? (hard!)
- Is the discrete solution convergent ? (relatively easy!)
- It is solvable on computers
- (the quadrature error, domain approximation error ignored)



- the finite element space V_h is dense in $H_0^1(\Omega)$ and $L^2(\Omega)$
- Ritz projection $R_h : H_0^1(\Omega) \rightarrow V_h$

$$(\nabla R_h v, \nabla v_h) = (\nabla v, \nabla v_h), \quad \forall v_h \in V_h$$

$$\lim_{h \rightarrow 0} \|v - R_h v\|_{H^1(\Omega)} = 0, \quad \forall v \in H_0^1(\Omega)$$

- Lagrange interpolation operator $I_h : C(\bar{\Omega}) \rightarrow V_h$

$$\lim_{h \rightarrow 0} \|v - I_h v\|_{C(\bar{\Omega})} = 0, \quad \forall v \in W^{1,p}(\Omega), p > d.$$



the discretization induces the error, does it affect the solution greatly ?

Does the discrete solution q_h^* approximate q^* ?

Answer: Yes, there exists a subsequence $(q_h^*)_{h>0}$ converges to a solution of the continuous problem!

argument:

- constant belongs to all $\mathcal{A}_h$: $(q_h^*)_{h>0}$ is uniformly bdd in $L^2(\Omega)$
- $\exists$ a subseq. $(q_h^* \in \mathcal{A}_h)_{h>0}$ converges weakly to q^* in $L^2(\Omega)$
- $u_h(q_h^*) \rightarrow u(q^*)$ weakly in $H^1(\Omega)$
- minimizing property of q_h^* (for smooth q)

$$J_{\alpha,h}(q_h^*) \leq J_{\alpha,h}(I_h q) \rightarrow J_{\alpha}(q) \text{ as } h \rightarrow 0$$



$$\int_{\Omega} \nabla u_h(q_h^*) \cdot \nabla v_h dx + \int_{\Omega} q_h^* u_h(q_h^*) v_h dx = \int_{\Omega} f v_h dx \quad \forall v_h \in V_h$$

- with $v_h = u_h(q_h^*) \Rightarrow \|u_h(q_h^*)\|_{H^1(\Omega)} \leq c$
- convergent sub-subseq. $u_h(q_h^*)$ to u^* weakly in $H^1(\Omega)$
- for any $v \in H_0^1(\Omega)$, take $v_h = R_h v \in V_h$ + weak conv. $\Rightarrow$

$$\int_{\Omega} \nabla u_h(q_h^*) \cdot \nabla v_h dx \rightarrow \int_{\Omega} \nabla u^* \cdot \nabla v dx \quad \text{as } h \rightarrow 0$$

- splitting

$$\begin{aligned} \int_{\Omega} q_h^* u_h v_h dx &= \int_{\Omega} q_h^* u^* v dx + \int_{\Omega} q_h^* u_h (v_h - v) dx + \int_{\Omega} q_h^* (u_h - u^*) v dx \\ &\rightarrow \int_{\Omega} q^* u^* v dx + 0 + 0 \end{aligned}$$



by the approximation property of R_h : $v_h \rightarrow v$ in $H^1(\Omega)$

$$\begin{aligned} \left| \int_{\Omega} q_h^* u_h (v_h - v) dx \right| &\leq \|q_h^*\|_{L^2(\Omega)} \|u_h\|_{L^4(\Omega)} \|v_h - v\|_{L^4(\Omega)} \\ &\leq c \|q_h^*\|_{L^2(\Omega)} \|u_h\|_{H^1(\Omega)} \|v_h - v\|_{H^1(\Omega)} \rightarrow 0 \end{aligned}$$

by the weak convergence of $u_h \rightarrow u^*$ in $H^1(\Omega)$ and compact embedding

$$\begin{aligned} \left| \int_{\Omega} q_h (u_h - u^*) v dx \right| &\leq \|q_h^*\|_{L^\infty(\Omega)} \|u_h - u^*\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq c \|q_h^*\|_{L^\infty(\Omega)} \|u_h - u^*\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)} \rightarrow 0 \end{aligned}$$

+ weak lower semicontinuity of norms

$$J_\alpha(q^*) \leq \lim_{h \rightarrow 0} J_{\alpha,h}(I_h q) = J_\alpha(q) \quad \text{for any smooth } q \in \mathcal{A}$$

a density argument shows this is true for any $q \in \mathcal{A}$



claim: $\lim_{h \rightarrow 0} J_{\alpha, h}(I_h q) = J_{\alpha}(q)$ for any smooth $q \in \mathcal{A}$
since q is smooth + error estimate for I_h ,

$$\|q - I_h q\|_{L^\infty(\Omega)} = 0$$

meanwhile, $u_h(I_h q) \rightarrow u(q)$ in $H^1(\Omega)$

$$\begin{aligned} \int_{\Omega} \nabla u_h(I_h q) \cdot \nabla v_h dx + \int_{\Omega} I_h q u_h(I_h q) v_h dx &= \int_{\Omega} f v_h dx \quad \forall v_h \in V_h \\ \int_{\Omega} \nabla u(q) \cdot \nabla v dx + \int_{\Omega} q u(q) v dx &= \int_{\Omega} f v dx \quad \forall v \in V \end{aligned}$$

$w_h = u(q) - u_h(I_h q)$ satisfies

$$\int_{\Omega} \nabla w \cdot \nabla v_h dx + \int_{\Omega} q w v_h dx = \int_{\Omega} (I_h q - q) u_h(I_h q) v_h dx, \quad \forall v_h \in V_h$$



projected gradient method

Now we need to solve the optimization problem over $q \in \mathcal{A}$

$$J_\alpha(q) = \frac{1}{2} \|F(q) - g\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|q\|_{L^2(\Omega)}^2$$

gradient descent method: given q^0 , solve for $k = 1, 2, \dots$

$$q^{k+1} = P_{\mathcal{A}}(q^k - \tau J'(q^k))$$

- $\tau > 0$: a step size (Amijo's rule etc)
- $J'_\alpha(q^k) \in L^2(\Omega)$ the gradient at q^k
- $P_{\mathcal{A}}$: the projection into $\mathcal{A}$
- needs suitable stopping criterion



so one essential cost per iteration is to compute the gradient

$$J'_\alpha(q^k) = \underbrace{F'(q^k)^*}_{\text{matrix}} \underbrace{(F(q) - g)}_{\text{vector}} + \alpha q^k$$

1st approach: compute the matrix directly and multiply with vector
matrix can be computed by finite difference etc

$$[F'(q^k)]_i \approx \frac{F(q^k + he_i) - F(q^k)}{h} \quad h \text{ small}$$

where

- $[\cdot]_i$: the i th column
- e_i : the i th basis function
- h : small but not too small ...
- cost: $O(n)$ forward solves $\sim$ very expensive (if the number of basis functions is large)



2nd approach: adjoint technique definition of the adjoint operator

$$(F'(q^k)^*(F(q^k) - g), h)_{L^2(\Omega)} = (F(q^k) - g, F'(q^k)h)_{L^2(\Omega)} \quad \forall h \in L^2(\Omega)$$

using the definition of $F'(q^k)h$

$$\int_{\Omega} \nabla(F'(q^k)h) \cdot \nabla v dx + \int_{\Omega} q^k(F'(q^k)h)v dx = - \int_{\Omega} h u v dx \quad \forall v \in H_0^1(\Omega) \quad (*)$$

this does not help since it does not involve $F(q^k) - g$ directly $\Rightarrow$

$$\int_{\Omega} \nabla v \cdot \nabla \mathbf{p} + \int_{\Omega} q^k v \mathbf{p} dx = \int_{\Omega} v(F(q^k) - g) dx \quad \forall v \in H_0^1(\Omega) \quad (**)$$

Taking $v = p$ in $(*)$ and $v = F'(q^k)h$ in $(**)$ $\Rightarrow$

$$(F(q^k) - g, F'(q^k)h)_{L^2(\Omega)} = -(up, h)_{L^2(\Omega)}$$



$$(F'(q)^*(F(q) - g), h)_{L^2(\Omega)} = -(up, h)_{L^2(\Omega)} \quad \forall h \in L^2(\Omega)$$

with p solving

$$\int_{\Omega} \nabla v \cdot \nabla p + \int_{\Omega} q^k v p dx = \int_{\Omega} v (F(q^k) - g) dx \quad \forall v \in H_0^1(\Omega)$$

$$F'(q^k)^*(F(q) - g) = -up \Rightarrow J'_{\alpha}(q^k) = -u(q^k)p + \alpha q^k$$

distinct features

- Only one forward solve
- the stiffness matrix is identical with the forward problem (no extra assembling cost)



what is open for the inverse problem ?

- convergence rates of the discretization scheme ?
- convergence rates for other approximations, e.g., neural networks ?
- ...