Convergence of IRGNM type methods under a tangential cone condition in Banach space

Mario Luiz Previatti de Souza
Barbara Kaltenbacher

Alpen-Adria-Universität Klagenfurt

mario.previatti@aau.at
barbara.kaltenbacher@aau.at

Colóquio UFSC, 6 de Outubro, 2017

Der Wissenschaftsfonds.

Overview

Setting

Continuous version

Discretized version

Error estimate

Model Example

FWF Der Wissenschaftsfonds. D-A-CH project

Regularization and Discretization of Inverse Problems for PDEs in Banach Spaces

Iteratively Regularized Gauss Newton Method (IRGNM)

Nonlinear ill-posed operator equation

$$F(x) = y^\delta, \quad \|y - y^\delta\| \leq \delta$$

Iteratively Regularized Gauss Newton Method (IRGNM)

Nonlinear ill-posed operator equation

$$F(x) = y^\delta, \quad \|y - y^\delta\| \le \delta$$

IRGNM-Tikhonov

$$x_{k+1}^\delta \in \text{argmin}_{x\in\mathcal{D}(F)} \|F'(x_k^\delta)(x - x_k^\delta) - (y^\delta - F(x_k^\delta))\|^p + \alpha_k \mathcal{R}(x)$$

$k = 0, 1, \ldots \qquad p \in [1, \infty)$

IRGNM-Ivanov

$$x_{k+1}^\delta \in \text{argmin}_{x\in\mathcal{D}(F)} \|F'(x_k^\delta)(x - x_k^\delta) - (y^\delta - F(x_k^\delta))\| \text{ s.t. } \mathcal{R}(x) \le \rho_k$$

$k = 0, 1, \ldots$

Setting

- $F : \mathcal{D}(F) \subset X \longrightarrow Y$, X and Y real Banach spaces
- $x^\dagger \in \mathcal{D}(F)$ exists, i.e., $F(x^\dagger) = y$
- $\mathcal{R}(x)$ is proper, convex and l.s.c. with $\mathcal{R}(x^\dagger) < \infty$
- discrepancy principle

$$k_* = \min\{k \in \mathbb{N}_0 : \|F(x_k^\delta) - y^\delta\| \leq \tau\delta\}, \; \tau > 1 \text{ fixed}$$

Setting

- $F : \mathcal{D}(F) \subset X \longrightarrow Y$, X and Y real Banach spaces
- $x^\dagger \in \mathcal{D}(F)$ exists, i.e., $F(x^\dagger) = y$
- $\mathcal{R}(x)$ is proper, convex and l.s.c. with $\mathcal{R}(x^\dagger) < \infty$
- discrepancy principle

$$k_* = \min\{k \in \mathbb{N}_0 : \|F(x_k^\delta) - y^\delta\| \leq \tau\delta\}, \ \tau > 1 \text{ fixed}$$

- tangential cone condition $c_{tc} < 1/3$

$$\|F(\tilde{x}) - F(x) - F'(x)(\tilde{x} - x)\| \leq c_{tc}\|F(\tilde{x}) - F(x)\|, \ \forall x \in \mathcal{B}_R \subset \mathcal{D}(F)$$

- α_k, ρ_k a priori

$$\alpha_k = \alpha_0\theta^k \text{ for some } \theta \in \left(\left(2\frac{c_{tc}}{1 - c_{tc}}\right)^p, 1\right) \text{ and } \rho_k \equiv \rho \geq \mathcal{R}(x^\dagger)$$

Convergence result (continuous version)

Theorem

Let for all $r \geq \mathcal{R}(x^{\dagger})$, the sublevel set $\mathcal{B}_r = \{x \in \mathcal{D}(F) : \mathcal{R}(x) \leq r\}$ be compact with respect to some topology τ on X. Let for all $x \in \mathcal{B}_R$, $F^{'}(x)$ and F be τ-to-norm continuous, for some chosen $R > \mathcal{R}(x^{\dagger})$.

Then, for fixed δ and y^{δ}, the iterations are well defined and remain in $\mathcal{B}_R$ and the stopping index k_* is finite.
Moreover, we have τ-convergence as $\delta \to 0$.
If the solution $x^{\dagger}$ of $F(x) = y$ is unique in $\mathcal{B}_R$, then $x^{\delta}_{k_*(\delta,y^{\delta})} \to x^{\dagger}$ as $\delta \to 0$.
Additionally, k_* satisfies the asymptotics $k_* = \mathcal{O}(|log 1/\delta|)$.

Proof idea

Proof idea

1. Existence of minimizer: direct method of calculus of variations

Proof idea

1. Existence of minimizer: direct method of calculus of variations
2. Find a recursive estimate formula for $\|F(x_k^\delta) - y^\delta\|^p$

Proof idea

1. Existence of minimizer: direct method of calculus of variations
2. Find a recursive estimate formula for $\|F(x_k^\delta) - y^\delta\|^p$
3. Find upper bound for $k_*(\delta, y^\delta) \leq \bar{k}(\delta) = \mathcal{O}(\log 1/\delta)$

$$k_*(\delta, y^\delta) \leq \bar{k}(\delta) = \frac{\left(p\log(1/\delta) + \log\left(d_0 + \frac{\alpha_0}{\theta - q}\mathcal{R}(x^\dagger)\right) - \log\left(\tilde{\tau} - \frac{C}{1-q}\right)\right)}{\log 1/\theta} - 1$$

Proof idea

1. Existence of minimizer: direct method of calculus of variations
2. Find a recursive estimate formula for $\|F(x_k^\delta) - y^\delta\|^p$
3. Find upper bound for $k_*(\delta, y^\delta) \leq \bar{k}(\delta) = \mathcal{O}(\log 1/\delta)$

$$k_*(\delta, y^\delta) \leq \bar{k}(\delta) = \frac{\left(p\log(1/\delta) + \log\left(d_0 + \frac{\alpha_0}{\theta - q}\mathcal{R}(x^\dagger)\right) - \log\left(\tilde{\tau} - \frac{C}{1-q}\right)\right)}{\log 1/\theta} - 1$$

4. Find upper bound for $\mathcal{R}(x_k^\delta) \leq R$ for $k \in \{1, \ldots, k_*(\delta, y^\delta)\}$

$$R := \theta\left(\frac{d_0}{\alpha_0} + \frac{\mathcal{R}(x^\dagger)}{\theta - q}\right)\left(1 + \frac{C}{1-q}\left(\tilde{\tau} - \frac{C}{1-q}\right)^{-1}\right)$$

Proof idea

1. Existence of minimizer: direct method of calculus of variations
2. Find a recursive estimate formula for $\|F(x_k^\delta) - y^\delta\|^p$
3. Find upper bound for $k_*(\delta, y^\delta) \leq \bar{k}(\delta) = \mathcal{O}(\log 1/\delta)$

$$k_*(\delta, y^\delta) \leq \bar{k}(\delta) = \frac{\left(p\log(1/\delta) + \log\left(d_0 + \frac{\alpha_0}{\theta - q}\mathcal{R}(x^\dagger)\right) - \log\left(\tilde{\tau} - \frac{C}{1-q}\right)\right)}{\log 1/\theta} - 1$$

4. Find upper bound for $\mathcal{R}(x_k^\delta) \leq R$ for $k \in \{1, \dots, k_*(\delta, y^\delta)\}$

$$R := \theta\left(\frac{d_0}{\alpha_0} + \frac{\mathcal{R}(x^\dagger)}{\theta - q}\right)\left(1 + \frac{C}{1-q}\left(\tilde{\tau} - \frac{C}{1-q}\right)^{-1}\right)$$

5. Repeat for IRGNM-Ivanov

Discretized version

IRGNM-Tikhonov

$$x^{\delta}_{k+1,h} \in \text{argmin}_{x \in \mathcal{D}(F) \cap X^k_h} \| F^{k'}_h(x^{\delta}_{k,h})(x - x^{\delta}_{k,h}) - (y^{\delta} - F^k_h(x^{\delta}_{k,h})) \|^p + \alpha_k \mathcal{R}(x)$$

$k = 0, 1, \ldots \qquad p \in [1, \infty)$

IRGNM-Ivanov

$$x^{\delta}_{k+1,h} \in \text{argmin}_{x \in \mathcal{D}(F) \cap X^k_h} \| F^{k'}_h(x^{\delta}_{k,h})(x - x^{\delta}_{k,h}) - (y^{\delta} - F^k_h(x^{\delta}_{k,h})) \| \text{ s.t. } \mathcal{R}(x) \leq \rho_k$$

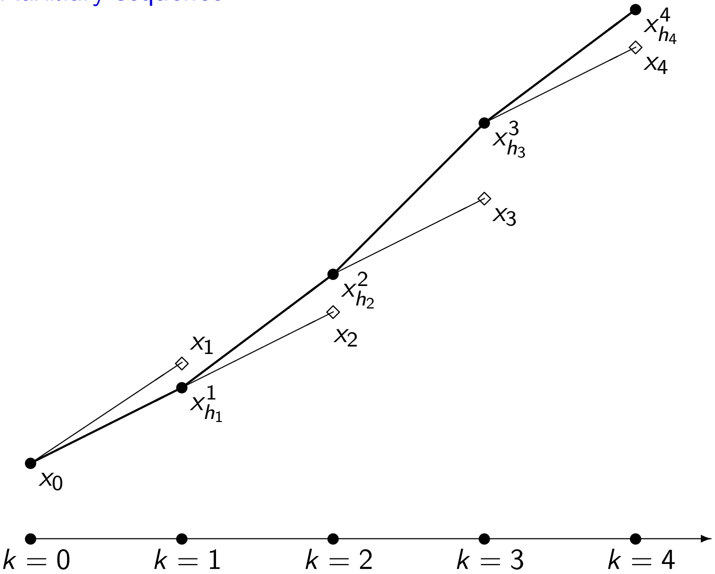
$k = 0, 1, \ldots$

Approach

Auxiliary sequence

Approach

Auxiliary sequence



Convergence result (discretized version)

Corollary

Assume the error estimates

$$\|F(x_{k,h}^\delta) - y^\delta\| - \|F(x_k^\delta) - y^\delta\| \le \eta_k$$
$$\|F_h^k(x_{k,h}^\delta) - y^\delta\| - \|F_h^k(x_k^\delta) - y^\delta\| \le \xi_k$$
$$\mathcal{R}(x_{k,h}^\delta) - \mathcal{R}(x_k^\delta) \le \zeta_k$$

hold with

$$\eta_k \le \bar{\tau}\delta, \quad \xi_k \le \hat{\tau}\delta, \quad \zeta_k \le \bar{\zeta} \qquad \text{(tolerance)}$$

for all $k \le k_*(\delta, y^\delta)$.

Then, the Theorem remains valid for $x^\delta_{k_*(\delta,y^\delta),h}$ in place of $x^\delta_{k_*(\delta,y^\delta)}$.

Goal oriented error estimators

$$F(x) = y \Leftrightarrow \begin{cases} A(x,u) = 0 \text{ model equation} \\ C(u) = y \text{ observation operator} \end{cases}$$

for $F = C \circ S$ such that $A(x, S(x)) = 0$, S parameter-to-state map

Goal oriented error estimators

$$F(x) = y \Leftrightarrow \begin{cases} A(x,u) = 0 \text{ model equation} \\ C(u) = y \text{ observation operator} \end{cases}$$

for $F = C \circ S$ such that $A(x, S(x)) = 0$, S parameter-to-state map

IRGNM-Tikhonov in a decoupled way

$(x^\delta_{k+1,h}, v^\delta_{k,h}, u^\delta_{k+1,h}, u^\delta_{k,h})$ solves

$$\min_{(x,v,u,\tilde{u})} \|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\|^p + \alpha_k \mathcal{R}(x)$$

$$\text{s.t.} \quad A'_x(x^\delta_{k,h}, \tilde{u})(x - x^\delta_{k,h}) + A'_u(x^\delta_{k,h}, \tilde{u})v = 0,$$

$$A(x^\delta_{k,h}, \tilde{u}) = 0, \quad A(x,u) = 0$$

Goal oriented error estimators

$$F(x) = y \Leftrightarrow \begin{cases} A(x,u) = 0 \text{ model equation} \\ C(u) = y \text{ observation operator} \end{cases}$$

for $F = C \circ S$ such that $A(x, S(x)) = 0$, S parameter-to-state map

IRGNM-Tikhonov in a decoupled way

$(x^\delta_{k+1,h}, v^\delta_{k,h}, u^\delta_{k+1,h}, u^\delta_{k,h})$ solves

$$\min_{(x,v,u,\tilde{u})} \|C^{'}(\tilde{u})v + C(\tilde{u}) - y^\delta\|^p + \alpha_k \mathcal{R}(x)$$

$$\text{s.t.} \quad A^{'}_x(x^\delta_{k,h}, \tilde{u})(x - x^\delta_{k,h}) + A^{'}_u(x^\delta_{k,h}, \tilde{u})v = 0,$$

$$A(x^\delta_{k,h}, \tilde{u}) = 0, \quad A(x,u) = 0$$

Quantities of interest

$I_1(x, v, u, \tilde{u}) = \|C(\tilde{u}) - y^\delta\|$
$I_2(x, v, u, \tilde{u}) = \|C(u) - y^\delta\| \rightsquigarrow$ just in theory
$I_3(x, v, u, \tilde{u}) = \mathcal{R}(x)$

Computing η_k, ξ_k, ζ_k for IRGNM-Tikhonov

- Lagrange functional $L(z)$, z its stationary point, $L^{'}(z)(dz) = 0, \forall dz$
- Auxiliary functional $M_i(z, \bar{z}) = I_i(z) + L^{'}(z)(\bar{z})$, $i = 1, 2, 3$, $\tilde{z} = (z, \bar{z})$ its stationary point

z_h, $\tilde{z}_h$ are the discrete stationary points of L and M

Computing η_k, ξ_k, ζ_k for IRGNM-Tikhonov

- Lagrange functional $L(z)$, z its stationary point, $L^{'}(z)(dz) = 0, \forall dz$
- Auxiliary functional $M_i(z, \bar{z}) = I_i(z) + L^{'}(z)(\bar{z})$, $i = 1, 2, 3$, $\tilde{z} = (z, \bar{z})$ its stationary point

z_h, $\tilde{z}_h$ are the discrete stationary points of L and M

[Becker,Vexler 2005] $\Rightarrow$

$$I_i^k(x, v, u, \tilde{u}) - I_i^k(x_h, v_h, u_h, \tilde{u}_h) = \underbrace{\frac{1}{2} M_i^{'}(\tilde{z}_h)(\tilde{z} - \hat{z}_h)}_{\epsilon_i^k} + \mathcal{O}(\|\tilde{z} - \tilde{z}_h\|^3), \quad \forall \hat{z}_h$$

$$\eta_{k+1} = \epsilon_1^{k+1} + \epsilon_2^k \text{ (theory)}, \quad \xi_k = \epsilon_1^k, \quad \zeta_k = \epsilon_3^k$$

Only one more Newton step for the stationary point of M

Optimality conditions

The minimizer $(x, v, u, \tilde{u}, \lambda, \tilde{\mu}, \mu)$ satisfies the optimality system
$p = 2$
...skipping h in notation

$$
\begin{aligned}
&-\Big(A_x'(x,u)^*\mu + A_x'(x_k^\delta,\tilde{u})^*\lambda\Big) \in \alpha_k \partial\mathcal{R}(x), && \forall dx \in X\\
&2\langle C'(\tilde{u})(dv), C'(\tilde{u})v + C(\tilde{u}) - y^\delta\rangle + \langle A_u'(x_k^\delta,\tilde{u})(dv),\lambda\rangle = 0, && \forall dv \in V\\
&\langle A_u'(x,u)(du),\mu\rangle = 0, && \forall du \in V\\
&\langle A_{xu}''(x_k^\delta,\tilde{u})(x - x_k^\delta, d\tilde{u}) + A_{uu}''(x_k^\delta,\tilde{u})(v, d\tilde{u}),\lambda\rangle + \langle A_u'(x_k^\delta,\tilde{u})(d\tilde{u}),\tilde{\mu}\rangle &&\\
&\quad + 2\langle C''(\tilde{u})(d\tilde{u}, v) + C'(\tilde{u})d\tilde{u}, C'(\tilde{u})v + C(\tilde{u}) - y^\delta\rangle = 0, && \forall d\tilde{u} \in V\\
&\langle A_x'(x_k^\delta,\tilde{u})(x - x_k^\delta) + A_u'(x_k^\delta,\tilde{u})v, d\lambda\rangle = 0, && \forall d\lambda \in W\\
&\langle A(x_k^\delta,\tilde{u}), d\tilde{\mu}\rangle = 0, && \forall d\tilde{\mu} \in W\\
&\langle A(x,u), d\mu\rangle = 0, && \forall d\mu \in W
\end{aligned}
$$

Optimality conditions...in practice

The minimizer $(x, v, \tilde{u}, \lambda, \tilde{\mu})$ satisfies the optimality system
$p = 2$
$\ldots$ skipping h in notation

$$
\begin{aligned}
&-A_x^{'}(x_k^\delta, \tilde{u})^*\lambda \in \alpha_k \partial \mathcal{R}(x), && \forall dx \in X \\
&2\langle C^{'}(\tilde{u})(dv), C^{'}(\tilde{u})v + C(\tilde{u}) - y^\delta \rangle + \langle A_u^{'}(x_k^\delta, \tilde{u})(dv), \lambda\rangle = 0, && \forall dv \in V \\
&\langle A_{xu}^{''}(x_k^\delta, \tilde{u})(x - x_k^\delta, d\tilde{u}) + A_{uu}^{''}(x_k^\delta, \tilde{u})(v, d\tilde{u}), \lambda\rangle + \langle A_u^{'}(x_k^\delta, \tilde{u})(d\tilde{u}), \tilde{\mu}\rangle \\
&\quad + 2\langle C^{''}(\tilde{u})(d\tilde{u}, v) + C^{'}(\tilde{u})d\tilde{u}, C^{'}(\tilde{u})v + C(\tilde{u}) - y^\delta\rangle = 0, && \forall d\tilde{u} \in V \\
&\langle A_x^{'}(x_k^\delta, \tilde{u})(x - x_k^\delta) + A_u^{'}(x_k^\delta, \tilde{u})v, d\lambda\rangle = 0, && \forall d\lambda \in W \\
&\langle A(x_k^\delta, \tilde{u}), d\tilde{\mu}\rangle = 0, && \forall d\tilde{\mu} \in W
\end{aligned}
$$

Computing stationary point

1st- $\langle A(x_k^\delta, \tilde{u}), d\tilde{\mu}\rangle = 0, \forall d\tilde{\mu} \in W$

2nd-

$$\begin{aligned} (x_{k+1,h}^\delta, v_{k,h}^\delta) \quad & \in argmin_{(x,v)\in\mathcal{D}(F)\times V} \| C^{'}(\tilde{u})v + C(\tilde{u}) - y^\delta\|^2 + \alpha_k \mathcal{R}(x) \\ \text{s.t. } \forall w \in W: \quad & \langle A_x^{'}(x_k^\delta, \tilde{u})(x - x_k^\delta) + A_u^{'}(x_k^\delta, \tilde{u})v, w\rangle_{W^*,W} = 0 \end{aligned}$$

3rd-

$$\begin{aligned} &\langle A_{xu}^{''}(x_k^\delta, \tilde{u})(x - x_k^\delta, d\tilde{u}) + A_{uu}^{''}(x_k^\delta, \tilde{u})(v, d\tilde{u}), \lambda\rangle + \langle A_u^{'}(x_k^\delta, \tilde{u})(d\tilde{u}), \tilde{\mu}\rangle \\ &\quad + 2\langle C^{''}(\tilde{u})(d\tilde{u}, v) + C^{'}(\tilde{u})d\tilde{u}, C^{'}(\tilde{u})v + C(\tilde{u}) - y^\delta\rangle = 0, \qquad \forall d\tilde{u} \in V \end{aligned}$$

Example

Model problem

Model equation $\begin{cases} -\Delta u + \kappa u^3 = x \text{ in } \Omega \subset \mathbb{R}^d \\ u = 0 \text{ on } \partial\Omega \end{cases}$

Observation operator $C(u) = y^\delta$, $C : W_0^{1,q'}(\Omega) \to L^2(\Omega)$, $q > d$

Example

Model problem

Model equation $\begin{cases} -\Delta u + \kappa u^3 = x \text{ in } \Omega \subset \mathbb{R}^d \\ u = 0 \text{ on } \partial\Omega \end{cases}$

Observation operator $C(u) = y^\delta$, $C : W_0^{1,q'}(\Omega) \to L^2(\Omega)$, $q > d$

IRGNM-Tikhonov

$$\min_{(x,v,u,\tilde{u})} \|C(v+\tilde{u}) - y^\delta\|^2_{L^2(\Omega)} + \alpha_k \|x\|_{\mathcal{M}(\Omega)}$$

s.t. $\forall w \in W_0^{1,q}(\Omega)$:

$$\int_\Omega (\nabla v \nabla w + 3\kappa \tilde{u}^2 v w) d\Omega = \int_\Omega w d(x - x_k),$$

$$\int_\Omega (\nabla \tilde{u} \nabla w + \kappa \tilde{u}^3 w) d\Omega = \int_\Omega w dx_k,$$

$$\int_\Omega (\nabla u \nabla w + \kappa u^3 w) d\Omega = \int_\Omega w dx.$$

Computing stationary point

(skipping h,δ)... in strong formulation

1st- solve the nonlinear equation

$$-\Delta \tilde{u} + \kappa \tilde{u}^3 = x^\delta_{k,h}$$

2nd- solve the linear case

$$\begin{aligned} (x^\delta_{k+1,h}, v^\delta_{k,h}) \quad & \in argmin_{(x,v)\in\mathcal{D}(F)\times V} \|v + \tilde{u} - y^\delta\|^2 + \alpha_k \|x\|_{\mathcal{M}(\Omega)} \\ \text{s.t.} \quad & -\Delta v + 3\kappa\tilde{u}^2 v = x - x^\delta_{k,h} \end{aligned}$$

3rd- solve the linear equation, it matters just for error computation

$$-\Delta\tilde{\mu} + 3\kappa\tilde{u}^2\tilde{\mu} = -6\kappa\tilde{u}v\lambda - 2(v + \tilde{u} - y^\delta)$$