Convergence guarantees for Newton type methods in tomographic problems via range invariance

Barbara Kaltenbacher

Alpen-Adria-Universität Klagenfurt

RICAM, Nov 28, 2022

FWF Der Wissenschaftsfonds.

Modeling - Analysis - Optimization
DOC.FUNDS DOCTORAL SCHOOL

UNIVERSITÄT KLAGENFURT

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs
  - electrical impedance tomography EIT:
    Identify conductivity $\sigma = \sigma(x)$ in
    $$-\nabla(\sigma \nabla \phi) = 0 \text{ in } \Omega$$
    from boundary observations of voltage/current $(\phi, \partial_\nu \phi)$ pairs.
  - full waveform inversion FWI / ultrasound tomography:
    Identify wave speed $c = c(x)$ in an ibvp for
    $$p_{tt} - c^2 \triangle p = 0 \text{ in } \Omega$$
    from additional boundary observations of the pressure p.

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs
  - electrical impedance tomography EIT:
    Identify conductivity $\sigma = \sigma(x)$ in
    $$-\nabla(\sigma\nabla\phi) = 0 \text{ in } \Omega$$
    from boundary observations of voltage/current $(\phi, \partial_\nu\phi)$ pairs.
  - full waveform inversion FWI / ultrasound tomography:
    Identify wave speed $c = c(x)$ in an ibvp for
    $$p_{tt} - c^2\triangle p = 0 \text{ in } \Omega$$
    from additional boundary observations of the pressure p.

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs from boundary observations.
- $\rightsquigarrow$ nonlinear ill-posed operator equation

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs from boundary observations.
- $\rightsquigarrow$ nonlinear ill-posed operator equation

all-at-once formulation

$$\left.\begin{array}{ll} A(q,u)=0 & \text{(model equation)} \\ Cu=y & \text{(observation equation)} \end{array}\right\} \quad \Leftrightarrow: \; \mathbb{F}(q,u)=(0,y)^T$$

reduced formulation, with parameter-to-state operator $S:\mathcal{D}(\mathbf{F})\to V$ defined by the first equation in

$$A(q,S(q))=0 \text{ and } C(S(q))=y \quad \Leftrightarrow: \; \mathbf{F}(q)=y.$$

$A:Q\times V\to W^*\ldots$ model operator, $\quad C\ \ldots$ observation op.
$Q\ldots$ parameter space, $V\ \ldots$ state space, $Y\ldots$ data space.

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs from boundary observations.
- $\rightsquigarrow$ nonlinear ill-posed operator equation

  all-at-once formulation

  $$\left.\begin{array}{ll} A(q,u)=0 & \text{(model equation)} \\ Cu=y & \text{(observation equation)} \end{array}\right\} \quad \Leftrightarrow: \ \mathbb{F}(q,u)=(0,y)^T$$

  reduced formulation, with parameter-to-state operator $S:\mathcal{D}(\mathbf{F})\to V$ defined by the first equation in

  $$A(q,S(q))=0 \text{ and } C(S(q))=y \quad \Leftrightarrow: \ \mathbf{F}(q)=y.$$

  $A:Q\times V\to W^*\ldots$ model operator, $\quad C\ \ldots$ observation op.
  $Q\ldots$ parameter space, $V\ \ldots$ state space, $Y\ldots$ data space.
- More generally, consider

  $$F(x)=y$$

  with $x=(q,u),\ F(x)=\mathbb{F}(q,u)$ or $x=q,\ F(x)=\mathbf{F}(q)$

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs from boundary observations.
- $\rightsquigarrow$ nonlinear ill-posed operator equation

$$F(x) = y$$

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs from boundary observations.
- $\rightsquigarrow$ nonlinear ill-posed operator equation

$$F(x) = y$$

- regularize!

Motivation: PDE based tomographic imaging

- Some tomographic imaging techniques lead to coefficient identification in PDEs from boundary observations.
- $\rightsquigarrow$ nonlinear ill-posed operator equation

$$F(x) = y$$

- regularize!
- convergence of iterative regularization:
  $\rightsquigarrow$ restrictions on nonlinearity of F
- convergence of variational regularization: need local convexity of Tikhonov functional $\rightsquigarrow$ restrictions on nonlinearity of F

The tangential cone condition and relatives

- tangential cone condition [Scherzer, 1995]

  $$\forall x, \tilde{x} \in U \,:\, \|F(x)-F(\tilde{x})-F'(x)(x-\tilde{x})\|_Y \leq c_{tc}\|F(x)-F(\tilde{x})\|_Y$$

  (in a neighborhood U of the exact solution) for the convergence of Landweber iteration [Hanke & Neubauer & Scherzer, 1995] and Newton type methods [Hanke, 1997], [BK & Previatti, 2018]
- *weak nonlinearity* condition [Chavent & Kunisch, 1996] for local convexity of the Tikhonov functional
- Newton-Mysovskii condition

  $$\forall x, \tilde{x} \in U \,:\, \|(F'(x) - F'(\tilde{x}))F'(x)^\dagger\|_X \leq C_{NM}\|x-\tilde{x}\|_X,$$

  [Deuflhardt & Engl & Scherzer, 1998] $^\dagger \ldots$ generalized inverse
- some relaxed versions of convexity, see e.g., [Kindermann, 2017]

The tangential cone condition and relatives

- tangential cone condition [Scherzer, 1995]

$$\forall x, \tilde{x} \in U \, : \, \|F(x) - F(\tilde{x}) - F'(x)(x - \tilde{x})\|_Y \le c_{\mathrm{tc}} \|F(x) - F(\tilde{x})\|_Y$$

- sufficient for this (and often used for its verification) is

$$\forall x, \tilde{x} \in U \, : \quad \mathrm{rng}(F'(x)^\star) = \mathrm{rng}(F'(\tilde{x})^\star).$$

where $^\star \ldots$ adjoint

The tangential cone condition and relatives

- tangential cone condition [Scherzer, 1995]

  $$\forall x, \tilde{x} \in U \ : \ \|F(x) - F(\tilde{x}) - F'(x)(x - \tilde{x})\|_Y \leq c_{\mathrm{tc}} \|F(x) - F(\tilde{x})\|_Y$$

- sufficient for this (and often used for its verification) is

  $$\forall x, \tilde{x} \in U \ : \quad \mathrm{rng}(F'(x)^\star) = \mathrm{rng}(F'(\tilde{x})^\star).$$

  where $^\star \ldots$ adjoint

- Let's skip the $\star$:

  $$\forall x, \tilde{x} \in U \ : \quad \mathrm{rng}(F'(x)) = \mathrm{rng}(F'(\tilde{x})).$$

  that is, range invariance of the linearized forward operator

The range invariance condition and relatives

- range invariance of the linearized forward operator

$$\forall x, \tilde{x} \in U : \quad \mathrm{rng}(F'(x)) = \mathrm{rng}(F'(\tilde{x})).$$

The range invariance condition and relatives

- range invariance of the linearized forward operator

$$\forall x, \tilde{x} \in U \ : \quad \mathrm{rng}(F'(x)) = \mathrm{rng}(F'(\tilde{x})).$$

- *affine covariant Lipschitz condition*

$$\forall x, \tilde{x} \in U \ : \ \|F'(x)^\dagger (F'(x) - F'(\tilde{x}))\|_X \leq C_{\mathrm{acL}} \|x - \tilde{x}\|_X ,$$

(compare to *Newton-Mysovskii*) allows to prove convergence of Newton and quasi Newton methods, cf., e.g., [Burger & BK 2006, Deuflhardt & Engl & Scherzer 1998, BK 1997, 1998].

The range invariance condition and relatives

- range invariance of the linearized forward operator

$$\forall x, \tilde{x} \in U \; : \quad \mathrm{rng}(F'(x)) = \mathrm{rng}(F'(\tilde{x})).$$

- *affine covariant Lipschitz condition*

$$\forall x, \tilde{x} \in U \; : \; \|F'(x)^{\dagger}(F'(x) - F'(\tilde{x}))\|_X \leq C_{\mathrm{acL}} \|x - \tilde{x}\|_X ,$$

(compare to *Newton-Mysovskii*) allows to prove convergence of Newton and quasi Newton methods, cf., e.g., [Burger & BK 2006, Deuflhardt & Engl & Scherzer 1998, BK 1997, 1998].

- a relaxed "difference instead of differential" version is

$$\exists x_0 \in U \, , \; K \in L(X,Y) \, \forall x \in U \, \exists r(x) \in X \; : \; F(x) - F(x_0) = Kr(x)$$

[BK 2022]

Why should $F'()$ invariance be easier to verify than $F'()^\star$ invariance?

Reduced setting: $\mathbf{F}(q) = CS(q)$ with (nonlinear) parameter-to-state operator S and linear observation operator C:

Why should $F'()$ invariance be easier to verify than $F'()^\star$ invariance?

Reduced setting: $\mathbf{F}(q) = CS(q)$ with (nonlinear) parameter-to-state operator S and linear observation operator C:

$$\operatorname{rng}(\mathbf{F}'(x)) = \operatorname{rng}(\mathbf{F}'(\tilde{x})) \impliedby \operatorname{rng}(S'(x)) = \operatorname{rng}(S'(\tilde{x}))$$

but

$$\operatorname{rng}(\mathbf{F}'(x)^\star) = \operatorname{rng}(\mathbf{F}'(\tilde{x})^\star) \not\Longleftarrow \operatorname{rng}(S'(x)^\star) = \operatorname{rng}(S'(\tilde{x})^\star)$$

Why should $F'()$ invariance be easier to verify than $F'()^\star$ invariance?

Reduced setting: $\mathbf{F}(q) = CS(q)$ with (nonlinear) parameter-to-state operator S and linear observation operator C:

$$\mathrm{rng}(\mathbf{F}'(x)) = \mathrm{rng}(\mathbf{F}'(\tilde{x})) \impliedby \mathrm{rng}(S'(x)) = \mathrm{rng}(S'(\tilde{x}))$$

but

$$\mathrm{rng}(\mathbf{F}'(x)^\star) = \mathrm{rng}(\mathbf{F}'(\tilde{x})^\star) \not\Longleftarrow \mathrm{rng}(S'(x)^\star) = \mathrm{rng}(S'(\tilde{x})^\star)$$

Similarly for all-at-once formulation

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} q(x)u_{tt}(x,t) - \triangle u(x,t) \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

PDE will actually be considered in a variational form with inital and boundary conditions

$C \in L(V,Y)\ldots$ boundary observations

$c\ldots$ wave speed / sound speed / (velocity)

$q = \frac{1}{c^2}\ldots$ squared slowness

$u\ldots$ pressure

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} q u_{tt} - \triangle u \\ C u \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

$$K[\underline{dq},\underline{du}] = \mathbb{F}'(q_0,u_0)[\underline{dq},\underline{du}] = \begin{pmatrix} \underline{dq}\, u_{0,tt} + q_0 \underline{du}_{tt} - \triangle \underline{du} \\ C\underline{du} \end{pmatrix}$$

$$\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = \begin{pmatrix} (q-q_0)\, u_{tt} + q_0 (u-u_0)_{tt} - \triangle (u-u_0) \\ C(u-u_0) \end{pmatrix}$$

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} q u_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

$$K[\underline{dq},\underline{du}] = \mathbb{F}'(q_0,u_0)[\underline{dq},\underline{du}] = \begin{pmatrix} \underline{dq}\, u_{0,tt} + q_0 \underline{du}_{tt} - \triangle \underline{du} \\ C\underline{du} \end{pmatrix}$$

$$\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = \begin{pmatrix} (q-q_0)\, u_{tt} + q_0 (u-u_0)_{tt} - \triangle (u-u_0) \\ C(u-u_0) \end{pmatrix}$$

Thus, $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds
(for arbitrary $C \in L(V,Y)$) with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} q u_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

$$K[\underline{dq},\underline{du}] = \mathbb{F}'(q_0,u_0)[\underline{dq},\underline{du}] = \begin{pmatrix} \underline{dq}\, u_{0,tt} + q_0 \underline{du}_{tt} - \triangle \underline{du} \\ C\underline{du} \end{pmatrix}$$

$$\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = \begin{pmatrix} (q-q_0)\, u_{tt} + q_0 (u-u_0)_{tt} - \triangle (u-u_0) \\ C(u-u_0) \end{pmatrix}$$

Thus, $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds (for arbitrary $C \in L(V,Y)$) with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} q u_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1+\frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u-u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

Division by $u_{0,tt}$?

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} dq \\ du \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

Division by $u_{0,tt}$?

No. We can choose $u_0 \neq S(q_0)$

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition
$\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

Lack of regularity of u_0, u?

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition
$\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

Lack of regularity of u_0, u?
No. All-at-once allows to choose parameter and state space freely "independent of PDE theory".

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

$\underline{dq}$ depends on x **and** t (via u, u_0)
but q is supposed to be a function of x only!

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

$\underline{dq}$ depends on x **and** t (via u, u_0)
but q is supposed to be a function of x only!
Yes. To cope with this, we extend q to be a function of (x,t).

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} qu_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

$\underline{dq}$ depends on x **and** t (via u, u_0)
but q is supposed to be a function of x only!
Yes. To cope with this, we extend q to be a function of (x,t).
But this destroys uniqueness!

An example: FWI/ ultrasound tomography in all-at-once form

$$\mathbb{F}(q,u) = \begin{pmatrix} q u_{tt} - \triangle u \\ Cu \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

The difference form range invariance condition $\mathbb{F}(q,u) - \mathbb{F}(q_0,u_0) = Kr(q,u)$ holds formally with

$$r(q,u) = \begin{pmatrix} \underline{dq} \\ \underline{du} \end{pmatrix} = \begin{pmatrix} (q-q_0)\left(1 + \frac{(u-u_0)_{tt}}{u_{0,tt}}\right) \\ u - u_0 \end{pmatrix}$$

What could possibly be wrong with this formula?

$\underline{dq}$ depends on x **and** t (via u, u_0)
but q is supposed to be a function of x only!
Yes. To cope with this, we extend q to be a function of (x,t).
But this destroys uniqueness!
Yes. To cope with this, we penalize time dependence by a term $\|P(q,u)\|^2 := \|q_t\|^2$ in the methods below.

Structure exploiting regularization methods

back to $x = (q, u)$ or $x = q \ldots$

Structure exploiting regularization methods

back to $x = (q, u)$ or $x = q \ldots$

$$\begin{aligned} F(x) &= y \quad \text{(possibly extended } x \text{ to enable range invariance)} \\ Px &= 0 \quad \text{(penalization to enforce uniqueness)} \end{aligned} \qquad \text{(IP)}$$

Structure exploiting regularization methods

back to $x = (q, u)$ or $x = q \ldots$

$$\begin{aligned} F(x) &= y \quad \text{(possibly extended } x \text{ to enable range invariance)} \\ Px &= 0 \quad \text{(penalization to enforce uniqueness)} \end{aligned} \qquad \text{(IP)}$$

Assume

$$\exists x_0 \in U\,,\ K \in L(X, Y)\, \forall x \in U\, \exists r(x) \in X\ :\ F(x) - F(x_0) = Kr(x),$$

Structure exploiting regularization methods

back to $x = (q, u)$ or $x = q \ldots$

$$\begin{aligned} F(x) &= y \quad \text{(possibly extended } x \text{ to enable range invariance)} \\ Px &= 0 \quad \text{(penalization to enforce uniqueness)} \end{aligned} \qquad \text{(IP)}$$

Assume

$$\exists x_0 \in U\,,\ K \in L(X,Y)\, \forall x \in U\, \exists r(x) \in X\ :\ F(x) - F(x_0) = Kr(x),$$

Then (IP) is equivalent to
(see also Rem 2.2 [Deuflhardt & Engl & Scherzer, 1998])

$$\begin{aligned} K\hat{r} &= y - F(x_0) && \text{linear ill-posed} \\ r(x) &= \hat{r} && \text{nonlinear well-posed} \\ \mathcal{P}(x) &:= \|P(x)\| = 0 \end{aligned}$$

Structure exploiting regularization methods

(IP) is equivalent to

$K\hat{r} = y - F(x_0)$	linear ill-posed
$r(x) = \hat{r}$	nonlinear well-posed
$\mathcal{P}(x) := \|P(x)\| = 0$	

Structure exploiting regularization methods

(IP) is equivalent to

$$K\hat{r} = y - F(x_0) \qquad \text{linear ill-posed}$$
$$r(x) = \hat{r} \qquad \text{nonlinear well-posed}$$
$$\mathcal{P}(x) := \|P(x)\| = 0$$

Variational Regularization

$$(\hat{r}^\delta_{\alpha,\beta}, x^\delta_{\alpha,\beta}) \in \text{argmin}_{(\hat{r},x)\in X\times U} J^\delta_{\alpha,\beta}(\hat{r}, x)$$
$$\text{where } J^\delta_{\alpha,\beta}(\hat{r}, x) := \|K\hat{r} + F(x_0) - y^\delta\|^p_Y + \alpha\mathcal{R}(\hat{r}) + \beta\|r(x) - \hat{r}\|^b_X + \mathcal{P}(x)$$

under some continuity/regularity conditions

- $r(x_n) \xrightarrow{\mathcal{T}} \hat{r} \;\Rightarrow\; \exists (x_{n_k})_{k\in\mathbb{N}} \subseteq U,\, x \in U:\; (x_{n_k} \xrightarrow{\mathcal{T}} x \text{ and } r(x) = \hat{r})$
- sublevel sets of $\mathcal{R}$ are $\mathcal{T}$ compact; $\mathcal{P}$ is $\mathcal{T}$ lower semicontinuous
- K is $\mathcal{T}$-to-$\mathcal{T}_Y$ continuous and $\|\cdot\|$ is $\mathcal{T}_Y$ lower semicontinuous

Structure exploiting regularization methods

(IP) is equivalent to

$$
\begin{aligned}
&K\hat{r} = y - F(x_0) && \text{linear ill-posed} \\
&r(x) = \hat{r} && \text{nonlinear well-posed} \\
&\mathcal{P}(x) := \|P(x)\| = 0
\end{aligned}
$$

Structure exploiting regularization methods

(IP) is equivalent to

$$K\hat{r} = y - F(x_0) \qquad \text{linear ill-posed}$$
$$r(x) = \hat{r} \qquad \text{nonlinear well-posed}$$
$$\mathcal{P}(x) := \|P(x)\| = 0$$

Frozen Newton

$$x_{n+1}^\delta \in \text{argmin}_{x\in U}\|K(x - x_n^\delta) + F(x_n^\delta) - y^\delta\|_Y^p + \alpha_n\mathcal{R}(x) + \mathcal{P}(x).$$

under the condition

$$\exists c \in (0,1)\,\forall x \in U \,:\, \|(r(x^\dagger) - r(x)) - (x^\dagger - x)\|_X \leq c\|x^\dagger - x\|_X$$

Structure exploiting regularization methods

(IP) is equivalent to

$$K\hat{r} = y - F(x_0) \qquad \text{linear ill-posed}$$
$$r(x) = \hat{r} \qquad \text{nonlinear well-posed}$$
$$\mathcal{P}(x) := \|P(x)\| = 0$$

Newton

$$(\hat{r}^\delta_{n+1}, x^\delta_{n+1}) \in \text{argmin}_{(\hat{r},x)\in X\times U} J^\delta_n(\hat{r}, x)$$

where $J^\delta_n(\hat{r}, x) := \|K\hat{r} + F(x_0) - y^\delta\|^p_Y + \alpha_n \mathcal{R}(\hat{r})$
$$+ \beta_n \|r(x^\delta_n) + r'(x^\delta_n)(x - x^\delta_n) - \hat{r}\|^b_X + \mathcal{P}(x)$$

under the condition

$r'(x_0)^{-1} \in L(X, X)$ and

$\exists L_r > 0\, \forall x \in U \,:\, \|r'(x^\dagger) - r'(x)\|_{L(X,X)} \le L_r \|x^\dagger - x\|_X < 1,$

Idea of proof for frozen Newton in Hilbert space

$$
\begin{aligned}
x_{n+1}^{\delta} - x^{\dagger} =& (K^{\star}K + P^{\star}P + \alpha_n I)^{-1} \\
& \Big(K^{\star}(y^{\delta} - y) + K^{\star}K\big((r(x^{\dagger}) - r(x_n^{\delta})) - (x^{\dagger} - x_n^{\delta})\big) \\
& \qquad + \alpha_n (x_0 - x^{\dagger})\Big)
\end{aligned}
$$

Idea of proof for frozen Newton in Hilbert space

$$x_{n+1}^\delta - x^\dagger = (K^\star K + P^\star P + \alpha_n I)^{-1} \Big(K^\star(y^\delta - y) + K^\star K\big((r(x^\dagger) - r(x_n^\delta)) - (x^\dagger - x_n^\delta)\big) + \alpha_n(x_0 - x^\dagger)\Big)$$

Under condition

$$\exists c \in (0,1)\, \forall x \in U \; : \; \|(r(x^\dagger) - r(x)) - (x^\dagger - x)\|_X \leq c\|x^\dagger - x\|_X$$

and using spectral calculus for $A := K^\star K + P^\star P$ we obtain

$$\|x_{n+1}^\delta - x^\dagger\|_X \leq \frac{\delta}{\sqrt{\alpha_n}} + c\|x_n^\delta - x^\dagger\|_X + a_n$$

with $a_n = \alpha_n \|(K^\star K + P^\star P + \alpha_n I)^{-1}(x_0 - x^\dagger)\|_X \to 0$ as $n \to \infty$
provided $x_0 - x^\dagger \in (\mathrm{nsp}(K) \cap \mathrm{nsp}(P))^\perp \subseteq \mathrm{nsp}(A)^\perp$.

Idea of proof for frozen Newton in Hilbert space

$$
\begin{aligned}
x_{n+1}^\delta - x^\dagger =& (K^\star K + P^\star P + \alpha_n I)^{-1} \\
& \Big(K^\star (y^\delta - y) + K^\star K \big((r(x^\dagger) - r(x_n^\delta)) - (x^\dagger - x_n^\delta) \big) \\
& \qquad + \alpha_n (x_0 - x^\dagger) \Big)
\end{aligned}
$$

Under condition

$$
\exists c \in (0,1) \, \forall x \in U \, : \; \|(r(x^\dagger) - r(x)) - (x^\dagger - x)\|_X \le c \|x^\dagger - x\|_X
$$

and using spectral calculus for $A := K^\star K + P^\star P$ we obtain

$$
\|x_{n+1}^\delta - x^\dagger\|_X \le \frac{\delta}{\sqrt{\alpha_n}} + c \|x_n^\delta - x^\dagger\|_X + a_n
$$

with $a_n = \alpha_n \|(K^\star K + P^\star P + \alpha_n I)^{-1}(x_0 - x^\dagger)\|_X \to 0$ as $n \to \infty$ provided $x_0 - x^\dagger \in (\text{nsp}(K) \cap \text{nsp}(P))^\perp \subseteq \text{nsp}(A)^\perp$.

$\rightsquigarrow$ verify this in applications by (existing) linearized uniqueness proofs.

some further examples

Combined diffusion and absorption identification

Identify $a(x)$ and $c(x)$ (that is, $q = (a, c)$) in

$$-\nabla \cdot (a \nabla u) + c u = 0 \text{ in } \Omega \qquad (1)$$

from the N-t-D maps $\Lambda_\lambda \in L(H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega))$ for all $\lambda \geq 0$;
that is, $(\mathrm{tr}_{\partial\Omega} u_{\lambda,n})_{\lambda \geq 0, n \in \mathbb{N}}$ where $u^{\lambda,n}$ solves
$-\nabla \cdot (a \nabla u) + (c - \lambda) u = 0$ with $\partial_\nu u^n = \varphi^n$ on $\partial\Omega$
for a basis of boundary currents $\varphi_n \in H^{-1/2}(\partial\Omega)$.

Combined diffusion and absorption identification

Identify $a(x)$ and $c(x)$ (that is, $q = (a, c)$) in

$$-\nabla \cdot (a\nabla u) + cu = 0 \text{ in } \Omega \tag{1}$$

from the N-t-D maps $\Lambda_\lambda \in L(H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega))$ for all $\lambda \geq 0$;
that is, $(\mathrm{tr}_{\partial\Omega} u_{\lambda,n})_{\lambda \geq 0, n \in \mathbb{N}}$ where $u^{\lambda,n}$ solves
$-\nabla \cdot (a\nabla u) + (c - \lambda)u = 0$ with $\partial_\nu u^n = \varphi^n$ on $\partial\Omega$
for a basis of boundary currents $\varphi_n \in H^{-1/2}(\partial\Omega)$.

steady-state diffuse optical tomography
[Arridge & Schotland 2009; Gibson & Hebden & Arridge 2005, Harrach 2012]
uniqueness: [CanutoKavian:2004]

Combined diffusion and absorption identification

Identify $a(x)$ and $c(x)$ (that is, $q = (a, c)$) in

$$-\nabla \cdot (a\nabla u) + c u = 0 \text{ in } \Omega \qquad (1)$$

from the N-t-D maps $\Lambda_\lambda \in L(H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega))$ for all $\lambda \geq 0$;
that is, $(\mathrm{tr}_{\partial\Omega} u_{\lambda,n})_{\lambda \geq 0, n \in \mathbb{N}}$ where $u^{\lambda,n}$ solves
$-\nabla \cdot (a\nabla u) + (c - \lambda)u = 0$ with $\partial_\nu u^n = \varphi^n$ on $\partial\Omega$
for a basis of boundary currents $\varphi_n \in H^{-1/2}(\partial\Omega)$.

steady-state diffuse optical tomography
[Arridge & Schotland 2009; Gibson & Hebden & Arridge 2005, Harrach 2012]
uniqueness: [CanutoKavian:2004]

Convergence of frozen Newton
with $a = a(x)$, $c = c(x) \to c(x, \lambda, n)$.

Reconstruction of a boundary coefficient

Identify the Robin coefficient $q = q(x)$ in the elliptic boundary value problem

$$\begin{aligned} -\Delta u &= \ell \text{ in } \Omega \\ \partial_\nu u + q \cdot \Phi(u) &= h \text{ on } \Gamma_R \subseteq \partial\Omega \\ \partial_\nu u &= h \text{ on } \Gamma_N \subseteq \partial\Omega \setminus \Gamma_R \\ u &= 0 \text{ on } \Gamma_D := \partial\Omega \setminus (\Gamma_R \cup \Gamma_N) \end{aligned} \tag{2}$$

from boundary observations $y = \mathrm{tr}_{\partial\Omega} u$.
Note the nonlinearity wrt u.

Reconstruction of a boundary coefficient

Identify the Robin coefficient $q = q(x)$ in the elliptic boundary value problem

$$\begin{aligned} -\Delta u &= \ell \text{ in } \Omega \\ \partial_\nu u + q \cdot \Phi(u) &= h \text{ on } \Gamma_R \subseteq \partial\Omega \\ \partial_\nu u &= h \text{ on } \Gamma_N \subseteq \partial\Omega \setminus \Gamma_R \\ u &= 0 \text{ on } \Gamma_D := \partial\Omega \setminus (\Gamma_R \cup \Gamma_N) \end{aligned} \tag{2}$$

from boundary observations $y = \mathrm{tr}_{\partial\Omega} u$.
Note the nonlinearity wrt u.

Convergence of frozen Newton
without any extension/penalization of q being needed.

Nonlinearity coefficient imaging

Identify the squared slowness $s = s(x)$ and the nonlinearity coefficient $\eta = \eta(x)$ in the fractionally damped Westervelt equation

$$\left(su - \eta u^2\right)_{tt} - \triangle u + \tilde{D}u = \tilde{r} \quad \text{in } \Omega \times (0, T)$$
$$\partial_\nu u + \gamma u = 0 \text{ on } \partial\Omega \times (0, T), \quad u(0) = 0, \quad u_t(0) = 0 \quad \text{in } \Omega.$$

from two boundary observations

$$h_i(t) = u_i(x_0, t)\,, \quad t \in (0, T)\,, \quad \text{for } r = r_i, \quad i = 1, 2. \tag{3}$$

Nonlinearity coefficient imaging

Identify the squared slowness $s = s(x)$ and the nonlinearity coefficient $\eta = \eta(x)$ in the fractionally damped Westervelt equation

$$\left(su - \eta u^2\right)_{tt} - \triangle u + \tilde{D}u = \tilde{r} \quad \text{in } \Omega \times (0,T)$$
$$\partial_\nu u + \gamma u = 0 \text{ on } \partial\Omega \times (0,T), \quad u(0) = 0, \quad u_t(0) = 0 \quad \text{in } \Omega.$$

from two boundary observations

$$h_i(t) = u_i(x_0, t)\,, \quad t \in (0,T)\,, \quad \text{for } r = r_i, \quad i = 1, 2. \tag{3}$$

see [BK & Rundell IPI 2021, Math.Comp. 2021]
uniqueness of $\eta = \eta(x)$ from N-t-D map: [Acosta & Uhlmann & Zhai 2022]

Nonlinearity coefficient imaging

Identify the squared slowness $s = s(x)$ and the nonlinearity coefficient $\eta = \eta(x)$ in the fractionally damped Westervelt equation

$$\left(su - \eta u^2\right)_{tt} - \triangle u + \tilde{D}u = \tilde{r} \quad \text{in } \Omega \times (0, T)$$
$$\partial_\nu u + \gamma u = 0 \text{ on } \partial\Omega \times (0, T), \quad u(0) = 0, \quad u_t(0) = 0 \quad \text{in } \Omega.$$

from two boundary observations

$$h_i(t) = u_i(x_0, t)\,, \quad t \in (0, T)\,, \quad \text{for } r = r_i, \quad i = 1, 2. \qquad (3)$$

see [BK & Rundell IPI 2021, Math.Comp. 2021]
uniqueness of $\eta = \eta(x)$ from N-t-D map: [Acosta & Uhlmann & Zhai 2022]

Convergence of frozen Newton
with $\eta = \eta(x)$, $s = s(x) \to (s_1(x,t), s_2(x,t))$
[BK & Rundell, 2022]

Thank you for your attention!