Regularization of linear integral equations with noisy data and noisy operator

Ismael Rodrigo Bleyer

Prof. Dr. Ronny Ramlau

Johannes Kepler Universität - Linz

Brussels - April 9, 2010.

supported by

Der Wissenschaftsfonds.

Overview

- Introduction
- Proposed method
- Main results
- Computational aspects
- Ongoing and future work

Overview

General problem

Consider a (linear) ill-posed problems of the form

$$K_0 f = g_0 \,,$$

where K_0 is a **integral operator**

$$\begin{aligned} K_0 : \quad \mathcal{U} \quad &\longrightarrow \quad \mathcal{H} \\ f \quad &\longmapsto \quad \boxed{g_0 = K_0 f}\,, \end{aligned}$$

where

$$(K_0 f)(s) := \int_\Omega k_0(s,t) f(t) dt \,.$$

The operator K_0 is generated by kernel function k_0.

"Some mathematicians still have a kind of fear whenever they encounter a Fredholm integral equation of the first kind".

Integral operator + kernel
($k \in L^2(\Omega^2)$ or $k \in C(\Omega^2)$)
$\Downarrow$
compact and **ill-posed**

Francesco Tricomi (1897-1978)

"Some mathematicians still have a kind of fear whenever they encounter a Fredholm integral equation of the first kind".

Integral operator $+$ kernel
$(k \in L^2(\Omega^2)$ or $k \in C(\Omega^2))$
$\Downarrow$
compact and **ill-posed**

Francesco Tricomi (1897-1978)

Difficulties

Measurements:

- instead of $g_0 \in \mathscr{R}(K_0)$ we have **noisy data** $g_\delta \in Y$ with

$$\left\| g_0 - g_\delta \right\| \leq \delta.$$

- instead of $K_0 \in \mathscr{L}(\mathcal{U}, \mathcal{H})$ we have a **noisy operator** $K_\epsilon \in \mathscr{L}(\mathcal{U}, \mathcal{H})$ where

$$\left\| K_0 - K_\epsilon \right\| \leq \epsilon.$$

Inverse problem:

- given g_δ find function f;
- given g_δ and k_ϵ find function f.

$$Kf = K(k, f) \qquad \text{Nonlinear!}$$

Difficulties

Measurements:

- instead of $g_0 \in \mathscr{R}(K_0)$ we have **noisy data** $g_\delta \in Y$ with

$$\left\| g_0 - g_\delta \right\| \le \delta.$$

- instead of $K_0 \in \mathscr{L}(\mathcal{U}, \mathcal{H})$ we have a **noisy operator** $K_\epsilon \in \mathscr{L}(\mathcal{U}, \mathcal{H})$ where

$$\left\| K_0 - K_\epsilon \right\| \le \epsilon.$$

Inverse problem:

- given g_δ find function f;
- given g_δ and k_ϵ find function f.

$$Kf = K(k, f) \qquad \text{Nonlinear!}$$

Overview

- Introduction
- Proposed method
- Main results
- Computational aspects
- Ongoing and future work

How to solve?

- **Tikhonov regularization** is the most widely applied methods for solving ill-posed problems

$$\begin{array}{rl} \text{minimize} & \left\| K_\epsilon f - g_\delta \right\| \\ \text{subject to} & \left\| Lf \right\| \leq M, \end{array}$$

- **Regularized total least square** is a method based on TLS Golub and Van Loan [1980], adding a stabilization term with respect to the solution x.

$$\begin{array}{rl} \text{minimize} & \left\| A - A_\epsilon \right\|_F + \left\| b - b_\delta \right\|_2 \\ \text{subject to} & \begin{cases} Ax = b \\ \left\| Lx \right\|_2 \leq M. \end{cases} \end{array}$$

Remark: discrete version $Ax = b$ (finite dimension).

Main idea

Adding constraint to **R-TLS**

$$\begin{aligned} &\text{minimize} \quad \left\|A - A_\epsilon\right\|_F + \left\|b - b_\delta\right\|_2 \\ &\text{subject to} \quad \begin{cases} Ax = b \\ \left\|Lx\right\|_2 \le M \,. \end{cases} \end{aligned}$$

- rewrite as unrestricted problem;
- apply for infinite dimensional case;
- penalize the kernel function instead of the operator;
- generalize the penalty term w.r.t. k.

Main idea

Adding constraint to **R-TLS**

$$\begin{aligned} &\text{minimize} \quad \left\|A - A_\epsilon\right\|_F + \left\|Ax - b_\delta\right\|_2 \\ &\text{subject to} \quad \begin{cases} \\ \left\|Lx\right\|_2 \le M \,. \end{cases} \end{aligned}$$

- rewrite as unrestricted problem;
- apply for infinite dimensional case;
- penalize the kernel function instead of the operator;
- generalize the penalty term w.r.t. k.

Main idea

Adding constraint to **R-TLS**

$$\begin{aligned} \text{minimize} \quad & \|A - A_\epsilon\|_F + \|Ax - b_\delta\|_2 \\ \text{subject to} \quad & \begin{cases} \|A\|_F \leq N \\ \|Lx\|_2 \leq M \,. \end{cases} \end{aligned}$$

- rewrite as unrestricted problem;
- apply for infinite dimensional case;
- penalize the kernel function instead of the operator;
- generalize the penalty term w.r.t. k.

Main idea

Adding constraint to **R-TLS**

$$\begin{aligned} &\text{minimize} \quad \left\|A - A_\epsilon\right\|_F + \left\|Ax - b_\delta\right\|_2 \\ &\text{subject to} \quad \begin{cases} \left\|A\right\|_F \leq N \\ \left\|Lx\right\|_2 \leq M \,. \end{cases} \end{aligned}$$

- rewrite as unrestricted problem;
- apply for infinite dimensional case;
- penalize the kernel function instead of the operator;
- generalize the penalty term w.r.t. k.

Motivated by R-TLS ideas, we proposed solve the following problem

$$\text{minimize} \;\; T(k,f) := \frac{1}{2} J(k,f) + \beta \mathcal{R}(k) \;, \tag{1}$$

where

$$J(k,f) = \left\| K(k,f) - g_\delta \right\|^2_{L^2(\Omega)} + \tau \left\| k - k_\epsilon \right\|^2_{L^2(\Omega^2)} + \alpha \left\| Lf \right\|^2_{L^2(\Omega)} \,,$$

α, β are the regularization parameters, τ is a weight parameter and $\mathcal{R} : X \to [0, +\infty]$ is

- proper **convex** function and
- weak lower semi-continuous (**w-lsc**).

Overview

Main results: theoretical

Proposition

Let T be the functional defined on (1) *and L be a* **positive defined** *operator. Then T is* **positive**, **weak lower semi-continuous** *and* **coercive** *functional.*

Theorem (existence)

Let the assumptions of Proposition 1 hold. Then there exists a **global minimum** *of*

$$\text{minimize} \;\; T\,(k, f)\; .$$

Theorem (stability)

- $\delta_j \rightarrow \delta$ *and* $\epsilon_j \rightarrow \epsilon$
- $g_{\delta_j} \rightarrow g_\delta$ *and* $k_{\epsilon_j} \rightarrow k_\epsilon$
- $\alpha, \beta > 0$
- (k^j, f^j) *is a minimizer of* T *with* g_{δ_j} *and* k_{ϵ_j}
- *Then there* **exists** *a convergent subsequence of* $(k^j, f^j)_j$

$$(k^{j_m}, f^{j_m}) \longrightarrow (\bar{k}, \bar{f})$$

where $(\bar{k}, \bar{f})$ *is a minimizer of* T *with* $g_\delta, k_\epsilon, \alpha$ *and* β.

Theorem (stability)

- $\delta_j \to \delta$ *and* $\epsilon_j \to \epsilon$
- $g_{\delta_j} \to g_\delta$ *and* $k_{\epsilon_j} \to k_\epsilon$
- $\alpha, \beta > 0$
- (k^j, f^j) *is a minimizer of* T *with* g_{δ_j} *and* k_{ϵ_j}
- *Then there* **exists** *a convergent subsequence of* $(k^j, f^j)_j$

$$(k^{j_m}, f^{j_m}) \longrightarrow (\bar{k}, \bar{f})$$

where $(\bar{k}, \bar{f})$ *is a minimizer of* T *with* $g_\delta, k_\epsilon, \alpha$ *and* β.

Theorem (stability)

- $\delta_j \to \delta$ *and* $\epsilon_j \to \epsilon$
- $g_{\delta_j} \to g_\delta$ *and* $k_{\epsilon_j} \to k_\epsilon$
- $\alpha, \beta > 0$
- (k^j, f^j) *is a minimizer of* T *with* g_{δ_j} *and* k_{ϵ_j}
- *Then there* **exists** *a convergent subsequence of* $(k^j, f^j)_j$

$$(k^{j_m}, f^{j_m}) \longrightarrow (\bar{k}, \bar{f})$$

where $(\bar{k}, \bar{f})$ *is a minimizer of* T *with* $g_\delta, k_\epsilon, \alpha$ *and* β*.*

For convergence results we need to define

Definition

We call $(k^\dagger, f^\dagger)$ a Φ-**minimizing solution** if

$$(k^\dagger, f^\dagger) = \arg\min_{(k,f)} \{\Phi(k,f) \;\mid\; K(k,f) = g_0, k = k_0\}.$$

where the equalities above hold almost everywhere.

We define the convex functional

$$\Phi(k,f) := \frac{1}{2}\left\|L\ f\right\|^2 + \eta \mathcal{R}\left(k\right)$$

Theorem (Convergence)

- $\delta_j \to 0$ *and* $\epsilon_j \to 0$
- $\|g_{\delta_j} - g_0\| \leq \delta_j$ *and* $\|k_{\epsilon_j} - k_0\| \leq \epsilon_j$
- $\alpha_j = \alpha(\epsilon_j, \delta_j)$ *and* $\beta_j = \beta(\epsilon_j, \delta_j)$, *s.t.* $\alpha_j \to 0$, $\beta_j \to 0$,

$$\lim_{j\to\infty} \frac{\delta_j^2 + \tau\epsilon_j^2}{\alpha_j} = 0 \quad \text{and} \quad \lim_{j\to\infty} \frac{\beta_j}{\alpha_j} = \eta$$

- (k^j, f^j) *is a minimizer of* T *with* g_{δ_j}, k_{ϵ_j}, α_j *and* β_j
- *Then there* **exists** *a convergent subsequence of* $(k^j, f^j)_j$

$$(k^{j_m}, f^{j_m}) \longrightarrow (k^\dagger, f^\dagger)$$

where $(k^\dagger, f^\dagger)$ *is a* Φ-**minimizing solution**.

Theorem (Convergence)

- $\delta_j \to 0$ *and* $\epsilon_j \to 0$
- $\left\|g_{\delta_j} - g_0\right\| \leq \delta_j$ *and* $\left\|k_{\epsilon_j} - k_0\right\| \leq \epsilon_j$
- $\alpha_j = \alpha(\epsilon_j, \delta_j)$ *and* $\beta_j = \beta(\epsilon_j, \delta_j)$, *s.t.* $\alpha_j \to 0$, $\beta_j \to 0$,

$$\lim_{j\to\infty} \frac{\delta_j^2 + \tau\epsilon_j^2}{\alpha_j} = 0 \quad \text{and} \quad \lim_{j\to\infty} \frac{\beta_j}{\alpha_j} = \eta$$

- (k^j, f^j) *is a minimizer of* T *with* g_{δ_j}, k_{ϵ_j}, α_j *and* β_j
- *Then there* **exists** *a convergent subsequence of* $(k^j, f^j)_j$

$$(k^{j_m}, f^{j_m}) \longrightarrow (k^\dagger, f^\dagger)$$

where $(k^\dagger, f^\dagger)$ *is a* Φ-**minimizing solution**.

Theorem (Convergence)

- $\delta_j \to 0$ *and* $\epsilon_j \to 0$
- $\left\|g_{\delta_j} - g_0\right\| \le \delta_j$ *and* $\left\|k_{\epsilon_j} - k_0\right\| \le \epsilon_j$
- $\alpha_j = \alpha(\epsilon_j, \delta_j)$ *and* $\beta_j = \beta(\epsilon_j, \delta_j)$, *s.t.* $\alpha_j \to 0$, $\beta_j \to 0$,

$$\lim_{j\to\infty} \frac{\delta_j^2 + \tau\epsilon_j^2}{\alpha_j} = 0 \quad \text{and} \quad \lim_{j\to\infty} \frac{\beta_j}{\alpha_j} = \eta$$

- (k^j, f^j) *is a minimizer of* T *with* g_{δ_j}, k_{ϵ_j}, α_j *and* β_j
- *Then there* **exists** *a convergent subsequence of* $(k^j, f^j)_j$

$$(k^{j_m}, f^{j_m}) \longrightarrow (k^\dagger, f^\dagger)$$

where $(k^\dagger, f^\dagger)$ *is a* Φ-**minimizing solution**.

Remark

One can generalizes the previous results on Banach settings

$$\text{minimize } T(k,f) := J(k,f) + \beta\mathfrak{R}(k) \ ,$$

where

$$J(k,f) = \frac{1}{p}\|F(k,f) - g_\delta\|_{X^p}^p + \frac{\tau}{r}\|k - k_\epsilon\|_{X^r}^r + \frac{\alpha}{p}\|f\|_{X^p}^p$$

where $F : X^r \times X^p \to X^p$ is a nonlinear operator and $R : \mathscr{D}(R) \to [0, +\infty]$ is a w-lsc and proper convex functional.

Assumptions

- X^r and X^p are, respectively, a p-smooth and r-smooth Banach space;
- the norms $\|\cdot\|_{X^p}$ and $\|\cdot\|_{X^r}$ are weakly sequentially lower semicontinuous;
- the exact data is attainable;
- the operator F is weakly sequentially lower semi-continuous;
- the operator is continuous;
- $\mathscr{D}(F)$ has non-empty interior and $\mathscr{D}(F) \cap dom\ \mathcal{R} \neq \emptyset$.

Overview

- Introduction
- Proposed method
- Main results
- Computational aspects
- Ongoing and future work

Optimality condition

If the pair $(\bar{k}, \bar{f})$ is a minimizer of $T(k, f)$, then

$$0 \in \partial T\left(\bar{k}, \bar{f}\right) = \partial\Big(J\left(\bar{k}, \bar{f}\right) + \beta \mathfrak{R}\left(\bar{k}\right)\Big)$$

We know

$$J'(k,f)(u,v) = 2\left\langle \begin{bmatrix} (K_f^* K_f + \tau I)k - (\tau k_\epsilon + K_f^* g_\delta) \\ (K_k^* K_k + \alpha L^* L)f - K_k^* g_\delta \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle_{L^2 \times L^2}.$$

and

$$\partial \mathfrak{R}\left(\bar{k}\right) = \{\xi \in L^2(\Omega^2) \;\mid\; \mathfrak{R}\left(k\right) \geq \mathfrak{R}\left(\bar{k}\right) + \langle \xi \,,\, k - \bar{k} \rangle\}$$

for all $k \in L^2(\Omega^2) \cap \mathscr{D}(R)$.

Optimality condition

If the pair $(\bar{k}, \bar{f})$ is a minimizer of $T(k, f)$, then

$$0 \in \partial T\left(\bar{k}, \bar{f}\right) = \partial\Big(J\left(\bar{k}, \bar{f}\right) + \beta\mathfrak{R}\left(\bar{k}\right)\Big)$$

We know

$$J'(k, f)(u, v) = 2\left\langle \begin{bmatrix} (K_f^* K_f + \tau I)k - (\tau k_\epsilon + K_f^* g_\delta) \\ (K_k^* K_k + \alpha L^* L)f - K_k^* g_\delta \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle_{L^2 \times L^2} .$$

and

$$\partial\mathfrak{R}\left(\bar{k}\right) = \{\xi \in L^2(\Omega^2) \ \mid \ \mathfrak{R}\left(k\right) \geq \mathfrak{R}\left(\bar{k}\right) + \langle \xi \ , \ k - \bar{k} \rangle\}$$

for all $k \in L^2(\Omega^2) \cap \mathscr{D}(R)$.

Example 1

For the choice: $\mathcal{R}(k) = \frac{1}{2}\|k\|_2^2$.

Candidates $(\bar{k}, \bar{f})$ for a minimizer of our problem have to fulfill the necessary optimality condition:

$$\begin{cases} (K_{\bar{k}}^* K_{\bar{k}} + \alpha L^* L)\bar{f} = K_{\bar{k}}^* g_\delta \\ (K_{\bar{f}}^* K_{\bar{f}} + \gamma I)\bar{k} = K_{\bar{f}}^* g_\delta + \tau k_\epsilon \end{cases}$$

where $K_k = K(k, \cdot)$ and $K_f = K(\cdot, f)$ are linear operator and $\gamma := \tau + \beta$

Example 2

For the choice: $\mathcal{R}(k) = \|k\|_1$ we know

$$\partial\mathcal{R}(k(s,t)) = \operatorname{sgn}(k(s,t)) \text{ for a.e. } (s,t) \in \Omega^2$$

where

$$\operatorname{sgn}(z) = \begin{cases} \left\{ \frac{z}{|z|} \right\} & \text{if } z \neq 0 \\ \{\xi \in \mathbb{C} \mid |\xi| \leq 1\} & \text{otherwise} \end{cases}$$

Necessary optimality condition

$$\begin{cases} (K_{\bar{k}}^* K_{\bar{k}} + \alpha L^* L)\bar{f} = K_{\bar{k}}^* g_\delta \\ (K_{\bar{f}}^* K_{\bar{f}} + \tau I)\bar{k} = K_{\bar{f}}^* g_\delta + \tau k_\epsilon - \beta \operatorname{sgn}(\bar{k}) \end{cases}$$

Example 2

For the choice: $\mathcal{R}(k) = \|k\|_1$ we know

$$\partial\mathcal{R}(k(s,t)) = \operatorname{sgn}(k(s,t)) \text{ for a.e. } (s,t) \in \Omega^2$$

where

$$\operatorname{sgn}(z) = \begin{cases} \left\{ \frac{z}{|z|} \right\} & \text{if } z \neq 0 \\ \{\xi \in \mathbb{C} \mid |\xi| \leq 1\} & \text{otherwise} \end{cases}$$

Necessary optimality condition

$$\begin{cases} (K_{\bar{k}}^* K_{\bar{k}} + \alpha L^* L)\bar{f} = K_{\bar{k}}^* g_\delta \\ (K_{\bar{f}}^* K_{\bar{f}} + \tau I)\bar{k} = K_{\bar{f}}^* g_\delta + \tau k_\epsilon - \beta \operatorname{sgn}(\bar{k}) \end{cases}$$

Remark: iterative development

- first equation: f depends of k and α.

$$f_\delta^\alpha(k) = \left(K_k^* K_k + \alpha L^* L\right)^{-1} K_k^* g_\delta.$$

- second equation: add k in both sides

$$k^{n+1} = \left(k^n + K_f^* g_\delta + \tau k_\epsilon - (K_f^* K_f + \tau I)k^n\right) - \beta\,\mathrm{sgn}(k^n)\,.$$

Iterative shrinkage-thresholding algorithm

Require: $L, g_\delta, k_\epsilon, \tau$ and $k^0 \in L^2(\Omega^2) \cap L^1(\Omega^2)$

1: $n = 0$
2: **repeat**
3: choose α and β
4: $k^{n+1} = \mathcal{S}_\beta\left(k^n + K^*_{f_\delta^\alpha(k^n)}(g_\delta - K_{f_\delta^\alpha(k^n)} k^n) + \tau(k_\epsilon - k^n)\right)$
5: **until** convergence

Example 3

Remark: Theorem 5 we have $\beta_n = \eta\alpha_n$.

$$\text{minimize}\,\frac{1}{2}J(k,f) + \alpha\Phi(k,f)$$

where

$$J(k,f) = \left\|K(k,f) - g_\delta\right\|^2 + \tau\left\|k - k_\epsilon\right\|^2$$

and

$$\Phi(k,f) = \frac{1}{2}\left\|L\ f\right\|^2 + \eta\mathcal{R}\left(k\right).$$

For a iterative scheme w.r.t. Bregman distance

$$(k^{n+1}, f^{n+1}) \in \arg\min\left\{\frac{1}{2}J(k,f) + \alpha_n D_\Phi^{\zeta^n}\left((k,f),(k^n,f^n)\right)\right\}.$$

Example 3

Remark: Theorem 5 we have $\beta_n = \eta\alpha_n$.

$$\text{minimize} \, \frac{1}{2} J(k,f) + \alpha\Phi(k,f)$$

where

$$J(k,f) = \left\| K(k,f) - g_\delta \right\|^2 + \tau \left\| k - k_\epsilon \right\|^2$$

and

$$\Phi(k,f) = \frac{1}{2} \left\| L \; f \right\|^2 + \eta \mathcal{R}\left(k\right).$$

For a iterative scheme w.r.t. Bregman distance

$$(k^{n+1}, f^{n+1}) \in \arg\min \left\{ \frac{1}{2} J(k,f) + \alpha_n D_\Phi^{\zeta^n}\left((k,f),(k^n,f^n)\right) \right\}.$$

Generalized Tikhonov with Bregman distance

Require: $(k^0, f^0) \in \mathscr{D}(F)$

1: $n = 0$

2: **repeat**

3: $(k^{n+1}, f^{n+1}) \in \arg\min \left\{ \frac{1}{2} J(k, f) + \frac{\alpha_n}{2} \left\| L(f - f^n) \right\|_2^2 + \beta_n D_{\mathcal{R}}^{\xi^n}(k, k^n) \right\}$

4: $\xi^{n+1} = \xi^n - \frac{1}{\beta_n} \Big(K_{f^n}^* (g_\delta - K_{f^n} k^n) + \tau (k_\epsilon - k^n) \Big)$

5: $n = n + 1$

6: $\alpha_n, \beta_n > 0$

7: **until** convergence

Overview

Ongoing and future work

- prove convergence rates;
- how to choose the best regularization parameter?
- *a priori* and *a posteriori* choice;
- implementations and numerical experiments.
- design Algorithm 1 for threshold on wavelet domain.

I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57(11):1413–1457, 2004. ISSN 0010-3640.

G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 17(6): 883–893, 1980. ISSN 0036-1429.

G. H. Golub, P. C. Hansen, and D. P. O'leary. Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl*, 21:185–194, 1999.

L. Justen and R. Ramlau. A general framework for soft-shrinkage with applications to blind deconvolution and wavelet denoising. *Applied and Computational Harmonic Analysis*, 26(1):43–63, 2009.

S. Lu, S. V. Pereverzev, and U. Tautenhahn. Regularized total least squares: computational aspects and error bounds. Technical Report 30, Ricam, Linz, Austria, 2007.

S. Lu, S. V. Pereverzev, and U. Tautenhahn. Dual regularized total least squeres and multi-parameter regularization. *Computational methods in applied mathematics*, 8(3):253–262, 2008.

Thank you for your kind attention!

Questions?