Regularization of linear integral equations with noisy data and noisy operator

Ismael Rodrigo Bleyer

Prof. Dr. Ronny Ramlau

Johannes Kepler Universität - Linz

Antalya - May 25-29, 2010.

Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

JKU
JOHANNES KEPLER UNIVERSITY LINZ

supported by
FWF
Der Wissenschaftsfonds.

Overview

- Introduction
- Proposed method
- Main results
- Computational aspects
- Numerical ilustration
- 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**

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

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

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

- 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

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

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

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

- 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

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

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

How to solve?

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

$$\begin{aligned} \text{minimize} \quad & \left\| A_\epsilon x - b_\delta \right\|_2 \\ \text{subject to} \quad & \left\| Lx \right\|_2 \leq M, \end{aligned}$$

- **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.

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

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

Main idea

Adding constraint to **R-TLS**

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

- rewrite as unrestricted problem;
- apply for infinite dimensional case;
- penalize the kernel function instead of the operator;
- generalize the new penalty term for a convex functional.

Main idea

Adding constraint to **R-TLS**

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

- rewrite as unrestricted problem;
- apply for infinite dimensional case;
- penalize the kernel function instead of the operator;
- generalize the new penalty term for a convex functional.

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 new penalty term for a convex functional.

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 new penalty term for a convex functional.

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

$$\text{minimize } T(k,f) := \frac{1}{2} J(k,f) + \frac{\alpha}{2} \|Lf\|^2_{L_2(\Omega)} + \beta \mathcal{R}(k) \,, \qquad (1)$$

where

$$J(k,f) = \|K(k,f) - g_\delta\|^2_{L_2(\Omega)} + \tau \|k - k_\epsilon\|^2_{L_2(\Omega^2)} \,,$$

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)

- $\delta_j \to \delta$ *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.

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

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

where the equalities above hold almost everywhere.

We define the convex functional

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

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* $\epsilon_j \to 0$
- $\|g_{\delta_j} - g_0\| \leq \delta_j$ *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**.

Remark

As a direct consequence, the choice $\beta = \eta\alpha$ satisfies the previous theorem. So, defining

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

and

$$\begin{array}{rcccc} B(k,f) & : & L_2(\Omega^2) \times L_2(\Omega) & \longrightarrow & L_2(\Omega) \times L_2(\Omega) \\ & & (k,f) & \longmapsto & (K(k,f),k) \end{array}$$

we can rewrite (1) for $\tau = 1$ as

$$\text{minimize}\, \frac{1}{2}\left\|B(k,f) - (g_\delta, k_\epsilon)\right\|^2 + \alpha\Phi(k,f)$$

where $\left\|(k,f) - (u,v)\right\|^2 := \left\|k-u\right\|^2 + \left\|f-v\right\|^2$.

Overview

Optimality condition

If the pair $(\bar{k}, \bar{f})$ is a minimizer of $T(k, f)$, then $0 \in \partial T(\bar{k}, \bar{f})$.

$$\partial\Big(\underbrace{J(\bar{k}, \bar{f}) + \alpha \left\| L\bar{f} \right\|^2}_{\text{Fréchet differentiable}} + \beta\,\mathcal{R}(\bar{k})\Big) = h'(k, f)(u, v) + \beta\,\partial\mathcal{R}(\bar{k})$$

where

$$h'(k, f)(u, v) = \left\langle \begin{bmatrix} (K_f^* K_f + \tau I)k - (\tau k_\epsilon + K_f^* g_\delta) \\ (K_k^* K_k + \alpha L^* L)f - K_k^* g_\delta \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle_{L_2 \times L_2}$$

for all $(u, v) \in L_2(\Omega^2) \times L_2(\Omega)$ and

$$\partial\mathcal{R}(\bar{k}) = \left\{ \xi \in L_2(\Omega^2) \;\middle|\; \mathcal{R}(k) \geq \mathcal{R}(\bar{k}) + \langle \xi\,, k - \bar{k} \rangle \right\}$$

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 $0 \in \partial T(\bar{k}, \bar{f})$.

$$\partial\Big(\underbrace{J(\bar{k}, \bar{f}) + \alpha \left\| L\bar{f} \right\|^2}_{\text{Fréchet differentiable}} + \beta\, \mathcal{R}(\bar{k}) \Big) = h'(k, f)(u, v) + \beta\, \partial\mathcal{R}(\bar{k})$$

where

$$h'(k, f)(u, v) = \left\langle \begin{bmatrix} (K_f^* K_f + \tau I)k - (\tau k_\epsilon + K_f^* g_\delta) \\ (K_k^* K_k + \alpha L^* L)f - K_k^* g_\delta \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle_{L_2 \times L_2}$$

for all $(u, v) \in L_2(\Omega^2) \times L_2(\Omega)$ and

$$\partial\mathcal{R}(\bar{k}) = \{\xi \in L_2(\Omega^2) \;\mid\; \mathcal{R}(k) \geq \mathcal{R}(\bar{k}) + \langle \xi \,,\, k - \bar{k} \rangle\}$$

for all $k \in L_2(\Omega^2) \cap \mathscr{D}(R)$.

Which convex function?

For the choice: $\mathcal{R}(k) = \left\|k\right\|_{L_1(\Omega^2)}$ 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: if $(\bar{k}, \bar{f})$ minimizes (1) then

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

Which convex function?

For the choice: $\mathcal{R}(k) = \|k\|_{L_1(\Omega^2)}$ 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: if $(\bar{k}, \bar{f})$ minimizes (1) then

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

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: solve as a point-fixed iteration

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

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: $k^{n+1} = \mathcal{S}_\beta\left(k^n + K^*_{f_\delta^\alpha(k^n)}(g_\delta - K_{f_\delta^\alpha(k^n)} k^n) + \tau(k_\epsilon - k^n)\right)$

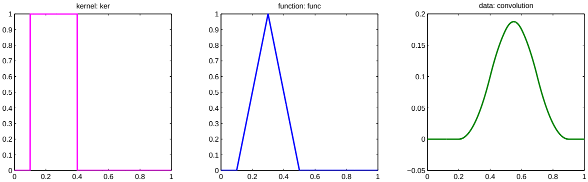
5: **until** convergence

Overview

First numerical result

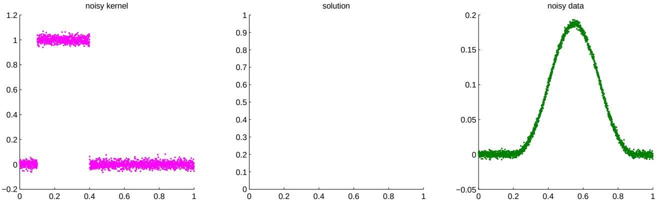
Basic problem: convolution operator in 1D

$$\int_\Omega k(s-t)f(t)dt = g(s)$$



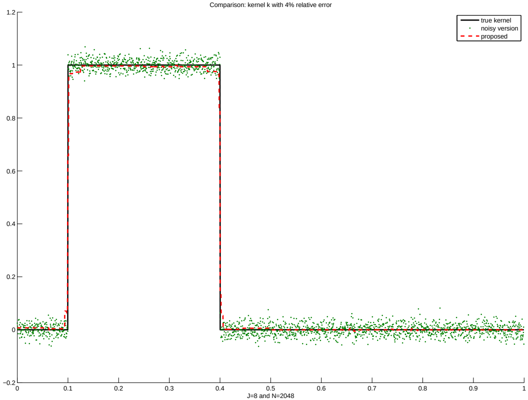
- space: $\Omega = [0,1]$, discretization: $N = 2048$ points
- initial guess: $k^0 = k_\epsilon$, $\tau = 1.0$
- relative error: 4% and 3%

Noisy kernel and data

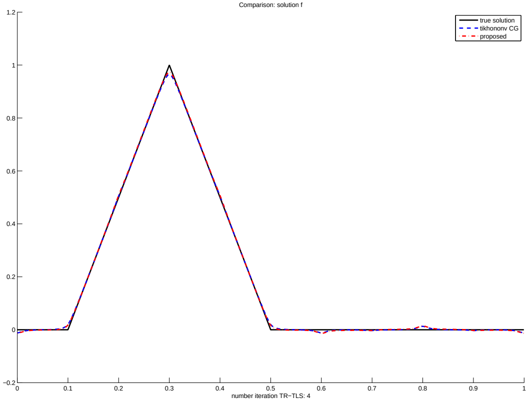


Tolerance: 10^{-7}. Iterations: Proposed 3 and Tikhonov (CG) 21.

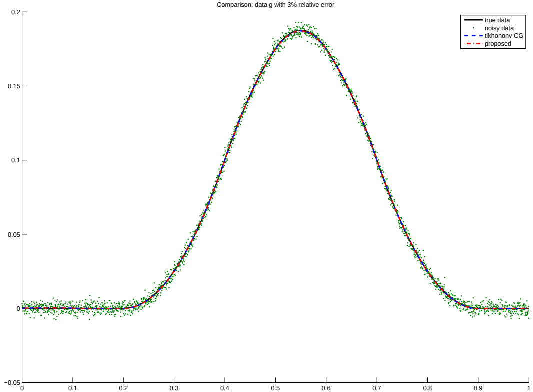
$\left\|k^n - k^{n+1}\right\|_2$	$\left\|f^n - f^{n+1}\right\|_2$	$\frac{\left\|k^n - k_0\right\|_2}{\left\|k_0\right\|_2}$	$\frac{\left\|f^n - f_0\right\|_2}{\left\|f_0\right\|_2}$
1.4479	16.4736	0.0194	0.0141
0.0001	0.1175	0.0194	0.0138
0.0000	0.0000	0.0194	0.0138



noisy kernel	0.0586
proposed	0.0194



Tikhonov	1.9954e-04
proposed	1.9013e-04



Tikhonov	5.6513e-06
proposed	5.7668e-06

Overview

Ongoing and future work

- study of source conditions;
- prove convergence rates;
- how to choose the best regularization parameter?
- *a priori* and *a posteriori* choice;
- implementations and numerical experiments (2D)

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?