Semi-discrete equations in Banach spaces: The approximate inverse approach

Andreas Rieder
Thomas Schuster
Frank Schöpfer
Saarbrücken
Oldenburg

FAKULTÄT FÜR MATHEMATIK – INSTITUT FÜR ANGEWANDTE UND NUMERISCHE MATHEMATIK

KIT – University of the State of Baden-Württemberg and National Research Center of the Helmholtz Association

www.kit.edu
AI in Banach spaces: pros and cons

Pros:
- No geometric assumptions on Banach spaces.
- Discretization effects included into the analysis.

Cons:
- Some components may not be easily available in concrete applications.
- Abstract concept; not easily presented.
Setting the stage

- $X, Y$ real Banach spaces
- $A \in \mathcal{L}(X, Y)$, $\text{R}(A) \neq \overline{\text{R}(A)}$

inverse problem: $Af = g$ where $g \in \text{R}(A)$

Remarks:

- The inverse problem is ill-posed in the sense of Nashed.
- There is no bounded inner inverse for $A$:
  \[ \forall B \in \mathcal{L}(Y, X) : ABA = A. \]
  Inner inverses are equation solvers: $\tilde{f} = Bg$ solves the inverse problem.
- In the Hilbert space setting $A^+$ (Moore-Penrose inverse) is an inner inverse which is bounded iff $\text{R}(A) = \overline{\text{R}(A)}$. 
Modelling including Measurement Process

\[ \Psi_n : Y \rightarrow \mathbb{R}^n \]

observation operator (linear, bounded)

**Example:** \( \psi_i \in Y' \) sensitivity profile of \( i \)-th detector

\[
\Psi_n g := (\langle \psi_1, g \rangle_{Y' \times Y}, \ldots, \langle \psi_n, g \rangle_{Y' \times Y})^T
\]

Semi-discrete problem (SDP):

\[
A_nf_n = g_n \quad \text{(No discretization!)}
\]

where

\[
A_n = \Psi_n A \in \mathcal{L}(X, \mathbb{R}^n), \quad g_n = \Psi_n g, \quad f_n \in X
\]

- SDP is highly under-determined!
- If \( X \) is uniformly convex then the minimum norm solution of SDP is well defined and unique.
The key concept: computing moments

Instead of trying to find a solution $f_n \in X$ of SDP we aim at computing moments:

$$\langle e_{d,i}, f_n \rangle_{X' \times X} \quad \text{for mollifiers } e_{d,i} \in X', \ i = 1, \ldots, d.$$

Then, we recover an approximate solution $E_d f_n$ of SDP by

$$E_d f_n := \sum_{i=1}^{d} \langle e_{d,i}, f_n \rangle_{X' \times X} b_{d,i}$$

where $\{b_{d,i}\}_{i=1}^{d} \subset X$.

$\{e_{d,i}\}_{i=1}^{d} \subset X'$ and $\{b_{d,i}\}_{i=1}^{d} \subset X$ have to be intertwined such that $E_d$ satisfies the mollifier property

$$\lim_{d \to \infty} \|E_d w - w\|_X = 0 \quad \text{for any } w \in X.$$
Mollifier property: an example

\[ X = (\mathcal{C}([0,1]), \| \cdot \|_\infty), \quad X' \cong BV_*(0,1): \quad f \mapsto \langle \mu, f \rangle_{X' \times X} := \int_0^1 f(x) \, d\mu(x). \]

Define

\[ E_d f = \sum_{i=0}^{d} \langle e_{d,i}, f \rangle_{X' \times X} b_{d,i} \quad \text{where} \quad \langle e_{d,i}, f \rangle_{X' \times X} = \frac{2}{d} \int_{(i-1)/d}^{(i+1)/d} f(x) \, dx. \]

Then,

\[ \lim_{d \to \infty} \| f - E_d f \|_\infty = 0 \quad \text{for any} \quad f \in \mathcal{C}(0,1). \]

Moreover,

\[ \| f - E_d f \|_\infty \leq C_E \, d^{-\alpha} \| f \|_{\mathcal{C}^\alpha(0,1)}, \quad 0 < \alpha \leq 1. \]
Reconstruction kernels: the basic idea

To evaluate $E_d f_n$ we need to calculate the moments $\langle e_{d,i}, f_n \rangle_{X' \times X}$. Let the equations

$$A^*_n v_{d,i}^n = e_{d,i}, \quad i = 1, \ldots, d, \quad (1)$$

have solutions. Then,

$$\langle e_{d,i}, f_n \rangle_{X' \times X} = \langle v_{d,i}^n, A_n f_n \rangle_2 = \langle v_{d,i}^n, g_n \rangle_2$$

and defining

$$\tilde{A}_{n,d}: \mathbb{R}^n \to X, \quad \tilde{A}_{n,d} \alpha := \sum_{i=1}^{d} \langle v_{d,i}^n, \alpha \rangle_2 b_{d,i}$$

we obtain

$$\tilde{A}_{n,d} A_n = E_d.$$  

$\tilde{A}_{n,d}$ approximate inverse of $A_n$, $v_{d,i}^n$ reconstruction kernel

However, (1) cannot be expected to have solutions.

(Louis, Maass 1990, Louis 1996)
Surrogate kernels from the underlying continuous setting

Let $A$ be injective. Given $\varepsilon > 0$ and $e \in X'$.

- If $X$ is reflexive, $\mathcal{R}(A^*)$ is dense in $X'$ and we find $v \in X'$:
  \[ \| A^* v - e \|_{X'} \leq \varepsilon. \]

- If $X$ is arbitrary, $\mathcal{R}(A^*)$ is weak-$\star$ dense in $X'$.
  Let $f \in X$ denote the (unique) solution of $Af = g$.
  Then, we find $v = v(f) \in Y'$:
  \[ |\langle A^* v - e, f \rangle_{X' \times X}| \leq \varepsilon \| f \|_X. \]

We call $v$ a continuous kernel with inaccuracy $\varepsilon$ (related to $A$, $f$, and $e$).
Let $A$ be injective. Given $\varepsilon > 0$ and $e \in X'$.

- If $X$ is reflexive, $\mathcal{R}(A^*)$ is dense in $X'$ and we find $v \in Y'$:
  \[
  \|A^*v - e\|_{X'} \leq \varepsilon.
  \]

- If $X$ is arbitrary, $\mathcal{R}(A^*)$ is weak-$\star$ dense in $X'$.
  Let $f \in X$ denote the (unique) solution of $Af = g$.
  Then, we find $v = v(f) \in Y'$:
  \[
  |\langle A^*v - e, f \rangle_{X' \times X}| \leq \varepsilon \|f\|_X.
  \]

We call $v$ a continuous kernel with inaccuracy $\varepsilon$ (related to $A$, $f$, and $e$).
Continuous kernels: an example

- $X = (C([0, 1]), \| \cdot \|_\infty)$, $X' \simeq BV_*(0, 1)$

- $A: X \to X$, $Af(x) = \int_0^x f(t)dt$

- $A^*: X' \to X'$, $A^* \mu(\xi) = \mu(1)\xi - A\mu(\xi)$.

- Let $\eta \in X'$ with $\int_0^1 \eta(t)dt = 1$ and $\eta(1) = 0$. Define $e(x) := A\eta(x)$.

  Then, $v(x) = -\eta(x)$ satisfies
  $$A^*v(\xi) = v(1)\xi - Av(\xi) = -\eta(1)\xi + A\eta(\xi) = e(\xi).$$

- Let $\eta_\gamma = \chi_{[z-\gamma, z+\gamma]} / (2\gamma)$ for a fixed $z \in ]0, 1[$ with $\gamma > 0$ suff. small. Then,

  \[
  \langle e_\gamma, f \rangle_{X' \times X} = \langle v_\gamma, Af \rangle_{X' \times X} = -\int_0^1 Af(t)d\eta_\gamma(t)
  = \frac{Af(z + \gamma) - Af(z - \gamma)}{2\gamma} \xrightarrow{\gamma \to 0} (Af)'(z) = f(z).
  \]
Surrogate kernels from the underlying continuous setting (cont’d)

Having continuous kernels $v_{d,i}$ satisfying

$$|\langle A^* v_{d,i} - e_{d,i}, f \rangle_{X' \times X}| \leq \varepsilon_{d,i} \|f\|_X, \quad i = 1, \ldots, d,$$

at our disposal we define the approximate inverse with surrogate kernels

$$\tilde{A}_{n,d} := \sum_{i=1}^d \langle \Phi_n v_{d,i}, \alpha \rangle_2 b_{d,i}.$$

Here, $\Phi_n : Y' \to \mathbb{R}^n$ is linear and continuous and needs to be defined such that

$$\lim_{n \to \infty} \lim_{d \to \infty} \|\tilde{A}_{n,d} A_n f - f\|_X = 0 \quad \text{for any } f \in X.$$
Interlude: overview of the involved operators

\[ \sum_{i=0}^{d} b_{d,i} \otimes e_{d,i} = E_d \]

\[ \text{span}\{b_{d,1}, \ldots, b_{d,d}\} \subset X \]

\[ E_d \approx \tilde{A}_{n,d} A_n \]
Theorem: If all building blocks are chosen adequately then

$$\|\tilde{A}_{n,d} A_n f - f\|_X \lesssim \|f - E_d f\|_X + \sigma(d) \left( \rho_n \max_{1 \leq i \leq d} \|v_{d,i}\|_{Y'} + \max_{1 \leq i \leq d} \varepsilon_{d,i} \right)$$

where $\rho_n \to 0$ as $n \to \infty$.

Four error components:
1. Discretization error in $X$ by mollification operator,
2. Discretization error in $Y$ by observation operator and
3. Error due to the surrogate kernels: $\rho_n = \rho_n(A f, \Psi_n, \Phi_n)$,
4. Accuracy of the continuous kernels ($A^* v_{d,i} \approx e_{d,i}$).

$$\left\| \sum_{i=1}^{d} \alpha_i b_{d,i} \right\|_X \leq \sigma(d) \max_{1 \leq i \leq d} \left| \alpha_i \right|, \quad \alpha \in \mathbb{R}^d$$
Corollary: Choosing $d = d(n) = d(n, f)$ such that $d(n) \to \infty$ as $n \to \infty$ as well as

$$
\sigma(d(n)) \rho_n \max_{1 \leq i \leq d(n)} \|v_{d(n),i}\|_{Y'} \to 0 \quad \text{as } n \to \infty
$$

and

$$
\sigma(d(n)) \max_{1 \leq i \leq d(n)} \varepsilon_{d(n),i} \to 0 \quad \text{as } n \to \infty
$$

we have the convergence

$$
\lim_{n \to \infty} \|\widetilde{A}_{n,d(n)}A_nf - f\|_X = 0.
$$
Regularization property

The noise model:

\[(\Psi^\delta_{n}y)_k = (\Psi_{n}y)_k + \delta_k, \quad |\delta_k| \leq \delta, \quad y \in Y, \quad (2)\]

where \(\delta > 0\) is the noise level.

**Theorem:** Adopt all previous assumptions.

Let \(d = d(n) = d(n, f) \to \infty\) as \(n \to \infty\) such that convergence holds in the noise-free setting.

If \(n = n_{\delta}\) such that \(n_{\delta} \to \infty\) when \(\delta \to 0\) as well as

\[\delta \sigma (d(n_{\delta})) \sqrt{n_{\delta}} \max_{1 \leq i \leq d(n_{\delta})} \| \Phi_{n_{\delta}} v_{d(n_{\delta}),i} \|_2 \to 0 \quad \text{as} \quad \delta \to 0\]

then

\[\limsup_{\delta \to 0} \left\{ \| \tilde{A}_{n_{\delta},d(n_{\delta})} \Psi^\delta_{n_{\delta}} Af - f \|_X : \Psi^\delta_{n_{\delta}} \text{ fulfills (2)} \right\} = 0.\]
Example: the integration operator

- $X = (C([0, 1]), \| \cdot \|_\infty)$, $A: X \to X$, $Af(x) = \int_0^x f(t)dt$

- Let $\psi_{n,0} = \chi_{[0,1]}$ and $\psi_{n,k} = \chi_{[k/n,1]}$, $k = 1, \ldots, n$. Then,

$$\psi_{n,k} \in X' \quad \text{and} \quad (\Psi_n g)_k = \langle \psi_{n,k}, g \rangle_{X' \times X} = g\left(\frac{k}{n}\right).$$

- $A_n: X \to \mathbb{R}^n$, $(A_n f)_k = \int_0^{k/n} f(t)dt$, $k = 0, \ldots, n$.

- $\tilde{A}_{n,d}w = \sum_{i=0}^{d} \langle \Phi_n v_{d,i}, w \rangle_2 b_{d,i}$, $v_{d,i} = -e'_{d,i}$

- $\Phi_n: Y' \to \mathbb{R}^{n+1}$ is given by

$$\left(\Phi_n v\right)_k = \langle v, b_{n,k} \rangle_{X' \times X}$$
Example: the integration operator (cont’d)

\[
\tilde{A}_{n,d}w = \sum_{i=0}^{d} \langle \Phi_n v_{d,i}, w \rangle _2 b_{d,i}, \quad v_{d,i} = -e'_{d,i},
\]

If \( n = d \) then \( \langle \Phi_n v_{d,i}, w \rangle _2 \) is easily evaluated to be

\[
\tilde{A}_{n,n}w\left(\frac{k}{n}\right) = \langle \Phi_n v_{n,k}, w \rangle _2 = n \begin{cases} 
  w_1 - w_0 & : k = 0, \\
  (w_{k+1} - w_{k-1})/2 & : k = 1, \ldots, n-1, \\
  w_n - w_{n-1} & : k = n.
\end{cases}
\]

Compare with the continuous setting

\[
\langle v_{n,k}, Af \rangle _{X' \times X} = \frac{Af\left(\frac{k+1}{n}\right) - Af\left(\frac{k-1}{n}\right)}{2 \frac{n}{n}}.
\]
Example: the integration operator (cont’d)

We have that

\[
\| \tilde{A}_{n,d} A_n f - f \|_\infty \lesssim \| f - E_d f \|_\infty + \sigma(d) \left( \max_{1 \leq i \leq d} \| v_{d,i} \|_{X'} + \max_{1 \leq i \leq d} \varepsilon_{d,i} \right)
\]

\[
\lesssim n^{-1} \left( \rho_n \max_{1 \leq i \leq d} \| v_{d,i} \|_{X'} + \max_{1 \leq i \leq d} \varepsilon_{d,i} \right) = 1 \sim d = 0
\]

Setting \( d(n) = n^{1-\gamma} \) for one \( 0 < \gamma < 1 \) yields

\[
\lim_{n \to \infty} \| f - \tilde{A}_{n,d(n)} A_n f \|_\infty = 0 \quad \text{for any } f \in C(0, 1).
\]

Consequence

We can recover about \( n \) moments of \( f \) from \( n \) observations of \( Af \).

Remark

In a comparable Hilbert space setting we can only recover about \( n^{2/3} \) moments from \( n \) observations.
Example: the integration operator (cont’d)

\[(\Psi_n^\delta g)_k = g\left(\frac{k}{n}\right) + \delta_k, \quad |\delta_k| \leq \delta\]  \hspace{1cm} (3)

Data error: \[\delta \sqrt{n} \max_{1 \leq i \leq d(n)} \|\Phi_{n,v^d(n),i}\|_2 \sim \delta \sqrt{n \cdot d}\]

Setting \[d(n) = n^{1-\gamma}, \quad 0 < \gamma < 1, \quad \text{and} \quad n_\delta \sim \delta^{\frac{\gamma - 1}{3/2 - \gamma}}\] yields

\[\limsup_{\delta \to 0} \left\{ \|\tilde{A}_{n_\delta,d(n_\delta)} \Psi_{n_\delta}^\delta A f - f\|_\infty : \Psi_{n_\delta}^\delta \text{ fulfills (3)} \right\} = 0.\]
What you should recall from this talk

Approximate Inverse

- requires no geometric assumptions on Banach spaces and
- discretization effects are included into the analysis.

Th. Schuster, A. Rieder, F. Schöpfer

*The approximate inverse in action IV: semi-discrete equations in a Banach space setting*

Inverse Problems 28 (2012)