Complexity reduction of ill-posed integral equations

Tim Jahn

Linz, 29.11.2022

Problem setting

Solve integral equation

$$g(x) = (Kf)(x) = \int k(x,y)f(y)\mathrm{d}y$$

from noisy point evaluations

$$g_m^\delta := \begin{pmatrix} g(\xi_{1,m}) + \delta Z_1 \\ \dots \\ g(\xi_{m,m}) + \delta Z_m \end{pmatrix} \in \mathbb{R}^m.$$

- $k(\cdot,\cdot)$ integral kernel,
- f unknown solution,
- m discretisation dimension,
- $\xi_{1,m}, \dots, \xi_{m,m}$ evaluation points,
- $Z_1, \dots, Z_m$ i.i.d. white noise,
- $\delta > 0$ noise level.

- *m* given by initial meausrement setup,
- Solution algorithm scale with *m*,
- Usually, *m* is not optimal for given δ and f (Pereverzyev/Mathé).

Goal:
Adapt (reduce) *m* to reduce computational complexity, without loosing accuracy.

(Discretised) SVD

$(\sigma_{j,m}, v_{j,m}, u_{j,m})$ is SVD of

$$\begin{aligned} K_m : & L^2 \to \mathbb{R}^m \\ & f \mapsto \left(\int k(\xi_{1,m}, y) f(y) \mathrm{d}y \right)_{j=1}^{m}. \end{aligned}$$

Representer Theorem: Let $(\eta_{j,m}, z_{j,m}, w_{j,m})$ be SVD of $T_m := \left(\int k(\xi_{i,m}, y) k(\xi_{j,m}, y) \mathrm{d}y \right)_{ij} \in \mathbb{R}^{m \times m}$. Then it holds that

$$\begin{aligned} \sigma_{j,m} &= \eta_{j,m}, \\ u_{j,m} &= w_{j,m}, \\ v_{j,m} &= \sum_{l=1}^{m} (z_{j,m})_l k(\cdot, \xi_l). \end{aligned}$$

Error decomposition for truncated SVD estimator

Set

$$f^{\delta}_{k,m} := \sum_{j=1}^{k} \frac{(g^{\delta}_{m}, u_{j,m})_{\mathbb{R}^m}}{\sigma_{j,m}} v_{j,m}.$$

Then

$$\mathbb{E}\|f^{\delta}_{k,m} - f\|^2 = \delta^2 \sum_{j=1}^{k} \frac{1}{\sigma^2_{j,m}} + \sum_{j=k+1}^{m} (f, v_{j,m})^2 + \|P_{\mathcal{N}(K_m)} f\|^2$$

Next step: Relate to SVD (σ_j, v_j, u_j) of K.

Example: Second anti-derivative

Let $k(x,y) := \min(x(1-y), y(1-x))$ on $[0,1]^2$ with uniform grid.

Theorem

$\sigma_j^2 = (\pi j)^{-4}$ and $v_j(x) = \sin(\sigma_j^{-1/2} x)$ are EV of K^*K and

$$\sigma_{j,m}^2 := \frac{1}{16(m+1)^3 \sin^4\left(\frac{\sigma_j^{-1/2}}{2(m+1)}\right)} \left(1 - \frac{2}{3}\sin^2\left(\frac{\sigma_j^{-1/2}}{2(m+1)}\right)\right)$$

and $z_{j,m} = \begin{pmatrix} \sin(\sigma_j^{-1/2}\xi_{1,m}) & \dots & \sin(\sigma_j^{-1/2}\xi_{m,m}) \end{pmatrix}$ are EV of $T_m^t T_m$.

Under source condtion $f = (K^*K)^{\nu/2}h$, $\|h\| \le \rho$, and $\nu \ge 1/4$ (then f is differentiable) it holds that

$$\mathbb{E}\|f_{k,m}^\delta - f\|^2 \approx \frac{\delta^2}{m}k^5 + k^{-2\nu}\rho^2 + \frac{\|f'\|^2}{m^4}$$

- If first two terms dominate (for optimal k) $\Longrightarrow$ disc. dim. m too large
- Problem: First term (data error) increases, when m decreases.

Reduce complexity by averaging

Let $l \leq m$ with $m/l =: m_l$.

Average data to decrease dimension and noise level (Anderssen/Hegland).

$$\overline{g}^{\delta}_{m_l} := \left(\frac{1}{l} \sum_{i=1}^{l} g^{\delta}_{(j-1)l+i,m} \right)_{j=1}^{m_l} \in \mathbb{R}^{m_l}$$

First,

$$\left(\overline{g}^{\delta}_{m_l}\right)_j =: \left(\overline{g}_{m_l}\right)_j + \delta \overline{Z}_j$$

with

$$\left(\overline{g}_{m_l}\right)_j = \frac{1}{l} \sum_{i=1}^{l} g(\xi_{(j-1)l+i,m}) \quad \text{and} \quad \overline{Z}_j = \frac{\delta}{l} \sum_{i=1}^{l} Z_{(j-1)l+i}.$$

Error for averaged data

- $\mathbb{E}[\overline{Z}_j^2] = \frac{1}{l}$ instead of $\mathbb{E}[Z_j^2] = 1$.
- For $g \in \mathcal{C}^2$, it holds that $|\overline{g}_j - g(\xi_{j,m_l})| \le \frac{\pi^3 l^3}{m^3} \int_{\xi_{(j-1)l,m}}^{\xi_{jl,m}} (g'')^2(x)\mathrm{d}x$.

Set $\overline{f}^{\delta}_{k,m_l} := \sum_{j=1}^{k} \frac{(\overline{g}^{\delta}_{m_l}, u_{j,m_l})_{\mathbb{R}^{m_l}}}{\sigma_{j,m_l}} v_{j,m_l}$.

Theorem

$$\mathbb{E}\|\overline{f}^{\delta}_{k,m_l} - f\|^2 \le \left(\frac{\delta^2}{l m_l} + \frac{\|g''\|^2}{m_l^4}\right) k^5 + k^{-2\nu}\rho^2 + \frac{\|f'\|^2_{\infty}}{m_l^4}$$

Remember that

$$\mathbb{E}\|f^{\delta}_{k,m} - f\|^2 \approx \frac{\delta^2}{m} k^5 + k^{-2\nu}\rho^2 + \frac{\|f'\|_{\infty}}{m^4}$$

If m_l is not too small, then $\arg\min_k \mathbb{E}\|f^{\delta}_{k,m} - f\|^2 \approx \arg\min_k \mathbb{E}\|\overline{f}^{\delta}_{k,m_l} - f\|^2$.

Example

Costs for calculating SVD of $m \times m$ matrix are $\approx m^3$

Theorem

Let $\delta = m^{-p}$ *and* $l := \left(\frac{m^3\delta^2}{\|g''\|^2}\right)^{\frac{1}{4}}$. *Then*

$$\arg\min_k \mathbb{E}\|\bar{f}^{\delta}_{k,m_l} - f\|^2 \approx \arg\min_k \mathbb{E}\|f^{\delta}_{k,m} - f\|^2,$$

and

$$\frac{\mathrm{Costs}(\bar{f}^{\delta}_{k,m_l})}{\mathrm{Costs}(f^{\delta}_{k,m})} \approx \frac{1}{l^3} \approx \left(m^3 m^{-2p}\right)^{-\frac{3}{4}} = m^{-\frac{9-6p}{4}}$$

$\Longrightarrow$ For $p < \frac{3}{2}$, substantially reduced costs.

Numerical experiment

- Initial dimension $m = m_1 = 2^{11}$
- $m_l = 2^{11}, \ldots, 2^6$
- Comparison of small/medium/large δ
- Optimal choice: $k_{\mathrm{opt}} := \arg\min_k \|\overline{f}^{\delta}_{k,m_l} - f\|$ (unfeasible)
- Adaptive choice: $\|K_{m_l}\overline{f}^{\delta}_{k_{\mathrm{dp}},m_l} - \overline{g}^{\delta}_{m_l}\| \approx \tau\|\overline{g}^{\delta}_{m_l} - K_{m_l}f\|$ (discrepancy principle), $\tau = 1.5$
- Averaged results of 50 runs

Results

Table:

	m_l	2^{11}	2^{10}	2^{9}	2^{8}	2^{7}	2^{6}
small δ	e_{opt}	3.0e-2	4.7e-2	7.3e-2	1.0e-1	1.3e-1	1.7e-1
	e_{dp}	**7.9e-2**	8.6e-2	1.2e-1	1.4e-1	1.9e-1	2.5e-1
	k_{opt}	25.7	17.1	10.5	8	6	4
	k_{dp}	**8**	7	5	4	3	2
medium δ	e_{opt}	6.6e-2	6.8e-2	7.7e-2	1.0e-1	1.4e-1	1.7e-1
	e_{dp}	1.8e-1	1.5e-1	**1.4e-1**	1.5e-1	1.9e-1	2.5e-1
	k_{opt}	11.7	11.5	9.8	8.1	6	4
	k_{dp}	3	3.9	**4**	3.9	3	2
large δ	e_{opt}	1.4e-1	1.4e-1	1.4e-1	1.4e-1	1.5e-1	1.8e-1
	e_{dp}	4.1e-1	4.1e-1	3.5e-1	2.5e-1	2.5e-1	**2.5e-1**
	k_{opt}	5.2	5.2	5.1	5	4.7	4.2
	k_{dp}	1	1	1.4	2	2	**2**

- Complexity reduction visible
- Adaptation: Choose m_l such that k_{dp} is maximal

Linear problems under white noise

Linear ill-posed equation

$$Kx = y^{\delta}.$$

- $K : X \to Y$ compact injective operator with dense range between separable Hilbert spaces.
- corrupted data $y^{\delta} = y^{\dagger} + \delta Z$
- $Z : Y \to L^2(\Omega)$ is centred white noise, i.e.
  - $\mathbb{E}[(Z, y)] = 0$ for $y \in Y$,
  - $\mathbb{E}[(Z, y)(Z, y')] = (y, y')$ for $y, y' \in Y$,
  - $(Z, y) \overset{d}{=} (Z, Y)$ for $y, y' \in Y$.

Spectral cut-off regularisation

We assume that SVD of K is known

- $(v_j)_{j\in\mathbb{N}} \subset X$ and $(u_j)_{j\in\mathbb{N}} \subset Y$ are ONBs, $\sigma_1 \geq \sigma_2 \geq \ldots > 0$
- $Kv_j = \sigma_j u_j$ and $K^* u_j = \sigma_j v_j$.

Spectral cut-off regularisation:

$$x_k^\delta := \sum_{j=1}^{k} \frac{(y^\delta, u_j)}{\sigma_j} v_j$$

approximates unknown solution $x^\dagger = K^{-1} y^\dagger$.

Determine optimal regularisation parameter $k = k(\delta, y^\delta)$.

CHOICE OF REGULARISATION PARAMETER

Discrepancy principle: Find k such that residual norm equals error norm, i.e. $\|y^\delta - Kx_k^\delta\| \approx \|y^\delta - y^\dagger\|$.

$\Longrightarrow$ not applicable since

$$\mathbb{E}\left[\|y^\delta - y^\dagger\|^2\right] = \delta^2 \mathbb{E}\left[\|Z\|^2\right] = \delta^2 \sum_{j=1}^{\infty} \mathbb{E}\left[(Z, u_j)^2\right] = \infty.$$

Vogel, Mathé, Blanchard, Reiß, Stankewitz ... studied modifications via

- Pre-smoothing: Replace K with K^*K and y^δ with K^*y^δ,
- Discretisation: Replace y^δ with $(y^\delta, e_1), ..., (y^\delta, e_m)$.

Discretised discrepancy principle

Goal: Determine optimal truncation level $k = k(\delta, y^\delta)$.

For $m \in \mathbb{N}$ there holds

$$\underbrace{\mathbb{E}\left[\sum_{j=1}^{m}(y^\delta - y, u_j)^2\right]}_{\text{discr. error norm}} = \delta^2 \sum_{j=1}^{m} \mathbb{E}[(Z, u_j)^2] = m\delta^2.$$

Discrepancy principle for discr. measurements $(y^\delta, u_1), \ldots, (y^\delta, u_m)$:

$$k_{dp}^\delta(m) := \min\left\{ k = 1, \ldots, m \ : \ \underbrace{\sqrt{\sum_{j=k+1}^{m}(y^\delta, u_j)^2}}_{\text{discr. residual norm}} \le \tau\sqrt{m}\delta \right\}$$

for $\tau > 1$.

How to choose m?

Optimal convergence

Set

$$k_{dp}^{\delta} := \max\left\{k_{dp}^{\delta}(m) \ : \ m \in \mathbb{N}\right\}.$$

Theorem: Minimax-optimal for general source conditions; oracle inequality for polynomially ill-posed problems
in probability.

$\Longrightarrow$ Benefical to use size of discretisation **as additional regulariation parameter**.

1. **J.**: *A probabilistic oracle inequality and quantification of uncertainty of a modified discrepancy principle for statistical inverse problems*, Electron. Trans. Numer. Anal., Vol. **57**, 35-56 (2022)
2. **J.**: *Optimal convergence of the discrepancy principle for polynomially and exponentially ill-posed operators under white noise*, Numer. Funct. Anal. Optim., Vol. **3** no. 2, 145-167 (2022)

GENERALISATIONS

- works with heuristic discrepancy principle instead of discrepancy principle (**J.**, arxiv 2022)
- works with Landweber iteration instead of TSVD (**J.**, arxiv 2022)