Parameter identification problems for PDEs with L^1 data fitting

Christian Clason[1] Bangti Jin[2]

[1]Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität Graz

[2]Department of Mathematics (IAMCS), Texas A&M University

SIAM Conference on Optimization
Darmstadt, May 19, 2011

L^1 fitting problem

$$(\mathcal{P}) \qquad \min_{u\in\mathcal{X}} \|S(u) - y^\delta\|_{L^1} + \frac{\alpha}{2}\|u\|_{\mathcal{X}}^2$$

- $S : \mathcal{X} \to \mathcal{Y} \subset L^1(\Omega)$ nonlinear forward operator
- $y^\delta \in L^\infty(\Omega)$ noisy measurements
- $\alpha > 0$ regularization parameter
- $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, Lipschitz boundary $\partial\Omega$

L^1 fitting problem

$$(\mathcal{P}) \qquad \min_{u\in\mathcal{X}} \|S(u) - y^\delta\|_{L^1} + \frac{\alpha}{2}\|u\|_{\mathcal{X}}^2$$

L^1 fitting more robust for non-Gaussian noise:

- large outliers
- Laplace-distributed noise
- impulsive noise (salt & pepper, random-valued)

⇝ Many applications in imaging

L^1 fitting problem

$$(\mathcal{P}) \qquad \min_{u\in\mathcal{X}} \|S(u) - y^\delta\|_{\mathrm{L}^1} + \frac{\alpha}{2}\|u\|_{\mathcal{X}}^2$$

Here: parameter identification problems for PDEs

Main assumptions:

- $S : \mathcal{X} \to \mathcal{Y}$ sufficiently differentiable
- $\mathcal{X}$ Hilbert space (e.g., L^2, H^1)
- $\mathcal{Y}$ embeds compactly into L^q, $q > 2$

Goal: Fast Newton-type methods for L^1 fitting

Elliptic model problems

1. Inverse potential problem: $S : L^2(\Omega) \to H^1(\Omega)$, $u \mapsto y$,

$$\langle \nabla y, \nabla v \rangle_{L^2} + \langle uy, v \rangle_{L^2} = \langle f, v \rangle_{L^2} \quad \text{for all } v \in H^1(\Omega)$$

2. Inverse Robin problem: $S : L^2(\Gamma_i) \to H^{1/2}(\Gamma_c)$, $u \mapsto y|_{\Gamma_c}$,

$$\langle \nabla y, \nabla v \rangle_{L^2} + \langle uy, v \rangle_{L^2(\Gamma_i)} = \langle f, v \rangle_{L^2(\Gamma_c)} \quad \text{for all } v \in H^1(\Omega)$$

3. Inverse conductivity problem, $S : H^1(\Omega) \cap L^\infty(\Omega) \to H^1_0(\Omega)$, $u \mapsto y$,

$$\langle u \nabla y, \nabla v \rangle_{L^2} = \langle f, v \rangle_{L^2} \quad \text{for all } v \in H^1_0(\Omega)$$

Common properties

(A1) S uniformly bounded in $\mathcal{X}$, $u_n \rightharpoonup u$ in $\mathcal{X}$ implies

$$S(u_n) \to S(u) \quad \text{in } L^2(\Omega)$$

(A2) S twice Fréchet differentiable

(A3) For all $u, h \in \mathcal{X}$,

$$\|S'(u)h\|_{L^2} \leq C\|h\|_{\mathcal{X}}$$
$$\|S''(u)(h,h)\|_{L^2} \leq C\|h\|_{\mathcal{X}}^2$$

$\rightsquigarrow$ sufficient conditions for approach, existence of minimizers u_α

Optimality conditions

S strictly diff., $\|\cdot\|_{L^1}$ convex, real-valued $\Rightarrow$ $(\mathcal{P})$ is Lipschitz continuous,

Theorem

For any local minimizer $u_\alpha \in \mathcal{X}$ of problem $(\mathcal{P})$, there exists a $p_\alpha \in L^\infty(\Omega)$ such that

$$\text{(OS)} \quad \begin{cases} S'(u_\alpha)^* p_\alpha + \alpha j(u_\alpha) = 0, \\ \langle S(u_\alpha) - y^\delta, p - p_\alpha \rangle_{L^2} \leq 0 \quad \text{for all } \|p\|_{L^\infty} \leq 1. \end{cases}$$

(j: duality mapping in $\mathcal{X}$)

Optimality conditions

Theorem

For any local minimizer $u_\alpha \in \mathcal{X}$ *of problem* $(\mathcal{P})$, *there exists a* $p_\alpha \in L^\infty(\Omega)$ *such that*

$$
\text{(OS)} \qquad \begin{cases} S'(u_\alpha)^* p_\alpha + \alpha j(u_\alpha) = 0, \\ \langle S(u_\alpha) - y^\delta, p - p_\alpha \rangle_{L^2} \leq 0 \quad \textit{for all } \|p\|_{L^\infty} \leq 1. \end{cases}
$$

Complementarity function for variational inequality: for any $c > 0$,

$$
\begin{aligned} S(u_\alpha) - y^\delta = {} & \max(0, S(u_\alpha) - y^\delta + c(p_\alpha - 1)) \\ & + \min(0, S(u_\alpha) - y^\delta + c(p_\alpha + 1)) \end{aligned}
$$

Optimality conditions

Complementarity function for variational inequality: for any $c > 0$,

$$\begin{aligned} S(u_\alpha) - y^\delta = \max(0, S(u_\alpha) - y^\delta + c(p_\alpha - 1)) \\ + \min(0, S(u_\alpha) - y^\delta + c(p_\alpha + 1)) \end{aligned}$$

Pointwise interpretation:

$$p_\alpha = \operatorname{sign}(S(u_\alpha) - y^\delta) = \begin{cases} 1 & S(u_\alpha) - y^\delta > 0 \\ -1 & S(u_\alpha) - y^\delta < 0 \\ \tau \in [-1, 1] & S(u_\alpha) - y^\delta = 0 \end{cases}$$

Optimality conditions

Pointwise interpretation:

$$p_\alpha = \operatorname{sign}(S(u_\alpha) - y^\delta) = \begin{cases} 1 & S(u_\alpha) - y^\delta > 0 \\ -1 & S(u_\alpha) - y^\delta < 0 \\ \tau \in [-1,1] & S(u_\alpha) - y^\delta = 0 \end{cases}$$

Reduced optimality system

(OS')

$$\alpha j(u_\alpha) + S'(u_\alpha)^*(\operatorname{sign}(S(u_\alpha) - y^\delta)) \ni 0$$

Regularization

sign not differentiable in any sense $\rightsquigarrow$ replace by sign_β for $\beta > 0$,

$$\text{sign}_\beta(u)(x) := \begin{cases} 1 & \text{if } u(x) > \beta, \\ -1 & \text{if } u(x) < -\beta, \\ \frac{1}{\beta} t & \text{if } |u(x)| \leq \beta, \end{cases}$$

(equivalent to Huber-smoothing of $(\mathcal{P})$, dual L^2 regularization)

Regularized optimality system

$$(\text{OS}_\beta) \qquad \alpha j(u_\beta) + S'(u_\beta)^*(\text{sign}_\beta(S(u_\beta) - y^\delta)) = 0$$

Regularization

Regularized optimality system

$$(\mathrm{OS}_\beta) \qquad \alpha j(u_\beta) + S'(u_\beta)^*(\mathrm{sign}_\beta(S(u_\beta) - y^\delta)) = 0$$

Theorem

(OS_β) *has a solution* u_β, *and sequence* $\{u_\beta\}_{\beta>0}$ *contains subsequence converging in* $\mathcal{X}$ *to solution* u_α *to* (OS').

$\rightsquigarrow$ Continuation strategy in $\beta \to 0$ for numerical solution

Semi-smooth Newton method

Consider (OS_β) as $F(u) = 0$ for $F : \mathcal{X} \to \mathcal{X}^*$,

$$F(u) = \alpha j(u) + S'(u)^*(\mathrm{sign}_\beta(S(u) - y^\delta))$$

$t \mapsto \mathrm{sign}_\beta(t)$ semi-smooth, $S(u) - y^\delta \in L^q$, S twice differentiable

$\Rightarrow P(u) = \mathrm{sign}_\beta(S(u) - y^\delta)$ semi-smooth, Newton derivative

$$\begin{aligned} D_N P(u)h &= \beta^{-1}(S'(u)h)\chi_{\mathcal{I}} \\ &= \begin{cases} \beta^{-1}(S'(u)h) & \text{if } |(S(u) - y^\delta)| \le \beta \\ 0 & \text{else} \end{cases} \end{aligned}$$

Semi-smooth Newton method

$\mathcal{X}$ Hilbert space, $S'(u)$ linear operator $\Rightarrow$ F semi-smooth

Semi-smooth Newton step for $\delta u = u^{k+1} - u^k$

$$\alpha j'(u^k)\delta u + (S''(u^k)\delta u)^* P(u^k) + \frac{1}{\beta} S'(u^k)^*(\chi_{\mathcal{I}^k} S'(u^k)\delta u) = -F(u^k)$$

Can be solved using matrix-free Krylov method
(given u^k, δu, rhs/lhs computed by solving forward, adjoint PDE)

Semi-smooth Newton method

But: superlinear convergence requires regularity condition, S nonlinear, functional not necessarily convex $\rightsquigarrow$ assume for $\gamma > 0$

Second order condition

$$\text{(S)} \qquad \langle S''(u_\beta)(h,h), P(u_\beta)\rangle_{L^2} + \alpha\|h\|_{\mathcal{X}}^2 \geq \gamma\|h\|_{\mathcal{X}}^2 \quad \text{for all } h \in \mathcal{X}$$

(compare second order sufficient condition)

Here: (S) holds if either

- α large (large noise)
- β large or residual small (small noise) ($\Rightarrow P(u_\beta)$ small)

Semi-smooth Newton method

Second order condition

(S) $\quad \langle S''(u_\beta)(h,h), P(u_\beta)\rangle_{L^2} + \alpha\|h\|_{\mathcal{X}}^2 \geq \gamma\|h\|_{\mathcal{X}}^2 \quad$ for all $h \in \mathcal{X}$

Theorem

If (S) *holds and u^0 is sufficiently close to u_β, then the iterates of the semi-smooth Newton method converge superlinearly to the solution u_β to* (OS_β).

Automatic parameter choice

Noise level δ unknown: choose α^* solving

Balancing equation

$$(\sigma - 1)\|S(u^\delta_{\alpha^*}) - y^\delta\|_{L^1} = \frac{\alpha^*}{2}\|u^\delta_{\alpha^*}\|^2_{\mathcal{X}}$$

($\sigma > 1$ fixed, depends on smoothness of data/solution, not noise)

Fixed point iteration

$$\alpha_{k+1} = (\sigma - 1)\frac{\|S(u^\delta_{\alpha_k}) - y^\delta\|_{L^1}}{\frac{1}{2}\|u^\delta_{\alpha_k}\|^2_{\mathcal{X}}}$$

Automatic parameter choice

Theorem

If starting value α_0 satisfies

$$(\sigma - 1)\|S(u^\delta_{\alpha_0}) - y^\delta\|_{L^1} - \frac{\alpha_0}{2}\|u^\delta_{\alpha_0}\|^2_{\mathcal{X}} < 0$$

then $\{\alpha_k\}$

- *is monotonically decreasing*
- *converges to solution of balancing equation*

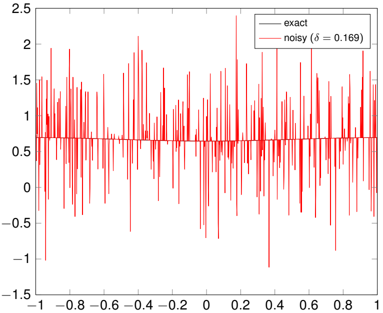
Numerical results for model problems

- Discretization using uniform linear finite elements
  1d: $N = 1001$, 2d: $N = 128 \times 128$ grid points
- Random impulsive noise: $y^\dagger = S(u^\dagger)$,

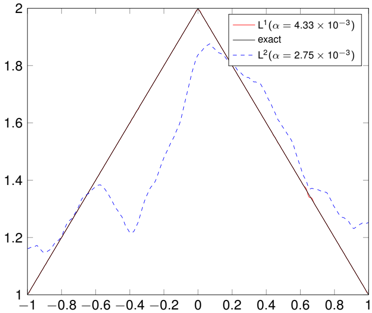
$$y^\delta = \begin{cases} y^\dagger + \|y^\dagger\|_{L^\infty}\xi, & \text{with probability } r \\ y^\dagger, & \text{otherwise} \end{cases}$$

  $\xi(x)$ normally distributed random variable
- α chosen using fixed point iteration (2–4 its.)
- Comparison with standard L^2 fitting (Newton method)

Inverse potential: $r = 0.3$

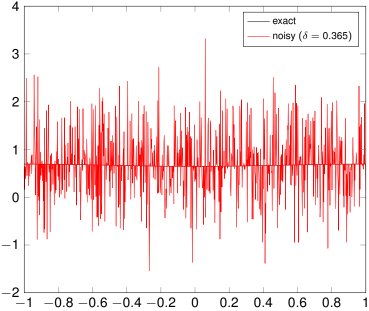


(a) data

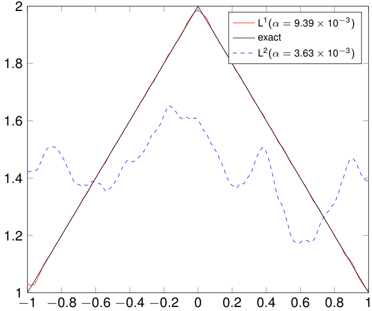


(b) reconstruction

Inverse potential: $r = 0.6$



(a) data



(b) reconstruction

Inverse potential: Balancing principle

r	δ	α_s	α_b	e_s	e_b
0.1	5.432e-2	4.508e-3	1.397e-3	2.696e-4	5.039e-4
0.2	1.037e-1	4.436e-3	2.667e-3	3.226e-4	4.879e-4
0.3	1.536e-1	4.510e-3	3.949e-3	5.229e-4	4.914e-4
0.4	2.189e-1	6.429e-3	5.629e-3	1.083e-3	1.020e-3
0.5	2.877e-1	8.449e-3	7.398e-3	5.348e-3	4.882e-3
0.6	3.201e-1	5.113e-3	8.233e-3	2.658e-3	9.632e-3
0.7	3.854e-1	6.155e-3	9.910e-3	1.039e-2	7.698e-3
0.8	4.117e-1	6.573e-3	1.058e-2	1.688e-2	3.909e-2
0.9	4.995e-1	1.478e-2	1.294e-2	5.486e-2	5.430e-2

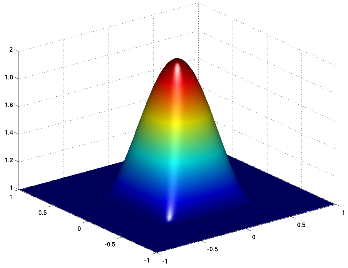
δ: noise level, e: L^2 error; $\cdot_s$: optimal choice, $\cdot_b$: balancing

Inverse potential: Performance

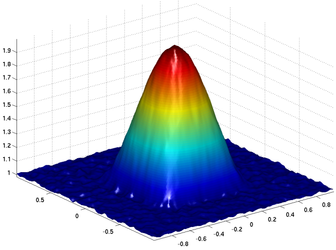
N	400	800	1600	3200	6400	12800
t_s	5.28	12.09	19.40	29.66	55.33	107.87
t_b	14.42	39.04	54.19	80.30	131.72	234.00
e	2.88e-3	9.17e-4	6.22e-4	3.52e-4	2.76e-4	2.78e-4

- N: number of elements
- t_s: computing time for semi-smooth Newton method including continuation in β (seconds, average of 10)
- t_b: computing time for fixed point iteration (choice of α)
- e: L^2 reconstruction error (average of 10)

Inverse potential (2d): $r = 0.3$

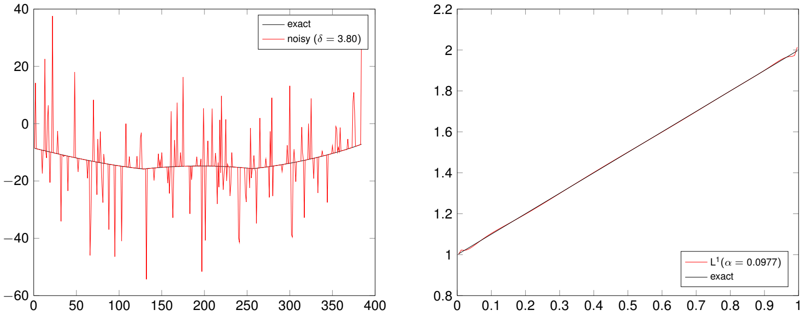


(a) exact



(b) reconstruction

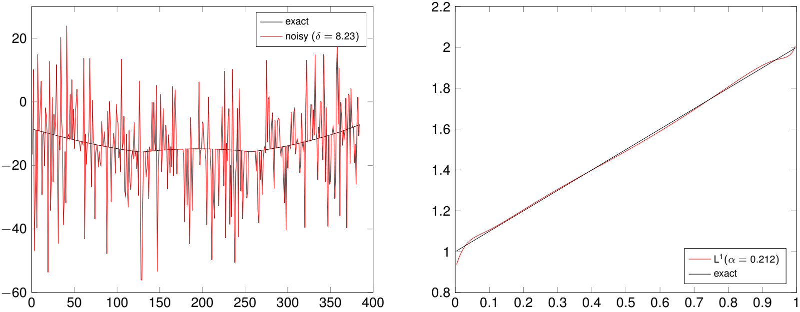
Inverse Robin: $r = 0.3$



(a) data

(b) reconstruction

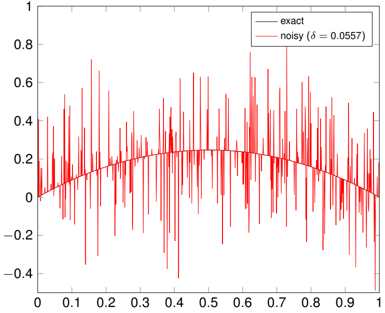
Inverse Robin: $r = 0.6$



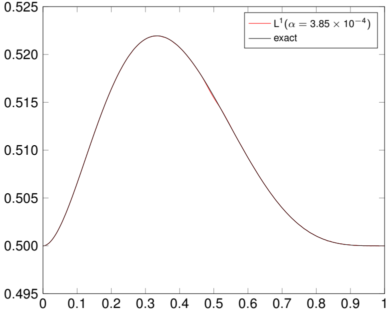
(a) data

(b) reconstruction

Inverse conductivity: $r = 0.3$

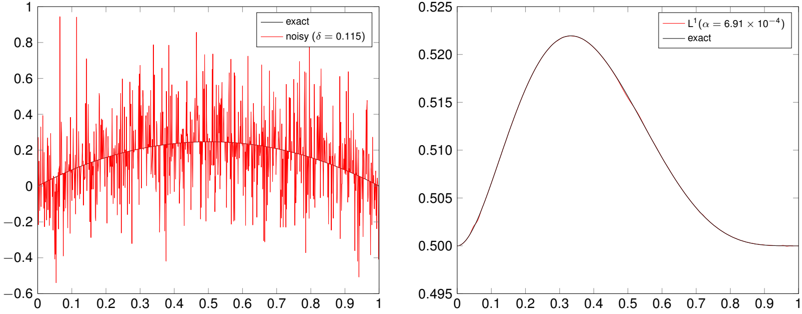


(a) data



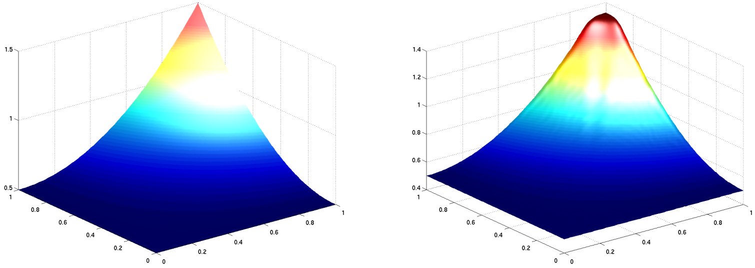
(b) reconstruction

Inverse conductivity: $r = 0.6$



(a) data (b) reconstruction

Inverse conductivity (2d): $r = 0.3$



(a) exact

(b) reconstruction

Conclusion

- Semi-smooth Newton methods for numerical solution of nonsmooth (Lipschitz) problems
- L^1 fitting very robust for impulsive noise
- Efficient heurisitic parameter choice by balancing principle

Future work

- Time dependent problems (require efficient FE solvers)
- Applications (magnetic induction, diffuse optical tomography)

Preprint, MATLAB code:

`http://www.uni-graz.at/~clason/publications.html`