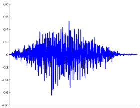
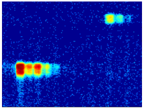
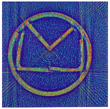
Inverse Problems with Sparsity Constraints

Peter Maaß

Center for Industrial Mathematics
University of Bremen

Helsinki, March 4th 2010

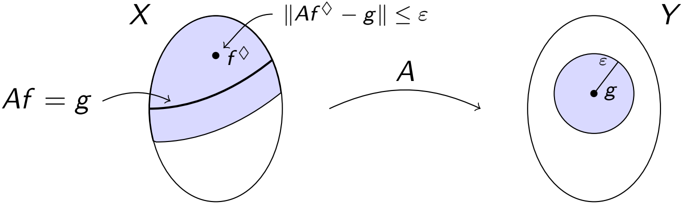



Outline

The usual setting

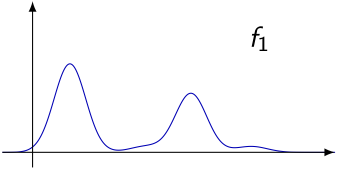
$$A : X \to Y, \quad g \in \operatorname{range}(A)$$

- solutions $\{f \in X \mid Af = g\}$
- noise $\|g^\varepsilon - g\|_Y \leq \varepsilon$
- approximate solutions $\{f \in X \mid \|Af - g^\varepsilon\| \leq \varepsilon\}$

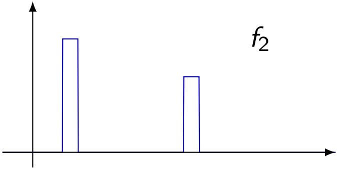


Set of possible solutions

$$\{f \mid \|Af - g^\varepsilon\| \le \varepsilon\}$$



$$\|Af_1 - g^\varepsilon\| \le \varepsilon$$



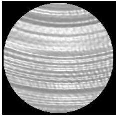
$$\|Af_2 - g^\varepsilon\| \le \varepsilon$$

$$\|f_1\| \quad < \quad \|f_2\|$$

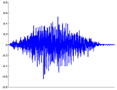
- Which one is the preferred approximation?

Sparse structures

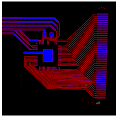
critical structures $\triangleq$ sparse representations



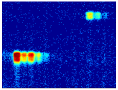
High precision surfaces



Linear guideways



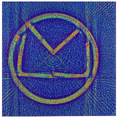
Quality control



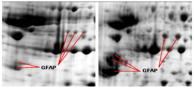
LC-MS spectra

Aero engines

Turning processes



Quality control



Test design

Sparse Decomposition

Definition

f is called sparse with respect to a basis/frame/dictionary $\{\varphi_i\}$, if

$$f = \sum_{i \in I} f_i \varphi_i, \quad |I| < \infty,$$

i.e. there exists a finitely supported decomposition of f.

Examples of bases:

- pixel basis
- Fourier basis
- object basis
- wavelets
- frames, dictionaries

Sparsity in Signal/ Image Processing

$$A = I$$

- sparsity with respect to basis $\{\varphi_i\}_{i\in\mathbb{Z}}$, i.e. $f = \sum f_i \varphi_i$
- variational approach/ Tikhonov functional

$$\min_f \|f - g^\varepsilon\|^2 + \alpha \|f\|_{\ell^p}^p$$

- shrinkage operator $f_\alpha^\varepsilon = \mathbf{S}_\alpha^p(\{g_i^\varepsilon\}) = \sum S_\alpha^p\left(\langle g^\varepsilon, \varphi_i\rangle\right) \varphi_i$, e.g.

$$S_\alpha^1(u_i) := \begin{cases} u_i - \alpha, & \text{for } u_i > \alpha, \\ 0 & \text{for } -\alpha \le u_i \le \alpha, \\ u_i + \alpha, & \text{for } u_i < \alpha. \end{cases}$$

D. Donoho. *De-noising by soft thresholding.* 1995. and many others

Overview

Sparsity and Inverse Problems

$$\min_f \|A f - g^\varepsilon\|^2 + \alpha \|f\|_{\ell^p}^p$$

- Regularization Properties

$$f_\alpha^\varepsilon \to f^\dagger, \quad \varepsilon \to 0$$

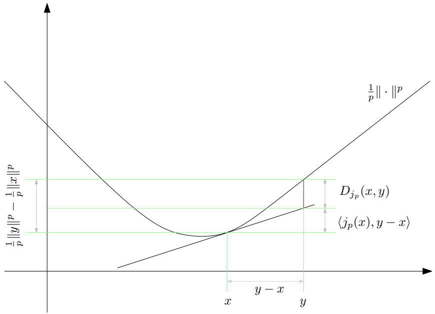
- Iterated Soft Shrinkage

$$f^{(k+1)} = \mathbf{S}_\alpha^p \left[f^{(k)} - A^* (A f^{(k)} - g^\varepsilon) \right]$$

Daubechies, Defrise, De Mol. *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint.* 2004.

Bregman distance

$$D_{j_p}(x,y) := \frac{1}{p}\|y\|^p - \frac{1}{p}\|x\|^p - \langle j_p(x), y-x\rangle$$



- measures the gap between the functional $\frac{1}{p}\|x\|^p$ and its linearization
- in $\ell^p, L^p, W^p_k, B^s_{p,q}$ with $1 < p,q < \infty$ can be bounded from above resp. below by some power of $\|x-y\|$
- better suited for analysis than the norm

Duality mapping

- may be multivalued (j_p selection of J_p)
- $\langle j_p(x), x\rangle = \|x\|^p$ and $\|j_p(x)\| = \|x\|^{p-1}$
- in $\ell^p, L^p, W_k^p, B_{p,q}^s$ with $1 < p, q < \infty$

$$J_p = \partial(\tfrac{1}{p}\|\cdot\|^p)$$

$$J_{p^*}(J_p(x)) = x \quad \text{and} \quad J_p(J_{p^*}(x^*)) = x^*$$

- needed for Bregman distances, Source Conditions, Minimization schemes

Major research directions

- Generalizations
  - Non-linear operators
  - X, Y Banach spaces
  - Iteration methods for sparse approximations
- Some open problems
  - Efficient algorithms for minimizing Tikhonov functionals
  - Source conditions
  - Applications

Other sparse/local reconstruction schemes

- BV-regularization, $A : BV(\Omega) \to L^2(\Omega)$,

$$\min \|Af - g^\varepsilon\| + \alpha \sup_{\|p\|_\infty \leq 1} \int_\Omega f \operatorname{div} p \, \mathrm{d}x$$

Rudin, Osher, Fatemi. *Nonlinear total variation based noise removal algorithms.* 1992.
Burger, Osher. *Convergence rates of convex variational regularization.* 2004.

- Compressive sampling Herrholz, Teschke. *Compressive sensing in inverse problems. 2010.*
- Sampling methods/ factorization schemes → A. Kirsch
- Mollifier methods → A.K. Louis

Overview

- Non-linear operators

$$\min_f \|A(f) - g^\varepsilon\|^2 + \alpha \|f\|_{\ell^p}^p$$

$$f^{(k+1)} = \mathbf{S}_\alpha^p \left[f^{(k)} - [A'(f^{(k)})]^* (A(f^{(k)}) - g^\varepsilon) \right]$$

Bredies, Bonesky, Lorenz, Maass. *A Generalized Conditional Gradient Method for Non-Linear Operator* 2007
Ramlau, Teschke. *An Iterative Algorithm for Nonlinear Inverse Problems* 2007.
Grasmair, Haltmeier, Scherzer. *Sparse Regularization with ℓ^q Penalty Term.* 2008.

- Non-linear operators with projections on ℓ_1 ball

$$\min_{f \in B_R} \|A(f) - g^\varepsilon\|^2$$

$$f^{(k+1)} = \mathbf{P}_{B_R} \left[f^{(k)} - \frac{\beta^k}{r} [A'(f^{(k+1)})]^* (A(f^{(k)}) - g^\varepsilon) \right]$$

Teschke, Borries. *Accelerated Projected Steepest Descent Methods,* 2010

Tikhonov-functionals T_α in Banach spaces

$$T_\alpha(x) = \frac{1}{r}\|Ax - y^\delta\|_Y^r + \alpha\frac{1}{p}\|x\|_X^p$$

X, Y Banach spaces

$$x_{n+1} = j_{p^*}(j_p(x_n) - \mu_n \partial T_\alpha(x_n)) \qquad x_{n+1}(x) = x_n - \mu_n j_{p^*}(\partial T_\alpha(x_n))$$

- update in the dual space or in the prinal space
- Tikhonov or generalized Landweber-iteration ($\alpha = 0$)

Schöpfer, Louis, Schuster. *Nonlinear iterative methods for linear ill-posed problems in Banach spaces.* 2006.

Hofmann, Kaltenbacher, Poeschl, Scherzer. *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators.* 2007.

Bonesky etal 2008, Kazimierski 2009

Efficient minimization Algortihms

$$\min_f \|Af - g^\varepsilon\|^2 + \alpha \|f\|_{\ell^p}^p$$

- active set methods
  Griesse, Lorenz. *A Semismooth Newton Method for Tikhonov Functionals with Sparsity Constraints.* 2008.
  Jin, Lorenz, Schiffler. *Elastic-Net Regularization: Error estimates and Active Set Methods.* 2009.
- iterative methods
  Beck, Teboulle. *A fast iterative shrinkage-thresholding algorithm for linear inverse problems.* 2009.
- gradient descent
  Bredies, Lorenz, Maass. *A generalized conditional gradient method and its connection to an iterative shrinkage method.* 2008.
  Figueiredo, Nowak, Wright. *Gradient Projection for Sparse Reconstruction: Applications to Compressed Sensing and Other Inverse Problems.* 2007.
- forward-backward splitting
  Combettes, Wajs. *Signal recovery by proximal forward-backward splitting.* 2005.
  Bredies. *A forward-backward splitting algorithm for the minimization of non-smooth convex functionals in Banach space.* 2009.
- semi-smooth Newton method Griesse-Lorenz(2008)
- FISTA Beck-Teboulle (2009), SpaRSA Wright-Nowak-Figueiredo (2009)

Source Conditions

Assumptions for ℓ^1 penalty

- K has finite basis injectivity property
- minimum-$\|\cdot\|_{\ell^1}$-norm-solution $f^\dagger$ is finitely supported
- source condition

$$\operatorname{range} A^* \cap \operatorname{Sign}(f^\dagger) \neq \emptyset.$$

Bredies, Lorenz. *Linear convergence of iterative soft-thresholding.* 2008.
Lorenz. *Convergence rates and source conditions for Tikhonov regularization with sparsity constraints.* 2008.
Grasmair, Haltmeier, Scherzer. *Sparse Regularization with ℓ^q Penalty Term.* 2008.

Extended concepts

- approximate source conditions Hofmann, Düvelmeyer, Krumbiegel 2006
- variational source conditions Hofmann, Kaltenbacher, Pöschl, Scherzer 2007, Hein

Overview

Impedance tomography (EIT)

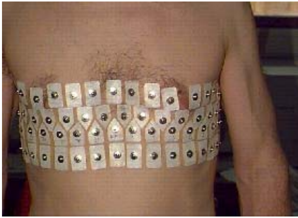
iterative soft shrinkage algorithm Daubechies et al (2004)

$$f^{n+1} = \mathcal{S}_\alpha \left(f^n - [A'(f^n)]^* (A(f^n) - g^\delta) \right)$$

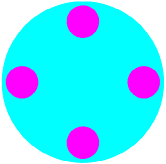
determine σ from EIT measurements

$$\begin{cases} -\mathrm{div}(\sigma \nabla u) = 0 \text{ in } \Omega \\ \sigma \frac{\partial u}{\partial n} = j \text{ on } \partial\Omega \end{cases}$$

forward solver: $u = F^\sigma(j)$
measurements (restr. of u on $\partial\Omega$) $g = \gamma_0 u$



Sparsity reconstruction Bangti Jin, T. Khan, P. Maass



- assumption: known background σ_0
- then calculating $\sigma^\dagger = \sigma_0 + \delta\sigma^\dagger$
- $\delta\sigma^\dagger$: localized inclusion

min-max problem Kohn-Vogelius (1986), Knowles (1998)

$$\min_{\sigma} \max_{j} \|\underbrace{\gamma_0 F^{\sigma}(j)}_{\text{simulation}} - \underbrace{\gamma_0 F^{\sigma^\dagger}(j)}_{\text{experiment}}\|$$

reconstruction algorithm

$$\delta\sigma^{n+1} = \mathcal{S}_\alpha\left(\delta\sigma^n - \left[\frac{\partial}{\partial\sigma}\gamma_0 F^{\sigma^n}(j)\right]^* (\gamma_0 F^{\sigma^n}(j) - \gamma_0 F^{\sigma^\dagger}(j))\right)$$

Analytic prerequisites

$$\frac{d}{d\sigma}F^{\sigma}(j) : L_{\infty}(\Omega) \to H^1(\Omega)$$

Kaipio et al (2000), Dorn-Natterer et al(2002), Rieder-Lechleiter(2008)

adjoint not defined

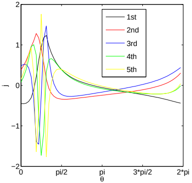
$$\frac{d}{d\sigma}F^{\sigma}(j) : L_{q}(\Omega) \to H^1(\Omega)$$

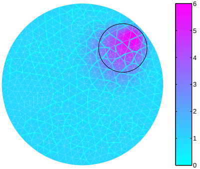
algorithm

- step size selection: Barzilai-Borwein rule
- Sobolev smoothing

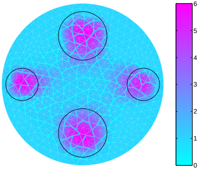
Ref: Jin B, Khan T, Maass P. Sparse reconstruction in electrical impedance tomography, preprint.

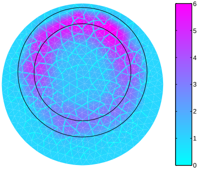
Results (continuum model)





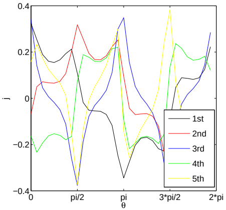
single inclusion (1% noise)

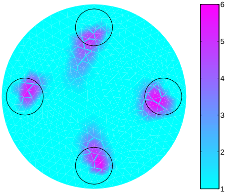




complex model (5% noise)

Results (../bilder/complete electrode model)





32 electrodes, 2% noise

- real data, 3D: Kaipio et al (2010),
- finite element analysis: Chen-Zou(1999), Jin-Zou(2009)