Solving inverse problems in imaging with Shearlab.jl

Héctor Andrade Loarca
(github: **arsenal9971**)

Notebook and Beamer:
https://github.com/arsenal997/Shearlab.jl/presentations/SIAM-IS

TU Berlin, BMS

8^{th} of June, 2018

Applied Functional Analysis Group
Technische Universität Berlin

Inverse problems in Imaging

Goal

Recover parameters characterizing a system under investigation from measurements (recover image from data).

Inverse problems in Imaging

Goal

Recover parameters characterizing a system under investigation from measurements (recover image from data).

Mathematical formulation

Recover $f_{\text{true}} \in X$ from data

$$g = \mathcal{T}(f_{\text{true}}) + \delta g$$

where $y \in Y$,$\mathcal{T} : X \longrightarrow Y$ and $\delta g \in Y$.

Inverse problems in Imaging

Goal

Recover parameters characterizing a system under investigation from measurements (recover image from data).

Mathematical formulation

Recover $f_{\text{true}} \in X$ from data

$$g = \mathcal{T}(f_{\text{true}}) + \delta g$$

where $y \in Y$,$\mathcal{T} : X \longrightarrow Y$ and $\delta g \in Y$.

- **Classical solution:** Minimization of the miss-fit against data:

  $$\min_{f \in X} \mathcal{L}(\mathcal{T}(f), g)$$

  $\mathcal{L} : Y \times Y \longrightarrow \mathbb{R}$ is a transformation of the negative data log-likelihood, e.g. $L(f) = ||\mathcal{T}(f) - g||_2^2$.

Ill-posedness and regularization

Hadamard well-posedness

Existance and uniqueness of solution for all data, and continuous dependence of solution on the data.

Ill-posedness and regularization

Hadamard well-posedness

Existance and uniqueness of solution for all data, and continuous dependence of solution on the data.

- Ill-posed problems produce overfitting when minimizing the data miss-fit, but they are the most common in applications (CT, EEG, MRI,...).

Ill-posedness and regularization

Hadamard well-posedness

Existance and uniqueness of solution for all data, and continuous dependence of solution on the data.

- Ill-posed problems produce overfitting when minimizing the data miss-fit, but they are the most common in applications (CT, EEG, MRI,...).
- **Regularization:** Set of methods to avoid overfitting by slightly modify the original problem to increase its regularity.

Ill-posedness and regularization

Hadamard well-posedness

Existance and uniqueness of solution for all data, and continuous dependence of solution on the data.

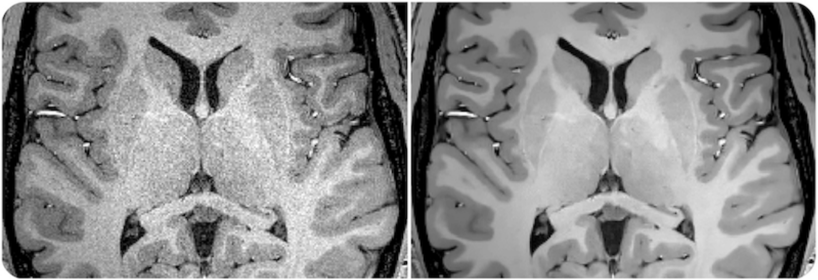
- Ill-posed problems produce overfitting when minimizing the data miss-fit, but they are the most common in applications (CT, EEG, MRI,...).
- **Regularization:** Set of methods to avoid overfitting by slightly modify the original problem to increase its regularity.
- **Variational regularization:** Introduces a "regularization functional" $\mathcal{S} : X \longrightarrow \mathbb{R}$ to encode a priori information about f_{true}, obtaining a new objective functional to minimize:

$$\min_{f \in X} \left[\mathcal{L}(\mathcal{T}(f), g) + \lambda \mathcal{S}(f) \right] \quad \text{for a fixed } \lambda \geq 0$$

Some examples

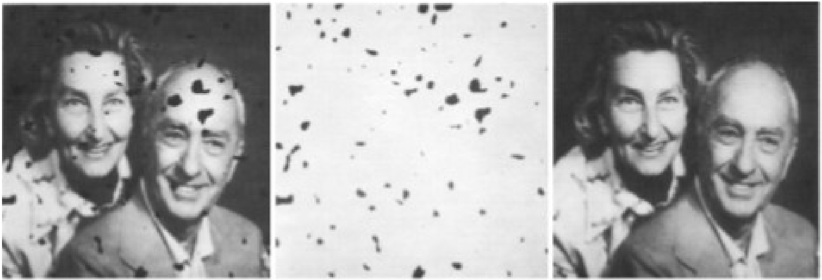
Denoising

$$\mathcal{T}(f)(x) = f(x) + \delta g(x)$$



Inpainting

$$\mathcal{T}(f) = P_K(f)$$

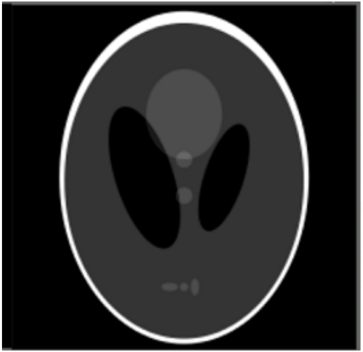


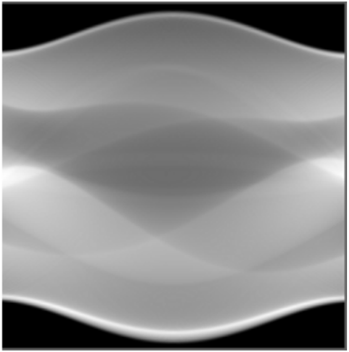
Deconvolution

$$\mathcal{T}(f)(x) = (h \circledast f)(x)$$

Computarized Tomography (CT)

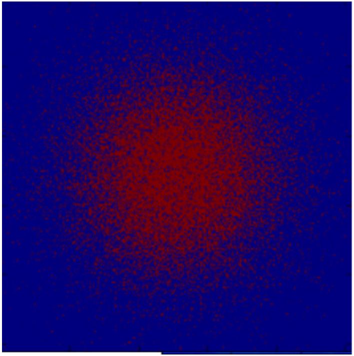
$$\mathcal{T}(f)(\theta, s) = \int_{-\infty}^{\infty} f(x_1, x_2)\delta(x_1 \cos(\theta) + x_2 \sin(\theta) - s)dx_1 dx_2$$





Magnetic Resonance Imaging (MRI)

$$\mathcal{T}(f) = (\text{k-space sampling})f$$



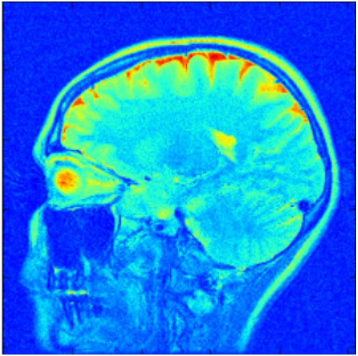


Image denoising

Goal

Recover an image $f \in X = \mathcal{E}^2(\mathbb{R}^2)$ from noisy data:

$$g = f + \delta g$$

where δg is Gaussian white noise, with sd. σ.

Image denoising

Goal

Recover an image $f \in X = \mathcal{E}^2(\mathbb{R}^2)$ from noisy data:

$$g = f + \delta g$$

where δg is Gaussian white noise, with sd. σ.

- The risk of an estimator $\tilde{f}$ is given by the MSE:

  $$\mathbb{E}||f - \tilde{f}||_2^2$$

  $\mathbb{E}$ is the expectation with respect of the probability distribution of δg.

Image denoising

Goal

Recover an image $f \in X = \mathcal{E}^2(\mathbb{R}^2)$ from noisy data:

$$g = f + \delta g$$

where δg is Gaussian white noise, with sd. σ.

- The risk of an estimator $\tilde{f}$ is given by the MSE:

  $$\mathbb{E}||f - \tilde{f}||_2^2$$

  $\mathbb{E}$ is the expectation with respect of the probability distribution of δg.

- The worst behaviour of the estimator is the supremum

  $$\sup_{f \in X} \mathbb{E}||f - \tilde{f}||_2^2$$

  the *Minimax* MSE will be

  $$\inf_{\tilde{f}} \sup_{f \in X} \mathbb{E}||f - \tilde{f}||_2^2$$

Minimax MSE

Frame

A frame for $\mathcal{E}^2(\mathbb{R}^2)$ is a collection $\Psi = \{\psi_i\}_{i\in\mathcal{I}} \subset \mathcal{E}^2(\mathbb{R}^2)$ satisfying

$$A||f||_2^2 \leq ||\{\langle f, \psi_i\rangle\}_{i\in\mathcal{I}}||^2_{\ell^2(\mathcal{I})} \leq B||f||_2^2 \quad \forall \in \mathcal{E}^2(\mathbb{R}^2)$$

for some $0 < A \leq B < \infty$.

Minimax MSE

Frame

A frame for $\mathcal{E}^2(\mathbb{R}^2)$ is a collection $\Psi = \{\psi_i\}_{i\in\mathcal{I}} \subset \mathcal{E}^2(\mathbb{R}^2)$ satisfying

$$A||f||_2^2 \leq ||\{\langle f, \psi_i\rangle\}_{i\in\mathcal{I}}||_{\ell^2(\mathcal{I})}^2 \leq B||f||_2^2 \quad \forall \in \mathcal{E}^2(\mathbb{R}^2)$$

for some $0 < A \leq B < \infty$.

- It is proven (Labate et al.,2012): If an image is sparse within a frame $\{\psi_i\}_{i\in\mathcal{I}}$, one can obtain a *Minimax* MSE estimator by thresholding the coefficients in the expansion of the noisy data:

$$g = \sum_{i\in\mathcal{I}} \langle g, \psi_i\rangle\psi_i$$

Image inpainting

Goal

Recover an image $f \in \mathcal{E}^2(\mathbb{R}^2)$ from known data:

$$g = P_K(f)$$

where P_K is and orthogonal projection onto the known subspace $(\mathcal{E}^2(\mathbb{R}^2))_K$.

Goal

Recover an image $f \in \mathcal{E}^2(\mathbb{R}^2)$ from known data:

$$g = P_K(f)$$

where P_K is and orthogonal projection onto the known subspace $(\mathcal{E}^2(\mathbb{R}^2))_K$.

- **Compressed sensing result:** If a signal (image) is sparse within a frame Ψ, it can be recovered from highly underdetermined, non-adaptive linear measurements by ℓ^1-regularization (Davenport et al., 2012), i.e.

$$\min_{\tilde{f} \in \mathcal{E}^2(\mathbb{R}^2)} ||\{\langle \tilde{f}, \psi_i \rangle\}_{i \in \mathcal{I}}||_{\ell^1(\mathcal{I})} \quad \text{s.t. } P_K(\tilde{f}) = g = P_K(f)$$

Error estimate

Let $\delta > 0$ and $\Lambda \subset \mathcal{I}$ be a δ-**cluster** for f with respect to Ψ.

If $\mu_c(\Lambda, P_M\Psi) < 1/2$ and f^* is the minimizer of the problem, then

$$||\{\langle f^* - f, \psi_i\rangle\}_{i\in\mathcal{I}}||_{\ell^1(\mathcal{I})} \leq \frac{2\delta}{1-\mu_c(\Lambda, P_M\Psi)}$$

where P_M is the projection onto the missing subspace and $\mu_c(\Lambda, P_M\Psi)$ the **cluster coherence**, defined by

$$\mu_c(\Lambda, P_M\Psi) := \max_{j\in\mathcal{I}} \sum_{i\in\Lambda} |\langle P_M\psi_i, P_M\psi_j\rangle|$$

Error estimate

Let $\delta > 0$ and $\Lambda \subset \mathcal{I}$ be a δ-**cluster** for f with respect to Ψ.

If $\mu_c(\Lambda, P_M\Psi) < 1/2$ and f^* is the minimizer of the problem, then

$$||\{\langle f^* - f, \psi_i\rangle\}_{i\in\mathcal{I}}||_{\ell^1(\mathcal{I})} \leq \frac{2\delta}{1 - \mu_c(\Lambda, P_M\Psi)}$$

where P_M is the projection onto the missing subspace and $\mu_c(\Lambda, P_M\Psi)$ the **cluster coherence**, defined by

$$\mu_c(\Lambda, P_M\Psi) := \max_{j\in\mathcal{I}} \sum_{i\in\Lambda} |\langle P_M\psi_i, P_M\psi_j\rangle|$$

- **Conclusion:** A sparsifying frame for images allows you to perform image denoising and inpainting, the reconstruction quality depends on the sparsifying level. **Problem:** Pick a good frame.

Cartoon-like functions

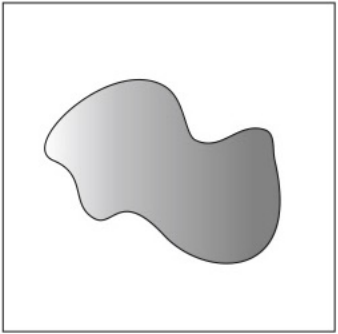
Definition

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{C}$, $f \in \mathcal{E}^2(\mathbb{R}^2)$ if $f = f_0 + \chi_B f_1$, with $B \subset [0,1]^2$, $\partial B \in C^2$ and with bounded curvature. Moreover, $f_i \in C^2(\mathbb{R}^2)$ with $||f_i||_{C^2} \leq 1$ and $\text{supp} f_i \subset [0,1]^2$ for $i = 0, 1$.

Cartoon-like functions

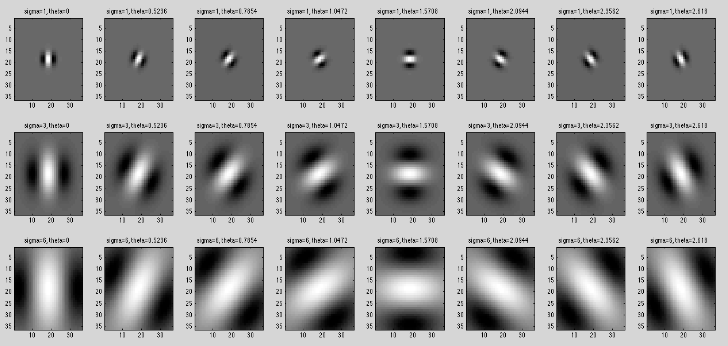
Definition

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{C}$, $f \in \mathcal{E}^2(\mathbb{R}^2)$ if $f = f_0 + \chi_B f_1$, with $B \subset [0,1]^2$, $\partial B \in C^2$ and with bounded curvature. Moreover, $f_i \in C^2(\mathbb{R}^2)$ with $||f_i||_{C^2} \leq 1$ and $\text{supp} f_i \subset [0,1]^2$ for $i = 0, 1$.



Examples of frames for images

- Gabor frames (Gabor, 1946).
- Wavelet frames (Morlet et al., 1984).
- Curvelet frames (Candès et al., 1999).
- Shearlet frames (Kutyniok et al., 2005).



Optimal approximation error for images

Best N-term approx. error (Donoho, 2001)

Let $\{\psi_\lambda\}_{\lambda\in\Lambda} \subset L^2(\mathbb{R}^2)$ a frame. The optimal best N-Term approximation error for any $f \in \mathcal{E}^2(\mathbb{R}^2)$ is

$$\sigma_N(f, \{\psi_\lambda\}_{\lambda\in\Lambda}) = O(N^{-1})$$

Optimal approximation error for images

Best N-term approx. error (Donoho, 2001)

Let $\{\psi_\lambda\}_{\lambda\in\Lambda} \subset L^2(\mathbb{R}^2)$ a frame. The optimal best N-Term approximation error for any $f \in \mathcal{E}^2(\mathbb{R}^2)$ is

$$\sigma_N(f, \{\psi_\lambda\}_{\lambda\in\Lambda}) = O(N^{-1})$$

Error of 2D-wavelets

$$\sigma_N(f, \{\psi_{j,m}\}_{j,m}) \sim N^{-1/2}$$

Optimal approximation error for images

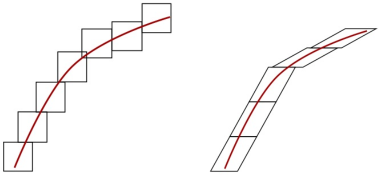
Best N-term approx. error (Donoho, 2001)

Let $\{\psi_\lambda\}_{\lambda\in\Lambda} \subset L^2(\mathbb{R}^2)$ a frame. The optimal best N-Term approximation error for any $f \in \mathcal{E}^2(\mathbb{R}^2)$ is

$$\sigma_N(f, \{\psi_\lambda\}_{\lambda\in\Lambda}) = O(N^{-1})$$

Error of 2D-wavelets

$$\sigma_N(f, \{\psi_{j,m}\}_{j,m}) \sim N^{-1/2}$$



$$A_j := \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}$$

$$S_k := \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

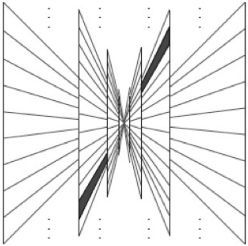
Shearlet Transform (Kutyniok, Guo, Labate, 2005)

Classical Shearlet Transform

$$\langle f, \psi_{j,k,m} \rangle = \int_{\mathbb{R}^2} f(x)\overline{\psi_{j,k,m}(x)}dx$$

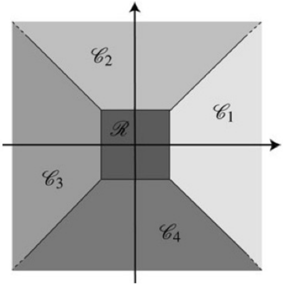
where

$$\mathcal{SH}(\psi) = \{\psi_{j,k,m}(x) = 2^{3j/4}\psi(S_k A_j x - m) : (j,k) \in \mathbb{Z}^2, m \in \mathbb{Z}^2\}$$

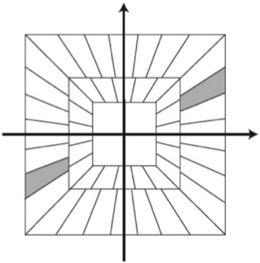


Cone-adapted shearlet transform

$$\mathcal{SH}(\phi, \psi, \tilde{\psi}, c) := \mathcal{P}_{\mathcal{R}}\Phi(\phi, c1) \cup \mathcal{P}_{\mathcal{C}_1}\Psi(\psi, c) \cup \mathcal{P}_{\mathcal{C}_2}\tilde{\Psi}(\tilde{\psi}, c)$$



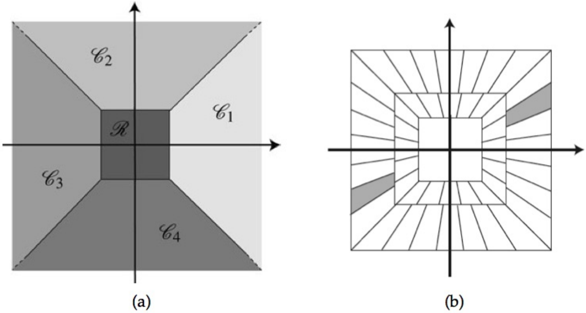
(a)



(b)

Cone-adapted shearlet transform

$$\mathcal{SH}(\phi,\psi,\tilde{\psi},c) := \mathcal{P}_{\mathcal{R}}\Phi(\phi,c1) \cup \mathcal{P}_{\mathcal{C}_1}\Psi(\psi,c) \cup \mathcal{P}_{\mathcal{C}_2}\tilde{\Psi}(\tilde{\psi},c)$$



- Best N-term approximation error

$$\sigma_N(f,\{\psi_{j,k,m}\}_{j,k,m}) \sim N^{-1}(\log(N))^{3/2}$$

Current software

- Matlab
  - FFST- Fast Finite Shearlet Transform (Häuser, Steidl,TU Keiserlautern)
    http://www.mathematik.uni-kl.de/imagepro/software/ffst/
  - 2D/3D Shearlet Toolbox (D. Labate, University of Houston)
    https://www.math.uh.edu/~dlabate/software.html
  - **Shearlab3D** (G. Kutyniok, W.-Q.Lim, R. Reisenhoffer, TU Berlin)
    http://www.shearlab.org/
- Python
  - pyShearLab (Stefan Loock, U Götingen)
    http://na.math.uni-goettingen.de/pyshearlab/
- **Julia**
  - **Shearlab.jl** (H. Andrade, TU Berlin)
    https://github.com/arsenal9971/Shearlab.jl

Why Julia?

- Extensive use of `fft`, well implemented in Julia.
- Fast vectorization and loops as well as JIT-compilation.
- Plenty of image filtering, import and rescaling functions with `Images.jl`, `Wavelets.jl`.
- Support of multithreading and painless GPU processing with `ArrayFire.jl`.

Thanks!

Questions?