SPOQ: A Novel Smoothed Norm Ratio for Sparse Signal Restoration.

Afef Cherni[1], Emilie Chouzenoux[2], Laurent Duval[3], and Jean-Christophe Pesquet[2]

[1] Aix-Marseille Univ., CNRS, Centrale Marseille, I2M, Marseille, France.
[2]CVN, CentraleSupélec, INRIA Saclay and Univ. Paris Saclay.
[3] IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison.

Overview

- **Observation model:** Restore the unknown sparse signal $\mathbf{x} \in \mathbb{R}^N$ from the observations $\mathbf{y} \in \mathbb{R}^M$ related to $\mathbf{x}$ through:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

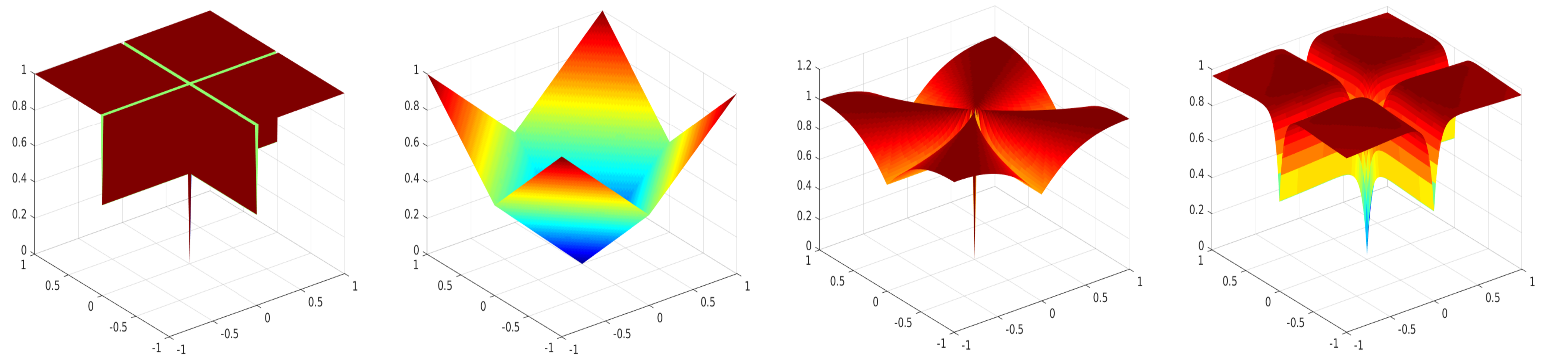
- **Variational approach:**

$$\underset{\mathbf{x}\in S}{\text{minimize}} \;\; \Psi(\mathbf{x}) \quad \text{with} \quad S = \{\mathbf{x} \in \mathbb{R}^N \mid \|\mathbf{H}\mathbf{x} - \mathbf{y}\| \leq \xi\}. \quad (2)$$

▷ $\Psi \colon \mathbb{R}^N \to]-\infty, +\infty]$: regularization function used to enforce sparsity on the solution.
▷ $\xi > 0$: parameter depending on the noise characteristics.

✗**Difficulties:** Choice of Ψ?

Motivation



ℓ_0 **pseudo-norm** | ℓ_1 **norm** | **SOOT** ℓ_1/ℓ_2 **[5]** | **SPOQ** $\ell_{1/2}/\ell_3$ **This work and [3]**

Proposed approach

Smoothed p-Over-q (SPOQ) penalty

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \Psi(\mathbf{x}) = \log\left(\frac{(\ell_{p,\alpha}^p(\mathbf{x}) + \beta^p)^{1/p}}{\ell_{q,\eta}(\mathbf{x})}\right), \quad (3)$$

where $p \in]0,2]$ and $q \in]2,+\infty[$ and:

$$\begin{cases} \ell_{p,\alpha}(\mathbf{x}) = \left(\sum_{n=1}^N \left((x_n^2 + \alpha^2)^{p/2} - \alpha^p\right)\right)^{1/p} \\ \ell_{q,\eta}(\mathbf{x}) = \left(\eta^q + \sum_{n=1}^N |x_n|^q\right)^{1/q} \end{cases}$$

for $(\alpha, \beta, \eta) \in]0,+\infty[^3$.

✓ $\ell_{p,\alpha}$ and $\ell_{q,\eta}$ are the smoothed version of ℓ_p and ℓ_q (quasi-)norms.
✓ Ψ (3) is a generalized version of the smoothed ℓ_1/ℓ_2 function [5].

Properties

✗ Problem (2) is non-convex.
✓ Ψ presents two properties:
- β Lipschitz-differentiable on $\mathbb{R}^N$
- Locally majorized by a quadratic function

By construction, for every $(\mathbf{x}, \mathbf{x}') \in \mathcal{B}_{q,\rho}^2$

$$\Psi(\mathbf{x}) \leq \Psi(\mathbf{x}') + \nabla\Psi(\mathbf{x}')^\top(\mathbf{x}-\mathbf{x}') + \frac{1}{2}(\mathbf{x}-\mathbf{x}')^\top \boldsymbol{A}_{q,\rho}(\mathbf{x}')(\mathbf{x}-\mathbf{x}')$$

where

$$\mathcal{B}_{q,\rho} = \{\mathbf{x} \in \mathbb{R}^N \mid \sum_{n=1}^N |x_n|^q \geq \rho^q\}$$

and

$$\boldsymbol{A}_{q,\rho}(\mathbf{x}) = \frac{1}{\ell_{p,\alpha}^p(\mathbf{x}) + \beta^p}\,\mathrm{Diag}\left((x_n^2 + \alpha^2)^{p/2-1}\right)_{1\leq n\leq N} + \frac{q-1}{(\eta^q + \rho^q)^{2/q}}\mathrm{I}_N.$$

Optimization Tools

- Variable-Metric Forward-Backward Algorithm [4, 1]
- Metric based on local Majoration-Minimization strategy.

☞ New Trust Region VMFB Algorithm

$\mathbf{x}_0 \in \mathbb{R}^N,\ B \in \mathbb{N}^*,\ \theta \in]0,1[,\ (\gamma_k)_{k\in\mathbb{N}} \in]0,2[$
For $k = 0, 1, \ldots$:
  For $i = 1, \ldots, B$:
    If $i = 1,\ \rho_{k,1} = \sum_{n=1}^N |x_{n,k}|^q$.
    If $i \in \{2, \ldots, B-1\},\ \rho_{k,i} = \theta\rho_{k,i-1}$.
    Else $\rho_{k,B} = 0$.
    Construct $\boldsymbol{A}_{k,i} = \boldsymbol{A}_{q,\rho_{k,i}}(\mathbf{x}_k)$
    $\mathbf{z}_{k,i} = \mathsf{P}_{\boldsymbol{A}_{k,i},S}\left(\mathbf{x}_k - \gamma_k(\boldsymbol{A}_{k,i})^{-1}\nabla\Psi(\mathbf{x}_k)\right)$
    If $\mathbf{z}_{k,i} \in \mathcal{B}_{q,\rho_{k,i}}$: Stop loop
  $\mathbf{x}_{k+1} = \mathbf{z}_{k,i}$

Application

Mass Spectrometry (MS)

Physico-chemical analysis:
- A fondamental technology of analytical chemistry.
- Used in structural biology, chemistry, pharmaceutical analysis, etc.

State of the art:
- Problem (2) in MS context was solved in [2] using a dictionary based-strategy.
- The dictionary based-strategy aimed at solving problem (2) using ℓ_1 norm.

✗ scale biases.

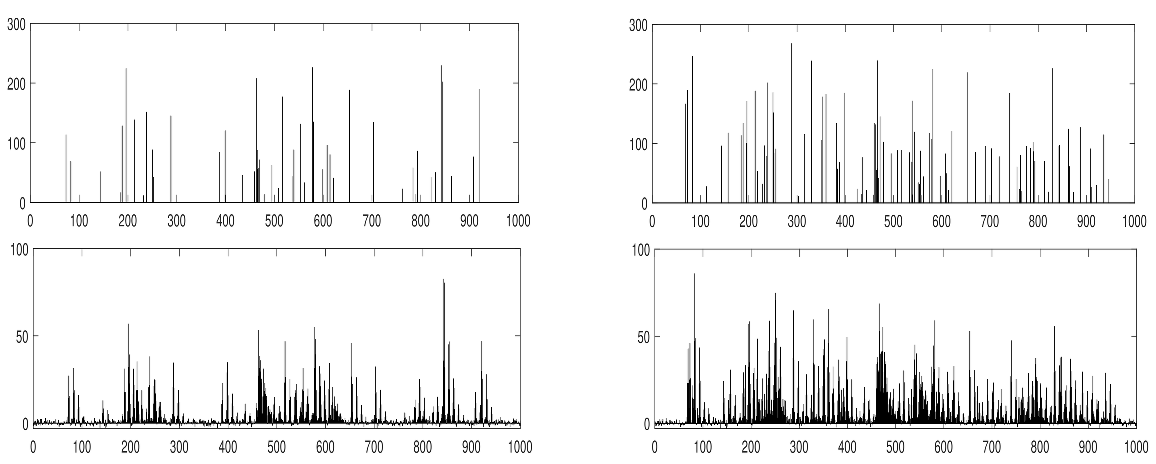
☞ Solve problem (2) with the dictionary based-strategy using proposed ℓ_p/ℓ_q penalty (with different p and q values) and various choices of Ψ (ℓ_1 norm, Cauchy and Welsh penalties).

Evaluation criteria

- SNR$= 20\log_{10}(\|\mathbf{x}\|/\|\mathbf{x} - \hat{\mathbf{x}}\|)$.
- TSNR: SNR computed only on the support of the sought sparse signal.
- $\hat{P}$: Estimated sparsity degree.

Simulated data

- Signal A: $N = 1000$, $P = 48$ (sparsity degree)
- Signal B: $N = 1000$, $P = 94$ (sparsity degree)



Original sparse signals and associated MS spectra of dataset A (left) and dataset B (right) (top: synthetic signal, bottom: noisy MS spectra)

- The associated MS spectra $\mathbf{y}$ are built with $\mathbf{H}$ being the dictionary-based method [2].
- A zero-mean Gaussian noise with standard deviation 10^{-2} is considered.

Results

		P	$\ell_{0.25}/\ell_2$	$\ell_{0.5}/\ell_2$	ℓ_1/ℓ_2	ℓ_1	Cauchy	Welsh
Signal A	SNR	48	**46.28**	41.91	40.91	43.16	42.84	27.54
			0.497	*0.436*	*0.910*	*0.654*	*0.572*	*0.461*
	TSNR	48	46.55	**47.71**	46.24	43.94	43.53	29.12
			0.571	*1.136*	*1.660*	*0.679*	*0.532*	*0.501*
	$\hat{P}$	48	**49**	129	365	80	883	259
			1.32	*11.85*	*10.13*	*9.46*	*10.57*	*8.08*
Signal B	SNR	94	**45.56**	42.74	41.31	43.02	42.71	30.99
			0.538	*1.266*	*1.298*	*1.260*	*1.194*	*0.488*
	TSNR	94	**47.26**	46.88	45.11	44.17	43.68	33.39
			0.639	*1.495*	*1.654*	*1.138*	*0.961*	*0.507*
	$\hat{P}$	94	**111**	216	410	165	952	342
			3.54	*12.43*	*11.03*	*17.41*	*6.66*	*11.72*

Means/stds of SNR, truncated SNR and sparsity level of the restored signals, computed on 10 noise realizations.

Conclusion & Perspectives

✓ ℓ_p/ℓ_q function ensures a good estimation for both dataset A and dataset B [3].
✓ For special values of (p, q), the ℓ_p/ℓ_q function appears better than others penalties.
✓ The special case of $p = 0.25$ and $q = 2$ seems to be an optimal choice in terms of estimation accuracy.
☞ Does any relationship between p and q can be established to make ℓ_p/ℓ_q the best sparsity promoter?

References

[1] F. Abboud, E. Chouzenoux, J.-C. Pesquet, J.-H. Chenot, and L. Laborelli. Dual block-coordinate forward-backward algorithm with application to deconvolution and deinterlacing of video sequences. *Journal of Mathematical Imaging and Vision*, 59(3):415–431, 2017.

[2] A. Cherni, E. Chouzenoux, and M.-A. Delsuc. Fast dictionary-based approach for mass spectrometry data analysis. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018)*, pages 1–5, 2018.

[3] A. Cherni, E. Chouzenoux, L. Duval, and J.-C. Pesquet. SPOQ: a smoothed scale-invariant ℓ_p-over-ℓ_q norm ratio penalty regularization for sparse signal recovery. *PREPRINT*, 2019.

[4] Emilie Chouzenoux, Jean-Christophe Pesquet, and Audrey Repetti. Variable metric forward–backward algorithm for minimizing the sum of a differentiable function and a convex function. *Journal of Optimization Theory and Applications*, 162(1):107–132, 2014.

[5] A. Repetti, M. Q. Pham, L. Duval, E. Chouzenoux, and J.-C. Pesquet. Euclid in a taxicab: Sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization. *j-ieee-spl*, 22(5):539–543, May 2015.

A*Midex Initiative d'excellence Aix-Marseille | Institut de Mathématiques de Marseille | Inria inventeurs du monde numérique | CentraleSupélec | IFP Energies nouvelles

SPARS 2019, Toulouse 1-4 July.