Luca Calatroni (calatroni@i3s.unice.fr)
Emmanuel Soubies (emmanuel.soubies@irit.fr)
October 13, 2022

Internship proposal

Exact Relaxations for Sparse Optimization.

Context

Sparse models are widely used in machine learning, statistics, and signal/image processing applications (e.g., coding, inverse problems, variable selection, or image decomposition). A typical formulation is given by

$$\hat{\mathbf{x}} \in \left\{ \arg\min_{\mathbf{x}\in\mathbb{R}^N} \mathcal{G}_0(\mathbf{x}) := \mathcal{D}(\mathbf{Hx};\mathbf{y}) + \lambda\|\mathbf{x}\|_0 \right\}, \tag{1}$$

where $\mathcal{D} : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$ measures the discrepancy between the model $\mathbf{Hx}$ and the data $\mathbf{y}$ (data-fidelity), $\|\cdot\|_0$ denotes the ℓ_0 pseudo-norm that counts the non-zero entries of its argument (sparsity prior), and $\lambda > 0$ controls the trade-off between data-fidelity and sparsity. However, this problem is non-convex and belongs the NP-complete class of complexity that makes its resolution a challenging task. Hence, the standard practice aims at replacing the ℓ_0 term in (1) by a continuous sparsity-promoting penalty $\Phi : \mathbb{R}^N \to \mathbb{R}$,

$$\hat{\mathbf{x}} \in \left\{ \arg\min_{\mathbf{x}\in\mathbb{R}^N} \tilde{\mathcal{G}}(\mathbf{x}) := \mathcal{D}(\mathbf{Hx};\mathbf{y}) + \lambda\Phi(\mathbf{x}) \right\}. \tag{2}$$

A popular penalty function Φ is the ℓ_1 norm that can be seen as the best compromise between sparsity and convexity [1, 2]. However, the bottleneck of this convex relaxation is that it introduces a bias on large coefficients [3].

This drawback has led to a growing interest in continuous, yet non-convex, relaxations $\tilde{\mathcal{G}}$ of the initial functional $\mathcal{G}_0$. The vast majority of existing relaxations have been proposed for the least-squares penalized problem, where $\mathcal{D}(\cdot,\mathbf{y}) = \frac{1}{2}\|\cdot-\mathbf{y}\|_2^2$. In particular, there exists a class of penalties Φ that lead to exact continuous relaxations of $\mathcal{G}_0$ [4, 5, 6] in the sense that

(P1) The global minimizers of $\mathcal{G}_0$ and those of the relaxation $\tilde{\mathcal{G}}$ coincide,
(P2) Each local minimizer of $\tilde{\mathcal{G}}$ is a local minimizer of $\mathcal{G}_0$.

In other words, such relaxations (termed as *exact continuous relaxations*) allow us to "reduce" the non-convexity of Problem (1) (e.g., less local minimizers, wider basins of attraction) while preserving its solution(s). These are appealing properties in the context of non-convex optimization.

Objectives

Existing works on exact continuous relaxations [4, 5, 6, 7] were only dedicated to the case of a quadratic data-fidelity terms $\mathcal{D}$. In a Bayesian framework, this corresponds to the negative log-likelihood describing the presence of additive white Gaussian noise. However, for many applications, the quadratic distance is not the choice that best complies with the nature of the data. For instance, the noise in the observed data $\mathbf{y}$ is generally not purely Gaussian but rather has a mixed [8] or signal-dependent nature [9, 10]. Alternative measures of fit include the generalized Kullback-Leibler divergence (a special instance of β-divergences [11] with $\beta = 1$), the logistic regression loss [12], or the Huber loss [13].

The main objective of this internship will be to generalize the theory developed in [4, 5] to other data-fidelity terms $\mathcal{D}$. In particular, one aspect of the work will be to define and analyze generic and tractable transformations which – when facing a problem of the form (1) – allow us to construct relaxations with properties (P1) and (P2) based, for instance, on generalised conjugation tools [14].

Practical aspects

We are looking for a highly motivated student, willing to continue with a PhD thesis, with a background in mathematics (optimization, probability an statistics, geometry) and/or electrical engineering (signal/image processing). Strong abilities in computer sciences will be appreciated.

The intern will be granted the usual stipend of ~ 600 euros/month. If the candidate is successful, this internship will be pursued by a PhD. Depending on the candidate's interest, this internship can take place either at

- i3S Laboratory in Sophia-Antipolis,
- IRIT laboratory in Toulouse.

It will be co-supervised by Luca Calatroni (CR, i3S) and Emmanuel Soubies (CR, IRIT).

Do not hesitate to contact us for more information.

Bibliography

[1] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, February 2006.

[2] David L. Donoho. For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59(6):797–829, June 2006.

[3] Marcus Carlsson, Daniele Gerosa, and Carl Olsson. An unbiased approach to compressed sensing. *arXiv:1806.05283 [math]*, June 2018. arXiv: 1806.05283.

[4] E. Soubies, L. Blanc-Féraud, and G. Aubert. A Continuous Exact l0 Penalty (CEL0) for Least Squares Regularized Problem. *SIAM Journal on Imaging Sciences*, 8(3):1607–1639, January 2015.

[5] E. Soubies, L. Blanc-Féraud, and G. Aubert. A Unified View of Exact Continuous Penalties for l2-l0 Minimization. *SIAM Journal on Optimization*, 27(3):2034–2060, January 2017.

[6] Marcus Carlsson. On Convex Envelopes and Regularization of Non-Convex Functionals without moving Global Minima. *arXiv:1811.03439 [math]*, November 2018. arXiv: 1811.03439.

[7] M. Lazzaretti, L. Calatroni, and C. Estatico. Weighted-cel0 sparse regularisation for molecule localisation in super-resolution microscopy with poisson data. *IEEE 2021 Symposium on Biomedical Imaging*, 2021.

[8] L. Calatroni, J. C. De Los Reyes, and C.-B. Schönlieb. Infimal convolution of data discrepancies for mixed noise removal. *SIAM J. Imaging Sci.*, 10(3), 2017.

[9] A. Jezierska, C. Chaux, J. Pesquet, H. Talbot, and G. Engler. An em approach for time-variant poisson-gaussian model parameter estimation. *IEEE Trans. on Sign. Process.*, 62(1), 2014.

[10] Mario Bertero, Patrizia Boccacci, and Valeria Ruggiero. *Inverse Imaging with Poisson Data*. 2053-2563. IOP Publishing, 2018.

[11] Andrzej Cichocki and Shun-ichi Amari. Families of alpha-beta-and gamma-divergences: Flexible and robust measures of similarities. *Entropy*, 12(6), 2010.

[12] Peter Bühlmann and Sara Van De Geer. *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media, 2011.

[13] Peter J Huber. *Robust statistics*, volume 523. 2004.

[14] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. 2009.