Metodi avanzati di ottimizzazione non lineare per l'elaborazione di immagini

G. Landi
Università di Bologna
germana.landi@unibo.it

Convegno e Assemblea GNCS 2020
12 febbraio 2020

Progetto INDAM-GNCS 2019

Titolo: **Metodi avanzati di ottimizzazione non lineare per l'elaborazione di immagini**

Coordinatore: Germana Landi

Membri:

- Silvia Bonettini, Simone Rebegoldi (UNIMORE)
- Valentina De Simone, Daniela di Serafino, Marco Viola (UNICAMPANIA)
- Martin Huska, Alessandro Lanza, Damiana Lazzaro, Serena Morigi, Monica Pragliola, Fiorella Sgallari, Fabiana Zama (UNIBO)

Il presente progetto si pone l'obiettivo di analizzare e sviluppare tecniche numeriche avanzate per specifici problemi di elaborazione di immagini. . . .
In particolare saranno oggetto della ricerca proposta metodi proximal gradient, metodi del gradiente, tecniche di majorization-minimization, metodi del gradiente coniugato non lineare e di tipo ϵ-subgradiente. I suddetti metodi saranno sperimentati in applicazioni nell'ambito della Risonanza Magnetica Nucleare, della Risonanza Magnetica Funzionale, della tomosintesi multienergetica e multimateriale e della termografia

Attività

Giornata di lavoro

- 31 Gennaio 2020, Firenze, progetti GNCS: “Tecniche adattive per metodi di ottimizzazione in Machine Learning” + “Metodi avanzati di ottimizzazione non lineare per l’elaborazione di immagini”

Programma Professori Visitatori

- Visista del Prof. Jacek Gondzio (School of Mathematics ,The University of Edinburgh, Scotland, UK) per due settimane presso l’Università della Campania “L. Vanvitelli”.

Partecipazione a conferenze

- European Conference on Numerical Mathematics and Advanced Applications (ENUMATH), The Netherlands, September 30 October 4, 2019
- Applied Inverse Problems Conference, France, July, 08th - 12th, 2019
- 9th International Conference on New Computational Methods for Inverse Problems (NCMIP 2019), France, May 24th, 2019.
- 9th International Congress on Industrial and Applied Mathematics (ICIAM 2019), Spain, July 15th-19th 2019.
- 2nd IMA and OR Society conference on Mathematics of Operational Research, UK, April 24th-27th 2019.

Pubblicazioni

1. Lazzaro D., Loli Piccolomini E., Zama F. (2019), A fast splitting method for efficient Split Bregman iterations, Applied Mathematics and Computation, 357, p. 139-146.
2. Lazzaro D., Piccolomini, E Loli, Zama, F. (2019), A nonconvex penalization algorithm with automatic choice of the regularization parameter in sparse imaging, Inverse Problems, 35 (8), p. 084002.
3. Keinert, F., Lazzaro, D., Morigi, S. (2019), A Robust Group-Sparse Representation Variational Method with Applications to Face Recognition, IEEE Transactions on Image Processing, 28(6), pp. 2785-2798 .
4. Chan R. H., Lazzaro D., Morigi S., Sgallari F. (2019), A Non-convex Nonseparable Approach to Single-Molecule Localization Microscopy. In: Scale Space and Variational Methods in Computer Vision. Lecture Notes in Computer Science, 11603, p. 498-509.
5. Rebegoldi S., Bonettini S. and Prato M., Efficient block coordinate methods for blind Cauchy denoising, Lecture Notes in Computer Science, in stampa, (*).
6. Bonettini S., Porta F., Prato M., Rebegoldi S., Ruggiero V. and Zanni L. (2019), Recent advances in variable metric first-order methods. In: Computational Methods for Inverse Problems in Imaging, M. Donatelli and S. Serra Capizzano (eds.), Springer INdAM Series 36, 1-31.
7. Corsaro S., De Simone V. (2019), Adaptive l1-regularization for short-selling control in portfolio selection, Computational Optimization and Applications, 72(2), pp. 457-478.
8. Corsaro S., De Simone V., Marino, Z. (2019), Fused Lasso approach in portfolio selection, Annals of Operations Research.
9. Corsaro S., De Simone V., Marino Z., Perla, F. (2019), l1-Regularization for multi-period portfolio selection, Annals of Operations Research.
10. Landi, G., Piccolomini, E.L., Nagy, J. (2019), Nonlinear conjugate gradient method for spectral tomosynthesis, Inverse Problems 35(9), 094003.
11. Bortolotti, V., Brizi, L., Fantazzini, P., Landi, G., Zama, F. (2019), Upen2DTool: A Uniform PENalty Matlab tool for inversion of 2D NMR relaxation data, SoftwareX, 10, 100302.
12. di Serafino D., Landi G., Viola M., ACQUIRE: an inexact iteratively reweighted norm approach for TV-based Poisson image restoration, Applied Mathematics and Computation, 364, 2020
13. Lanza A., Morigi S., Selesnick I.W., Sgallari F. (2019), Sparsity-Inducing Non- convex Nonseparable Regularization for Convex Image Processing, SIAM Journal on Imaging Sciences, 12(2), pp. 1099-1134.
14. Calatroni L., Lanza A., Pragliola M., Sgallari F. (2019), A exible Space-variant Anisotropic Regularization for Image Restoration with Automated Parameter Se- lection, SIAM Journal on Imaging Sciences, 12(2), pp. 1001-1037.
15. Huska M., Lanza A. , Morigi S. , Sgallari F. (2019), Convex non-convex seg- mentation of scalar fields over arbitrary triangulated surfaces, Journal of Com- putational and applied Mathematics, 349, pp. 438-451.
16. Lanza A., Morigi S., Pragliola M., Sgallari F. (2019), Space-variant generalised Gaussian regularisation for image restoration, Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization, 7, pp. 490-503.
17. Calatroni L., Lanza A., Pragliola M., Sgallari F. (2019), Space-Adaptive Anisotropic Bivariate Laplacian Regularization for Image Restoration, Lecture Notes in Computational Vision and Biomechanics, 34, pp. 67-76.
18. A. Abdullahi Hassan, V. Cardellini, P. D'Ambra, D. di Serafino, S. Filippone, Efficient Algebraic Multigrid Preconditioners on Clusters of GPUs, Parallel Processing Letters, 29 (1), 1950001, 2019.
19. V. De Simone, D. di Serafino, M. Viola, A subspace-accelerated split Bregman method for sparse data recovery with joint l1-type regularizers, submitted.
20. D. di Serafino, D. Orban, Constraint-Preconditioned Krylov Solvers for Regularized Saddle-Point Systems, Cahier du GERAD G-2019-72, GERAD, Montral, QC, Canada

Object of the talk

- Image restoration from Poisson data
- An inexact IRN-based minimization algorithm
- Joint work with Daniela di Serafino and Marco Viola (Università della Campania “L. Vanvitelli”)

Mathematical model of image formation

The entries y_j of the measured data $\mathbf{y} \in \mathbb{R}^m$ are samples from independent Poisson random variables Y_j:

$$Y_j \sim \text{Poisson}\left((A\mathbf{x} + \mathbf{b})_j\right)$$

- $A \in \mathbb{R}^{m \times n}$, $a_{ij} \geq 0 \forall i, j$, $\sum_{i=1}^{m} a_{ij} = 1 \forall j$ $\rightarrow$ observation mechanism
- $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{b} > 0$ $\rightarrow$ background radiation
- $\mathbf{x}$ $\rightarrow$ imaged object to be restored

Several applications: fluorescence microscopy, X-ray computed tomography (CT), positron emission tomography (PET), astronomical imaging, ...

Mathematical model of image restoration

In a variational approach, $\mathbf{x}$ is estimated as the solution of the minimization problem

$$\begin{array}{ll} \text{minimize} & D_{KL}(\mathbf{x}) + \lambda TV(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in \mathcal{S} \end{array}$$

- $D_{KL}(\mathbf{x})$ is the Kullback-Leibler divergence of $A\mathbf{x} + \mathbf{b}$ from $\mathbf{y}$:

$$D_{KL}(\mathbf{x}) = \sum_{j=1}^{m} \left(y_j \ln \frac{y_j}{(A\mathbf{x} + \mathbf{b})_j} + (A\mathbf{x} + \mathbf{b})_j - y_j \right)$$

- $TV(\mathbf{x})$ is a discrete version of the Total Variation (TV) functional:

$$TV(\mathbf{x}) = \sum_{i=1}^{n} \|D_i \mathbf{x}\|, \quad D_i = \begin{pmatrix} \mathbf{e}^T_{(l-1)r+k+1} - \mathbf{e}^T_{(l-1)r+k} \\ \mathbf{e}^T_{lr+k} - \mathbf{e}^T_{(l-1)r+k} \end{pmatrix}$$

- $\mathcal{S}$ is a nonempty, closed, convex subset of the non-negative orthant: $\mathcal{S} = \mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$, $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}, \mathbf{e}^T \mathbf{x} = \beta\}$

Smoothed TV

- Smoothed TV version, based on a Huber-like function

$$TV_\mu(\mathbf{x}) = \sum_{i=1}^{n} \phi_\mu \left(\|D_i \mathbf{x}\| \right), \quad \phi_\mu(z) = \begin{cases} z & \text{if } |z| > \mu, \\ \frac{1}{2}\left(\frac{z^2}{\mu} + \mu\right) & \text{otherwise.} \end{cases}$$

- $\mathbf{x}$ is estimated as the solution of the minimization problem

$$\begin{array}{ll} \text{minimize} & D_{KL}(\mathbf{x}) + \lambda TV_\mu(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in \mathcal{S} \end{array}$$

- The problem admits a solution, which is unique if $\mathbf{y} > 0$ and $\mathcal{N}(A) = \{\mathbf{0}\}$

Proposed approach: exploit second-order information

- Inexactly minimize a sequence of quadratic models obtained by
  - quadratic approximations of D_{KL}
  - iteratively reweighted norm (IRN) approximations of TV_μ
- Employ a line search procedure

Quadratic model

- Quadratic approximations of D_{KL} in an iterate $\mathbf{x}^{(k)} \in \mathcal{S}$

$$D_{KL}^{(k)}(\mathbf{x}) = D_{KL}(\mathbf{x}^{(k)}) + (\mathbf{x} - \mathbf{x}^{(k)})^T \nabla D_{KL}(\mathbf{x}^{(k)}) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^{(k)})^T (\nabla^2 D_{KL}(\mathbf{x}^{(k)}) + \gamma I)(\mathbf{x} - \mathbf{x}^{(k)}),$$

γI ensures that $D_{KL}^{(k)}$ is strongly convex; in practice γ is very small.

We can set $\gamma = 0$ if $\mathbf{y} > \mathbf{0}$, $\mathcal{N}(A) = \{\mathbf{0}\}$ and $\mathcal{S}$ is bounded.

Quadratic model (cont'd)

- Quadratic approximations of TV_μ in an iterate $\mathbf{x}^{(k)} \in \mathcal{S}$ obtained using an IRN approach [Rodríguez and Wohlberg '09]:

$$TV_\mu^{(k)}(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{n} w_i^{(k)} \|D_i\mathbf{x}\|^2 + \frac{1}{2} TV_\mu(\mathbf{x}^{(k)}),$$

where

$$w_i^{(k)} = \begin{cases} \|D_i\mathbf{x}^{(k)}\|^{-1} & \text{if } \|D_i\mathbf{x}^{(k)}\| > \mu, \\ \mu^{-1} & \text{otherwise.} \end{cases}$$

- $TV_\mu^{(k)}(\mathbf{x}^{(k)}) = TV_\mu(\mathbf{x}^{(k)}), \quad \nabla TV_\mu^{(k)}(\mathbf{x}^{(k)}) = \nabla TV_\mu(\mathbf{x}^{(k)});$
- $\nabla^2 TV_\mu^{(k)}(\mathbf{x}^{(k)}) = \sum_{i=1}^n w_i^{(k)} D_i^T D_i \approx \nabla^2 TV_\mu(\mathbf{x}^{(k)}) = \sum_{i=1}^n \nabla^2\phi_\mu\left(\|D_i\mathbf{x}^{(k)}\|\right)$

$$\nabla^2\phi_\mu(\|D_i\mathbf{x}\|) = \begin{cases} w_i D_i^T D_i - \dfrac{(D_i^T D_i\mathbf{x})(D_i^T D_i\mathbf{x})^T}{\|D_i\mathbf{x}\|^3} & \text{if } \|D_i\mathbf{x}\| > \mu \\ w_i D_i^T D_i & \text{if } \|D_i\mathbf{x}\| < \mu \end{cases}$$

Proposed method

choose $\mathbf{x}_0 \in \mathcal{S}$, $\eta \in (0,1)$, $\delta \in (0,1)$, $\{\varepsilon_k\}$ s.t. $\varepsilon_k > 0$ and $\lim_{k\to\infty} \varepsilon_k = 0$

for $k = 1, 2, \dots$ **do**

compute an approx $\widehat{\mathbf{x}}^{(k)}$ to the solution $\overline{\mathbf{x}}^{(k)}$ of the quadratic problem,

$$\min_{\mathbf{x}\in\mathcal{S}} F_k(\mathbf{x}) = D_{KL}^{(k)}(\mathbf{x}) + \lambda TV_\mu^{(k)}(\mathbf{x})$$

such that

$$\|\widehat{\mathbf{x}}^{(k)} - \overline{\mathbf{x}}^{(k)}\| \leq \varepsilon_k \quad \text{and} \quad F_k(\widehat{\mathbf{x}}^{(k)}) \leq F_k(\mathbf{x}^{(k)})$$

$\alpha_k := 1$

$\mathbf{d}^{(k)} := \widehat{\mathbf{x}}^{(k)} - \mathbf{x}^{(k)}$

$\mathbf{x}_\alpha^{(k)} := \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$

while $F(\mathbf{x}_\alpha^{(k)}) > F(\mathbf{x}^{(k)}) + \eta\alpha_k \nabla F(\mathbf{x}^{(k)})^T \mathbf{d}^{(k)}$ **do**

$\alpha_k := \delta\alpha_k$

$\mathbf{x}_\alpha^{(k)} := \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$

end while

$\mathbf{x}^{(k+1)} = \mathbf{x}_\alpha^{(k)}$

end for

ACQUIRE

ACQUIRE: Algorithm based on Consecutive QUadratic and Iterative REweighted norm approximations

- well posed (i.e., a steplength α_k satisfying the Armijo condition can be found in a finite number of iterations)
- convergent
- does not require the exact solution of the quadratic problem and condition $\|\widehat{\mathbf{x}}^{(k)} - \overline{\mathbf{x}}^{(k)}\| \leq \varepsilon_k$ can be replaced by one which is simple to verify

Well-posedness

Theorem

Let $\delta \in (0,1)$. There exist $\overline{\alpha} > 0$ independent of k and an integer $j_k \geq 0$ such that for $\alpha_k = \delta^{j_k}$

$$F(\mathbf{x}_\alpha^{(k)}) \leq F(\mathbf{x}^{(k)}) + \eta\alpha_k \nabla F(\mathbf{x}^{(k)})^T(\widehat{\mathbf{x}}^{(k)} - \mathbf{x}^{(k)}),$$

$$\alpha_k \geq \overline{\alpha}.$$

Convergence theory

Theorem

Let $\{\mathbf{x}^{(k)}\}$ be the sequence generated by the algorithm. Then there exists a subsequence $\{\mathbf{x}^{(k_j)}\}$ such that

$$\lim_{k_j \to \infty} \mathbf{x}^{(k_j)} = \overline{\mathbf{x}},$$

where $\overline{\mathbf{x}} \in \mathcal{S}$ is such that $\nabla_{\mathcal{S}} F(\overline{\mathbf{x}}) = 0$.

$\nabla_{\mathcal{S}} f(\mathbf{x}) = \arg\min \{\|\mathbf{v} + \nabla f(\mathbf{x})\| \;\; \text{s.t. } \mathbf{v} \in T_{\mathcal{S}}(\mathbf{x})\}$ projected gradient

$T_{\mathcal{S}}(\mathbf{x})$ tangent cone at $\mathbf{x}$

Convergence theory

Theorem

Let $\{\mathbf{x}^{(k)}\}$ be the sequence generated by the algorithm. Then there exists a subsequence $\{\mathbf{x}^{(k_j)}\}$ such that

$$\lim_{k_j \to \infty} \mathbf{x}^{(k_j)} = \bar{\mathbf{x}},$$

where $\bar{\mathbf{x}} \in \mathcal{S}$ is such that $\nabla_{\mathcal{S}} F(\bar{\mathbf{x}}) = 0$.

$\nabla_{\mathcal{S}} f(\mathbf{x}) = \arg\min \{\|\mathbf{v} + \nabla f(\mathbf{x})\| \ \text{ s.t. } \mathbf{v} \in T_{\mathcal{S}}(\mathbf{x})\}$ projected gradient

$T_{\mathcal{S}}(\mathbf{x})$ tangent cone at $\mathbf{x}$

Theorem

Assume that the function F is strictly convex. Then the sequence $\{\mathbf{x}^{(k)}\}$ generated by AQUIRE converges to a point $\bar{\mathbf{x}} \in \mathcal{S}$ such that $\nabla_{\mathcal{S}} F(\bar{\mathbf{x}}) = 0$.

Stopping criterion for inner iterations

Theorem

Assume that

$$\|\nabla_{\mathcal{S}} F_k(\widehat{\mathbf{x}}^{(k)})\| \leq \theta^k \|\nabla_{\mathcal{S}} F_k(\mathbf{x}^{(0)})\|$$

for some $\theta \in (0,1)$. *Then, there exists* $\{\varepsilon_k\}$, *with*

$$\varepsilon_k > 0, \quad \lim_{k\to\infty} \varepsilon_k = 0,$$

such that

$$\|\widehat{\mathbf{x}}^{(k)} - \overline{\mathbf{x}}^{(k)}\| \leq \varepsilon_k.$$

The theorem suggests a stopping condition which is simple to verify when the projected gradient can be easily computed, e.g., when $\mathcal{S} = \mathbb{R}^n_+$ or $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0},\, \mathbf{e}^T\mathbf{x} = \beta\}$

Comparison with other methods

- SGP (Scaled Gradient Projection method) [Bonettini, Zanella, Zanni '09]
- PDAL (Primal Dual method) [Chambolle, Pock '11, Malitsky, Pock '18]
- SPIRAL-TAP (Sparse Poisson Intensity Reconstruction ALgorithms - proximal gradient method) [Harmany, Marcia, Willet '12]
- Split-Bregman [Getreuer '12]
- VMILA (Variable Metric Inexact Line-search Algorithm - proximal gradient method) [Bonettini, Loris, Porta, Prato '16]

Aim: analyze the effect of employing second-order information and smoothing the TV

Implementation details

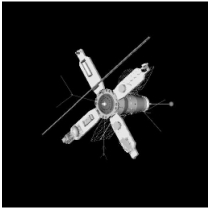
- SGP used for the solution of the quadratic subproblem
  - monotone line search
  - adaptive Barzilai-Borwein rule for the steplength selection
  - stopping condition for SGP: $\|\nabla_{\mathcal{S}} F_k(\widehat{\mathbf{x}}^{(k)})\| \leq \theta^k \|\nabla_{\mathcal{S}} F_k(\mathbf{x}^{(0)})\|$, with $\theta = 0.1$, num. iters ≤ 10
- strong convexity parameter $\gamma = 10^{-5}$
- nonmonotone line search proposed [Grippo, Lampariello, Lucidi 1986], memory length $= 5$, $\eta = 10^{-5}$, $\delta = 0.5$
- starting guess $\mathbf{x}^{(0)} = \mathbf{y}$
- stopping criterion: $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leq \text{Tol}\,\|\mathbf{x}^{(k)}\|$ or max time ≥ 25 sec

Test problems

- cameraman and satellite image

CAMERAMAN

SATELLITE



- convolution with a Gaussian PSF, $\sigma = 1.4, 2$(cameraman,satellite)
- addition of constant background $\mathbf{b} = 10^{-10}$
- degradation with Poisson noise SNR $=$ 35,40

$$\text{SNR} = 10 \log_{10} \frac{N^*}{\sqrt{N^* + N^{\text{back}}}}$$

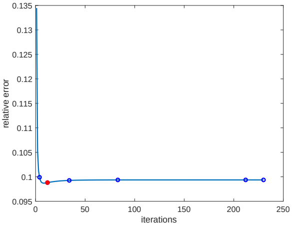
N^*, N^{back} $=$ total # photons in the original image $\mathbf{x}^*$ and in the background

- selection of the regularization parameter by trial and error

Restored images

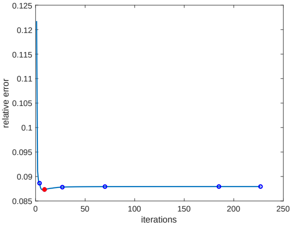
SNR = 35

RESTORED - SNR = 35



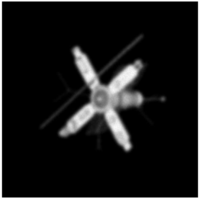
SNR = 40

RESTORED - SNR = 40

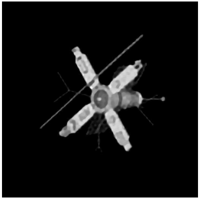


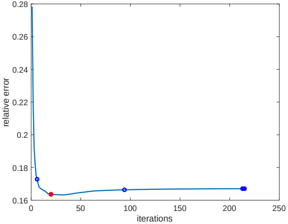
Restored images

SNR = 35

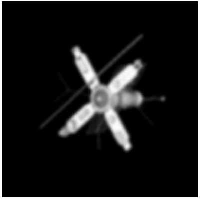


RESTORED - SNR = 35

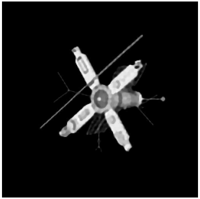


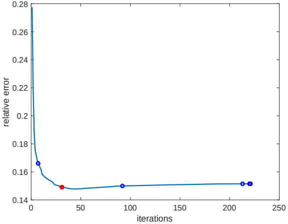


SNR = 40



RESTORED - SNR = 40





Comparison with other methods

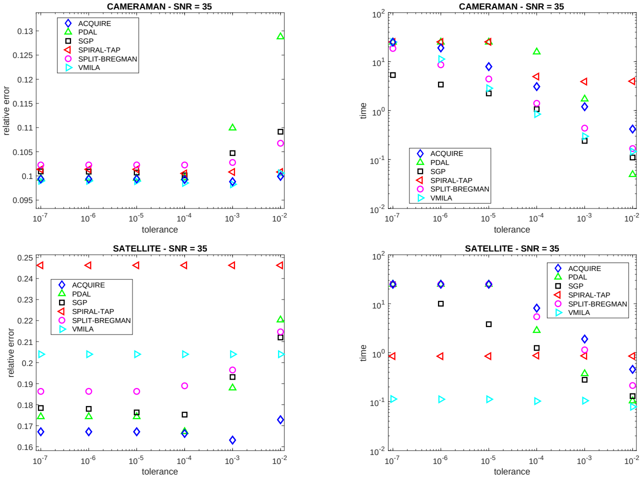


Figura: Test problems with SNR = 35: relative error (left) and execution time (right) versus tolerance, for all methods.

Comparison with other methods

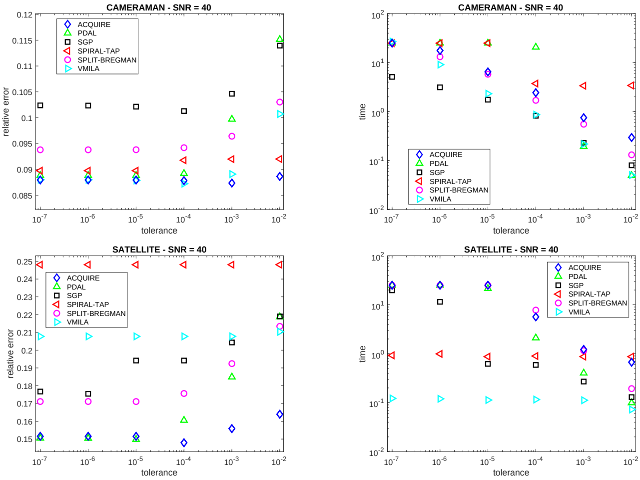


Figura: Test problems with SNR = 40: relative error (left) and execution time (right) versus tolerance, for all methods.

Method	Min rel err	MSSIM	Iters	Time	Tol
		cameraman			
ACQUIRE	9.88e−02	8.01e−01	11	1.20e+00	1.00e−03
PDAL	9.95e−02	8.02e−01	2717	2.50e+01	1.00e−06
SGP	9.86e−01	8.01e−01	67	1.07e+00	1.00e−04
SPIRAL-TV	1.01e−01	7.99e−01	62	4.95e+00	1.00e−04
SPLIT-BREGMAN	1.02e−01	8.04e−01	116	1.41e+00	1.00e−04
VMILA	9.83e−02	8.00e−01	18	2.98e−01	1.00e−03
		satellite			
ACQUIRE	1.63e−01	9.62e−01	20	1.90e+00	1.00e−03
PDAL	1.67e−01	9.60e−01	318	2.85e+00	1.00e−04
SGP	1.65e−01	9.61e−01	84	1.25e+00	1.00e−04
SPIRAL-TV	2.46e−01	9.11e−01	51	8.53e−01	1.00e−02
SPLIT-BREGMAN	1.86e−01	9.45e−01	2132	2.50e+01	1.00e−05
VMILA	2.04e−01	9.40e−01	9	7.87e−02	1.00e−02

Tabella: Test problems with SNR = 35: minimum relative error achieved by each method and corresponding MSSIM value, number of iterations, execution time and tolerance.

Method	Min Rel Err	MSSIM	Iters	Time	Tol
cameraman					
ACQUIRE	8.73e−02	8.42e−01	8	7.49e−01	1.00e−03
PDAL	8.88e−02	8.20e−01	2788	2.50e+01	1.00e−05
SGP	1.01e−01	8.42e−01	53	8.20e−01	1.00e−04
SPIRAL-TV	8.97e−02	8.36e−01	287	2.50e+01	1.00e−05
SPLIT-BREGMAN	9.38e−02	8.41e−01	2137	2.50e+01	1.00e−07
VMILA	8.72e−02	8.42e−01	58	8.66e−01	1.00e−04
satellite					
ACQUIRE	1.48e−01	9.70e−01	59	5.65e+00	1.00e−04
PDAL	1.50e−01	9.67e−01	2418	2.14e+01	1.00e−05
SGP	1.75e−01	9.69e−01	674	1.15e+01	1.00e−06
SPIRAL-TV	2.48e−01	9.08e−01	51	8.71e−01	1.00e−02
SPLIT-BREGMAN	1.71e−01	9.53e−01	2135	2.50e+01	1.00e−05
VMILA	2.08e−01	9.37e−01	10	1.12e−01	1.00e−03

Tabella: Test problems with SNR = 40: minimum relative error achieved by each method and corresponding MSSIM value, number of iterations, execution time and tolerance.

Conclusions

The proposed ACQUIRE

- requires low accuracy in the solution of the quadratic subproblem
- uses second order information and seems able to strongly reduce the reconstruction error in the first iterations
- is competitive with well-established methods since it achieves a tradeoff between accuracy and efficiency
- can be easily generalized to other l1-norm based regularization functionals

Thank you for listening!

. . . questions?