Proximal methods: tools for solving inverse problems on a large scale

Emilie Chouzenoux
LIGM, UPEM

Journée du Labex Bézout:

Data Science and Massive Data Analysis

12 June 2014

UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE

Joint work with




J.-C. Pesquet A. Repetti F. Abboud
LIGM, University Paris East

A. Florescu
Dunarea de Jos University
Galati, Romania

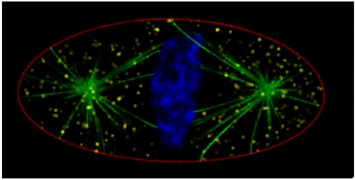


J.-H. Chenot L. Laborelli
Institut National Audiovisuel (INA)

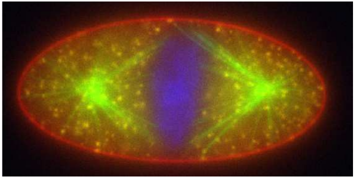
P. Ciuciu
CEA, NeuroSpin center

Inverse problems and large scale optimization

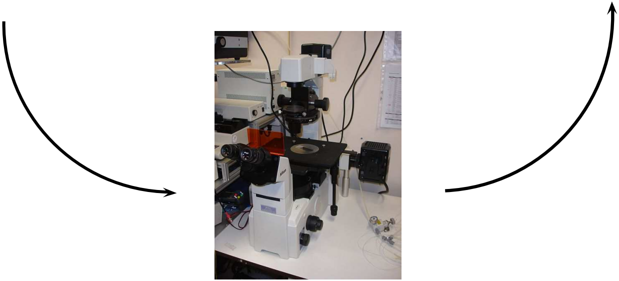
[Microscopy, ISBI Challenge 2013, F. Soulez]



Original image

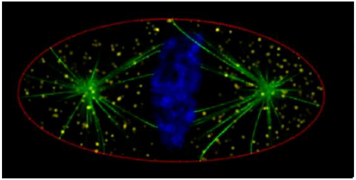


Degraded image



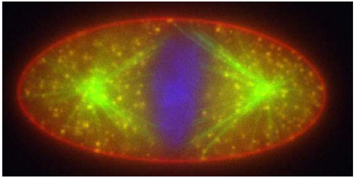
Inverse problems and large scale optimization

[Microscopy, ISBI Challenge 2013, F. Soulez]



Original image

$\overline{x} \in \mathbb{R}^N$



Degraded image

$z = \mathcal{D}(H\overline{x}) \in \mathbb{R}^M$

- $H \in \mathbb{R}^{M \times N}$: matrix associated with the degradation operator.
- $\mathcal{D}\colon \mathbb{R}^M \to \mathbb{R}^M$: noise degradation.

Inverse problem:

Find a good estimate of $\overline{x}$ from the observations z, using some *a priori* knowledge on $\overline{x}$ and on the noise characteristics.

Inverse problems and large scale optimization

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{D}(H\overline{x})$.

- Inverse filtering (if $M = N$ and H is invertible)

$$\begin{aligned}\widehat{x} &= H^{-1}z \\ &= H^{-1}(H\overline{x} + b) \quad \leftarrow \quad \text{if } b \in \mathbb{R}^M \text{ is an additive noise} \\ &= \overline{x} + H^{-1}b\end{aligned}$$

$\rightarrow$ Closed form expression, but amplification of the noise if H is ill-conditioned (*ill-posed problem*).

Inverse problems and large scale optimization

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{D}(H\overline{x})$.

- ~~Inverse filtering~~
- Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \underbrace{f_1(x)}_{\text{Data fidelity term}} + \underbrace{f_2(x)}_{\text{Regularization term}}$$

Inverse problems and large scale optimization

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{D}(H\overline{x})$.

- ~~Inverse filtering~~
- Variational approach

$$\widehat{x} \in \operatorname*{Argmin}_{x\in\mathbb{R}^N} \underbrace{f_1(x)}_{\text{Data fidelity term}} + \underbrace{f_2(x)}_{\text{Regularization term}}$$

Examples of data fidelity term

- Gaussian noise

$$(\forall x \in \mathbb{R}^N) \qquad f_1(x) = \frac{1}{\sigma^2}\|Hx - z\|^2$$

- Poisson noise

$$(\forall x \in \mathbb{R}^N) \qquad f_1(x) = \sum_{m=1}^{M} \Big([Hx]^{(m)} - z^{(m)} \log([Hx]^{(m)})\Big)$$

Examples of regularization terms (1)

- Admissibility constraints

$$\text{Find} \qquad x \in C = \bigcap_{m=1}^{M} C_m$$

where $(\forall m \in \{1, \ldots, M\})\ C_m \subset \mathbb{R}^N$.

Examples of regularization terms (1)

- Admissibility constraints

$$\text{Find} \qquad x \in C = \bigcap_{m=1}^{M} C_m$$

where $(\forall m \in \{1, \ldots, M\})\ C_m \subset \mathbb{R}^N$.

- Variational formulation

$$(\forall x \in \mathbb{R}^N) \qquad f_2(x) = \sum_{m=1}^{M} \iota_{C_m}(x)$$

where, for all $m \in \{1, \ldots, M\}$, ι_{C_m} is the indicator function of C_m:

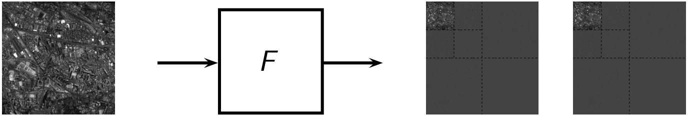
$$(\forall x \in \mathbb{R}^N) \qquad \iota_{C_m}(x) = \begin{cases} 0 & \text{if } x \in C_m \\ +\infty & \text{otherwise.} \end{cases}$$

Examples of regularization terms (2)

- ℓ_1 norm (analysis approach)

$$(\forall x \in \mathbb{R}^N) \qquad f_2(x) = \sum_{k=1}^{K} \left| [Fx]^{(k)} \right| = \|Fx\|_1$$

$F \in \mathbb{R}^{K \times N}$: Frame decomposition operator ($K \geq N$)



signal x frame coefficients

Examples of regularization terms (2)

- ℓ_1 norm (analysis approach)

$$(\forall x \in \mathbb{R}^N) \qquad f_2(x) = \sum_{k=1}^{K} \left| [Fx]^{(k)} \right| = \|Fx\|_1$$

- Total variation

$$(\forall x = (x^{(i_1,i_2)})_{1 \le i_1 \le N_1, 1 \le i_2 \le N_2} \in \mathbb{R}^{N_1 \times N_2})$$

$$f_2(x) = \mathrm{tv}(x) = \sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \|\nabla x^{(i_1,i_2)}\|_2$$

$\nabla x^{(i_1,i_2)}$: discrete gradient at pixel (i_1, i_2).

Inverse problems and large scale optimization

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{D}(H\overline{x})$.

- ~~Inverse filtering~~
- Variational approach (more general context)

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \sum_{i=1}^{m} f_i(x)$$

where f_i may denote a data fidelity term / a (hybrid) regularization term / constraint.

Inverse problems and large scale optimization

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{D}(H\overline{x})$.

- ~~Inverse filtering~~
- Variational approach (more general context)

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \sum_{i=1}^{m} f_i(x)$$

where f_i may denote a data fidelity term / a (hybrid) regularization term / constraint.

$\rightarrow$ Often no closed form expression or solution expensive to compute (especially in large scale context).

- **Need for an efficient iterative minimization strategy !**

Outline

1. Proximal-based algorithms
    - Proximity operator
    - Forward-Backward algorithm
    - Acceleration via metric change
    - Acceleration via block alternation

2. Applications
    - Parallel magnetic resonance imaging
    - Phase retrieval
    - Blind deconvolution of television video
    - Multi-channel image restoration

Proximal-based algorithms

Gradient and subgradient algorithms

Optimization problem: Minimization of function $f \in \Gamma_0(\mathbb{R}^N)$ on $\mathbb{R}^N$.

- If f has a β-Lipschitz gradient with $\beta \in]0, +\infty[$

$$(\forall \ell \in \mathbb{N}) \qquad x_{\ell+1} = x_\ell - \gamma_\ell \nabla f(x_\ell) \qquad \text{explicit step}$$

with $0 < \inf_{\ell \in \mathbb{N}} \gamma_\ell$ and $\sup_{\ell \in \mathbb{N}} \gamma_\ell < 2\beta^{-1}$.

Gradient and subgradient algorithms

Optimization problem: Minimization of function $f \in \Gamma_0(\mathbb{R}^N)$ on $\mathbb{R}^N$.

- If f has a β-Lipschitz gradient with $\beta \in]0, +\infty[$

$$(\forall \ell \in \mathbb{N}) \qquad x_{\ell+1} = x_\ell - \gamma_\ell \nabla f(x_\ell)$$ explicit step

with $0 < \inf_{\ell \in \mathbb{N}} \gamma_\ell$ and $\sup_{\ell \in \mathbb{N}} \gamma_\ell < 2\beta^{-1}$.

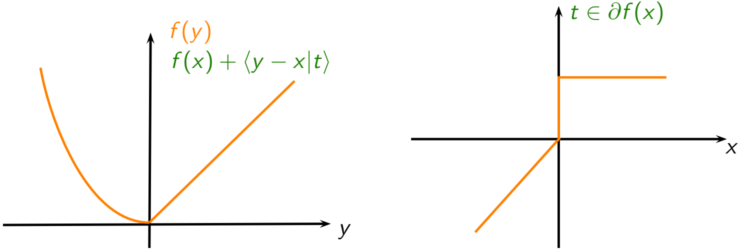
- When f is nonsmooth, replace gradient with subgradient

$$\partial f(x) = \left\{ t \in \mathbb{R}^N | (\forall y \in \mathbb{R}^N) \;\; f(y) \geq f(x) + \langle t | y - x \rangle \right\}$$

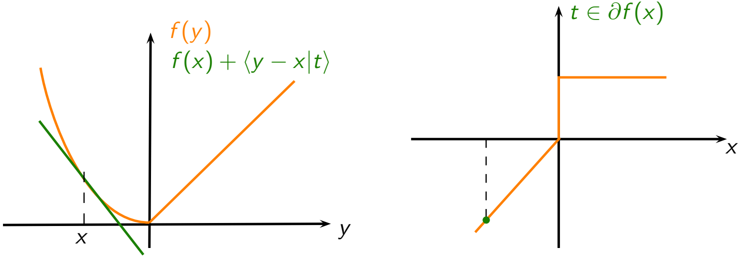
$t \in \partial f(x)$: subgradient at $x \in \mathbb{R}^N$

$\partial f : \mathbb{R}^N \to 2^{\mathbb{R}^N}$: subdifferential

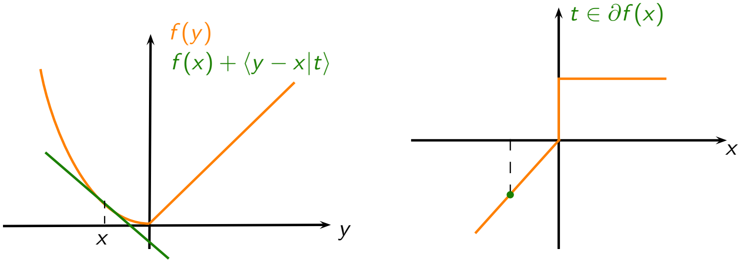
Subdifferential



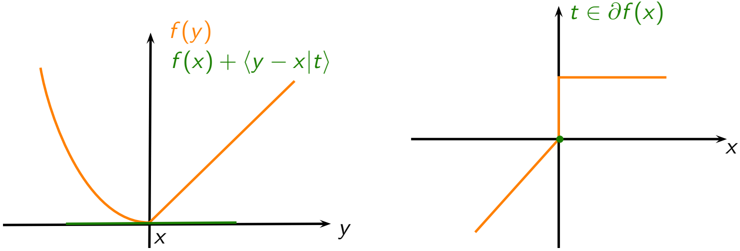
Subdifferential



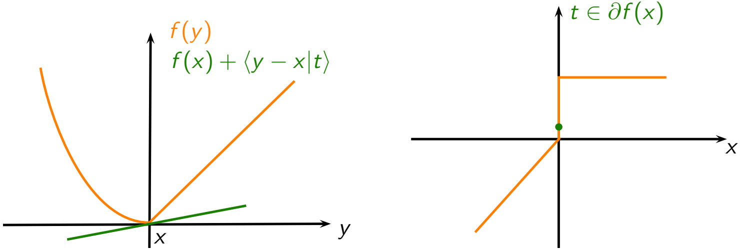
Subdifferential



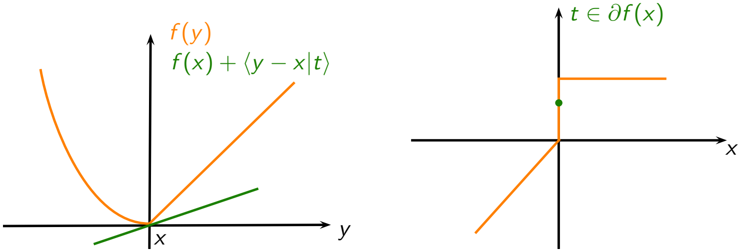
Subdifferential



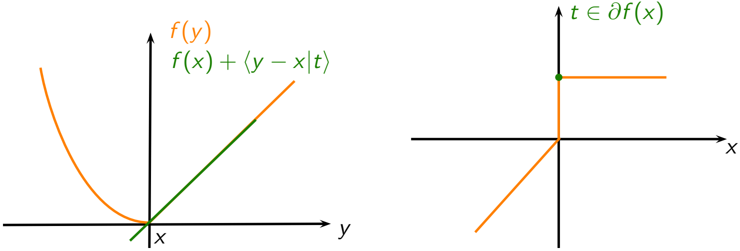
Subdifferential



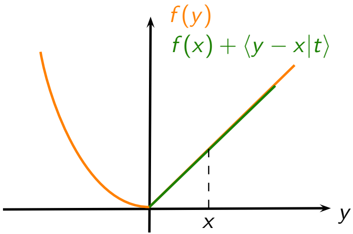
Subdifferential

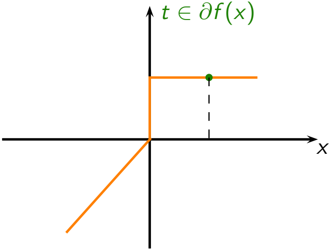


Subdifferential



Subdifferential





Example of subdifferential

Example:

- If f is differentiable at $x \in \mathbb{R}^N$ then $\partial f(x) = \{\nabla f(x)\}$.
- If $f = |\cdot|$ then

$$(\forall x \in \mathbb{R}) \qquad \partial f(x) = \begin{cases} \{\text{sign}(x)\} & \text{if } x \neq 0 \\ [-1, 1] & \text{if } x = 0 \end{cases}$$

From the subgradient algorithm ...

Optimization problem: Minimization of function $f \in \Gamma_0(\mathbb{R}^N)$ on $\mathbb{R}^N$.

Subgradient algorithm [Shor,1979]

$$(\forall \ell \in \mathbb{N}) \qquad x_{\ell+1} = x_\ell - \gamma_\ell t_\ell, \quad t_\ell \in \partial f(x_\ell)$$

where $(\forall \ell \in \mathbb{N})$ $\gamma_\ell \in]0,+\infty[$ such that $\sum_{\ell=0}^{+\infty} \gamma_\ell^2 < +\infty$ and $\sum_{\ell=0}^{+\infty} \gamma_\ell = +\infty$.

From the subgradient algorithm ...

Optimization problem: Minimization of function $f \in \Gamma_0(\mathbb{R}^N)$ on $\mathbb{R}^N$.

Subgradient algorithm [Shor,1979]

$$(\forall \ell \in \mathbb{N}) \qquad x_{\ell+1} = x_\ell - \gamma_\ell t_\ell, \quad t_\ell \in \partial f(x_\ell)$$

where $(\forall \ell \in \mathbb{N})\ \gamma_\ell \in]0, +\infty[$ such that $\sum_{\ell=0}^{+\infty} \gamma_\ell^2 < +\infty$ and $\sum_{\ell=0}^{+\infty} \gamma_\ell = +\infty$.

Implicit form

$$(\forall \ell \in \mathbb{N}) \qquad x_{\ell+1} = x_\ell - \gamma_\ell t'_\ell, \quad t'_\ell \in \partial f(x_{\ell+1})$$

$$\Leftrightarrow \quad x_\ell - x_{\ell+1} \in \gamma_\ell \partial f(x_{\ell+1})$$

... to the origins of the proximity operator!

Property

Let $\varphi \in \Gamma_0(\mathbb{R}^N)$. For all $x \in \mathbb{R}^N$, there exists a unique vector $\widehat{x} \in \mathbb{R}^N$ such that $x - \widehat{x} \in \partial\varphi(\widehat{x})$.

- Let $\widehat{x} = \mathrm{prox}_{\varphi}(x)$.
- $\mathrm{prox}_{\varphi} : \mathbb{R}^N \to \mathbb{R}^N$: proximity operator

Proximal point algorithm

$$(\forall \ell \in \mathbb{N}) \qquad x_\ell - x_{\ell+1} \in \gamma_\ell \partial f(x_{\ell+1})$$
$$\Leftrightarrow \quad x_{\ell+1} = \mathrm{prox}_{\gamma_\ell f}(x_\ell)$$

where $\inf_{\ell \in \mathbb{N}} \gamma_\ell > 0$ such that $\sum_{\ell=0}^{+\infty} \gamma_\ell = +\infty$.

Another definition of the proximity operator

Property

Let $f \in \Gamma_0(\mathbb{R}^N)$.
For all $x \in \mathbb{R}^N$, $\mathrm{prox}_f(x)$ is the unique minimizer of

$$y \mapsto f(y) + \frac{1}{2}\|x - y\|^2.$$

Example:

Let C a closed non empty subset of $\mathbb{R}^N$. Then, prox_{ι_C} reduces to the projection operator on the set C.

Some other examples

- Explicit form for objective functions associated to the usual log-concave probability densities [Chaux *et al.* - 2007]

  - Laplace
  - Generalized gaussian
  - maximum entropy
  - gamma
  - uniform
  - Weibull
  - Generalized inverse gaussian
  - Gaussian
  - Huber
  - Smoothed Laplace
  - chi
  - triangular
  - Pearson type I
  - ...

- And many other functions ! [Combettes, Pesquet - 2010]

Forward-backward algorithm

Optimization problem:
Minimization of $f + g$ on $\mathbb{R}^N$, assuming that g has a β-Lipschitz gradient.

Forward-backward algorithm

$$
\begin{aligned}
(\forall \ell \in \mathbb{N}) \qquad & x_{\ell+1} = x_\ell - \gamma_\ell (t'_\ell + \nabla g(x_\ell)), \quad t'_\ell \in \partial f(x_{\ell+1}) \\
\Leftrightarrow \quad & x_{\ell+1} = \mathrm{prox}_{\gamma_\ell f}(x_\ell - \gamma_\ell \nabla g(x_\ell))
\end{aligned}
$$

Forward-backward algorithm

Optimization problem:
Minimization of $f + g$ on $\mathbb{R}^N$, assuming that g has a β-Lipschitz gradient.

Forward-backward algorithm

$$(\forall \ell \in \mathbb{N}) \qquad \boxed{x_{\ell+1} = x_\ell + \lambda_\ell \big(\operatorname{prox}_{\gamma_\ell f}(x_\ell - \gamma_\ell \nabla g(x_\ell)) - x_\ell \big)}$$

Convergence of $(x_\ell)_{\ell\in\mathbb{N}}$ if $0 < \inf_{\ell\in\mathbb{N}} \gamma_\ell$, $\sup_{\ell\in\mathbb{N}} \gamma_\ell < 2\beta^{-1}$, $0 < \inf_{\ell\in\mathbb{N}} \lambda_\ell$ and $\sup_{\ell\in\mathbb{N}} \lambda_\ell \leq 1$.

- f and g convex [Chen,Rockafellar,1997][Combettes,Wajs,2005]
- f and g nonconvex (under Kurdyka-Łojasiewicz assumption) [Attouch *et al.* - 2011]

How to make the forward-backward algorithm efficient for big data optimization ?

First trick: Majoration-Minimization strategy

MM point of view

Majorize-Minimize Assumption

- For every $\ell \in \mathbb{N}$, there exists a symmetric positive definite (SPD) matrix $A_\ell(x_\ell) \in \mathbb{R}^{N\times N}$ such that for every $x \in \mathbb{R}^N$

$$Q(x, x_\ell) = g(x_\ell) + (x - x_\ell)^\top \nabla g(x_\ell) + \frac{1}{2}(x - x_\ell)^\top A_\ell(x_\ell)(x - x_\ell),$$

is a **majorant function** of g at x_ℓ on $\operatorname{dom} f$, i.e.,

$$g(x_\ell) = Q(x_\ell, x_\ell) \quad \text{and} \quad (\forall x \in \operatorname{dom} f) \quad g(x) \le Q(x, x_\ell).$$

MM point of view

Majorize-Minimize Assumption

- For every $\ell \in \mathbb{N}$, there exists a symmetric positive definite (SPD) matrix $A_\ell(x_\ell) \in \mathbb{R}^{N\times N}$ such that for every $x \in \mathbb{R}^N$

$$Q(x, x_\ell) = g(x_\ell) + (x - x_\ell)^\top \nabla g(x_\ell) + \frac{1}{2}(x - x_\ell)^\top A_\ell(x_\ell)(x - x_\ell),$$

is a **majorant function** of g at x_ℓ on dom f, i.e.,

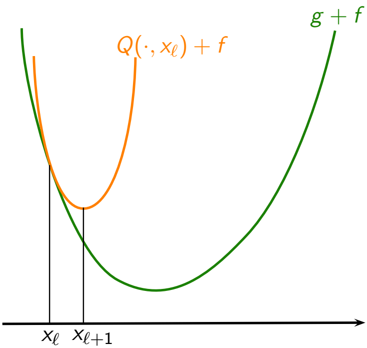
$$g(x_\ell) = Q(x_\ell, x_\ell) \quad \text{and} \quad (\forall x \in \operatorname{dom} f) \quad g(x) \leq Q(x, x_\ell).$$

g is differentiable with a β-Lipschitzian gradient on a convex subset of $\mathbb{R}^N$ $\Rightarrow$ $A_\ell(x_\ell) \equiv \beta\,\mathrm{Id}$ satisfies the above assumption [Bertsekas - 1999]

MM algorithm [Jacobson and Fessler - 2007]

MM Algorithm

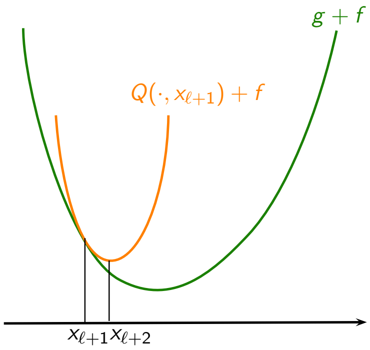
$$x_{\ell+1} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}}\ f(x) + Q(x, x_\ell)$$



MM algorithm [Jacobson and Fessler - 2007]

MM Algorithm

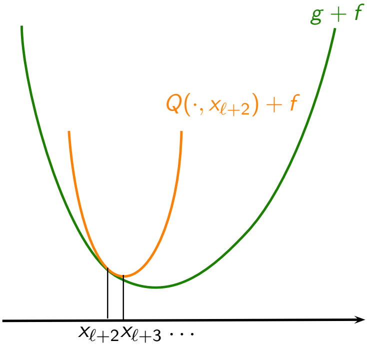
$$x_{\ell+1} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \; f(x) + Q(x, x_\ell)$$



MM algorithm [Jacobson and Fessler - 2007]

MM Algorithm

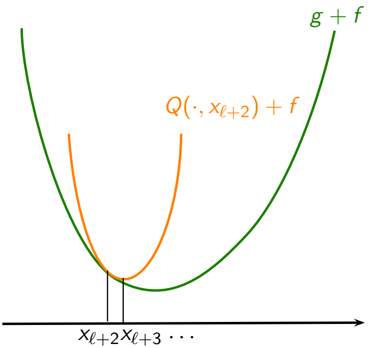
$$x_{\ell+1} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \; f(x) + Q(x, x_\ell)$$



MM algorithm [Jacobson and Fessler - 2007]

MM Algorithm

$$x_{\ell+1} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \; f(x) + Q(x, x_\ell)$$



$\Leftrightarrow$ Forward-backward algorithm with

- $A_\ell(x_\ell) \equiv \beta \, \mathrm{Id}$
- $\lambda_\ell \equiv 1$
- $\gamma_\ell \equiv 1$

$\rightsquigarrow$ Why not trying more sophisticated matrices $(A_\ell)_{\ell \in \mathbb{N}}$?

▶ **Variable metric forward-backward algorithm !**

Acceleration via metric change

Definition

Let $x \in \mathbb{R}^N$. Let A be a SPD matrix. The proximity operator relative to the metric induced by A is defined by

$$\operatorname{prox}_{\gamma^{-1}A,\, f}(x) = \underset{y \in \mathbb{R}^N}{\operatorname{Argmin}}\, f(y) + \frac{1}{2\gamma}\|y - x\|_A^2.$$

Variable metric forward-backward algorithm

$$(\forall \ell \in \mathbb{N}) \qquad x_{\ell+1} = \operatorname{prox}_{\gamma_\ell^{-1} A_\ell(x_\ell),\, f}\left(x_\ell - \gamma_\ell A_\ell(x_\ell)^{-1} \nabla g(x_\ell)\right).$$

Convergence of $(x_\ell)_{\ell \in \mathbb{N}}$

- f and g convex [Combettes *et al.* - 2012]
- f and g nonconvex [Chouzenoux *et al.* - 2013]

▶ **Significant acceleration in practice !**

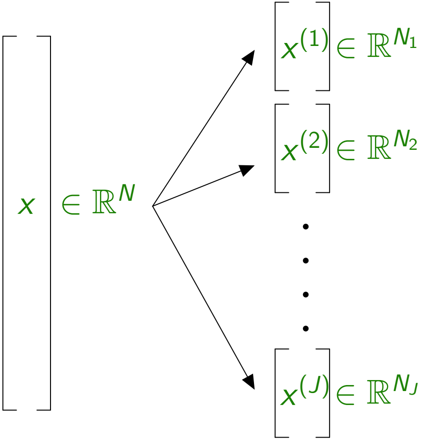
Second trick: Block alternation

Acceleration via block alternation

- Assumption: f is an additively block separable function.

Acceleration via block alternation

- Assumption: f is an additively block separable function.



$$N = \sum_{j=1}^{J} N_j$$

Acceleration via block alternation

- Assumption: f is an additively block separable function.

$$f\left(\begin{bmatrix} x \end{bmatrix}\right) = f\left(\begin{matrix} \begin{bmatrix} x^{(1)} \end{bmatrix} \\ \begin{bmatrix} x^{(2)} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} x^{(J)} \end{bmatrix} \end{matrix}\right) = \sum_{j=1}^{J} f_j(x^{(j)})$$

Acceleration via block alternation

Block coordinate forward-backward algorithm

$(\forall \ell \in \mathbb{N})$, pick a block $j_\ell \in \{1, \dots, J\}$, and update:

$$\begin{cases} x_{\ell+1}^{(j_\ell)} = \mathrm{prox}_{\gamma_\ell f_{j_\ell}} \left(x_\ell^{(j_\ell)} - \gamma_\ell \nabla_{j_\ell} g(x_\ell) \right) \\ x_{\ell+1}^{(\bar{\jmath}_\ell)} = x_\ell^{(\bar{\jmath}_\ell)} \end{cases}$$

- Convergence of $(x_\ell)_{\ell \in \mathbb{N}}$ (assuming a cyclic update rule) established in [Bolte *et al.* - 2013] for possibly nonconvex functions f and g verifying Kurdyka-Łojasiewicz assumption.

Acceleration via block alternation

Block coordinate forward-backward algorithm

$(\forall \ell \in \mathbb{N})$, pick a block $j_\ell \in \{1, \dots, J\}$, and update:

$$\begin{cases} x_{\ell+1}^{(j_\ell)} = \operatorname{prox}_{\gamma_\ell f_{j_\ell}} \left(x_\ell^{(j_\ell)} - \gamma_\ell \nabla_{j_\ell} g(x_\ell) \right) \\ x_{\ell+1}^{(\bar{\jmath}_\ell)} = x_\ell^{(\bar{\jmath}_\ell)} \end{cases}$$

- Convergence of $(x_\ell)_{\ell \in \mathbb{N}}$ (assuming a cyclic update rule) established in [Bolte *et al.* - 2013] for possibly nonconvex functions f and g verifying Kurdyka-Łojasiewicz assumption.
- **Block alternation presents several advantages:**
  - ✓ more flexibility,
  - ✓ reduced computational cost at each iteration,
  - ✓ reduced memory requirement.

Combining first and second trick ...

Acceleration via block alternation and metric change

Block coordinate variable metric forward-backward algorithm
$(\forall \ell \in \mathbb{N})$, pick a block $j_\ell \in \{1, \ldots, J\}$, and update

$$\begin{cases} x_{\ell+1}^{(j_\ell)} = \text{prox}_{\gamma_\ell^{-1} A_{j_\ell}(x_\ell), f_{j_\ell}} \left(x_\ell^{(j_\ell)} - \gamma_\ell A_{j_\ell}(x_\ell)^{-1} \nabla_{j_\ell} g(x_\ell) \right) \\ x_{\ell+1}^{(\bar{\jmath}_\ell)} = x_\ell^{(\bar{\jmath}_\ell)} \end{cases}$$

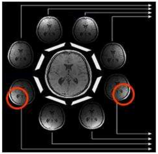
- Convergence of $(x_\ell)_{\ell \in \mathbb{N}}$ (assuming a quasi cyclic update rule) established in [Chouzenoux *et al.* - 2013] for nonconvex functions f and g verifying Kurdyka-Łojasiewicz assumption.

▶ **Benefits from the advantages of both acceleration techniques!**

Applications

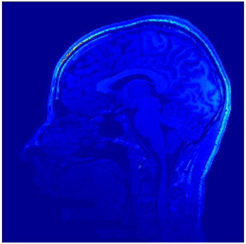
Parallel Magnetic Resonance Imaging [Florescu *et al.* - 2014]

Challenges:

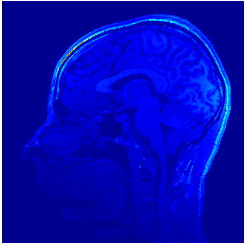


- Parallel acquisition and compressive sensing
- Complex-valued signals

Results:



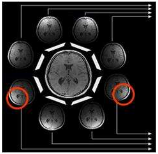
Original



Proposed method

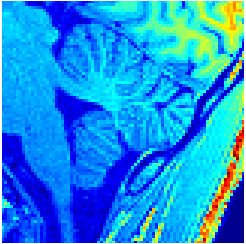
Parallel Magnetic Resonance Imaging [Florescu *et al.* - 2014]

Challenges:

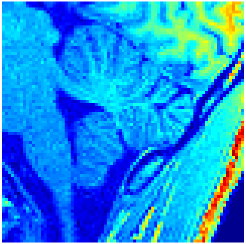


- Parallel acquisition and compressive sensing
- Complex-valued signals

Results:



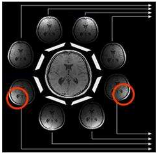
Original



Proposed method

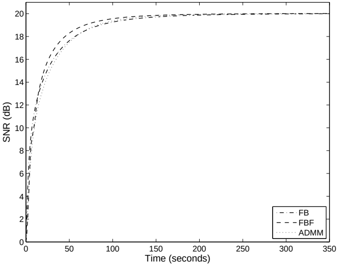
Parallel Magnetic Resonance Imaging [Florescu *et al.* - 2014]

Challenges:



- Parallel acquisition and compressive sensing
- Complex-valued signals

Results:



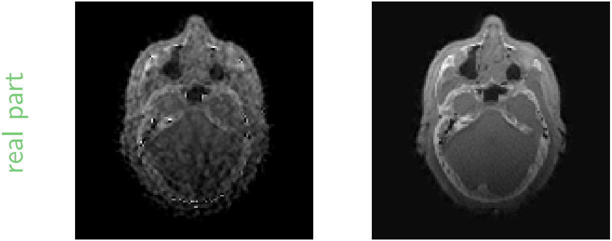
Convergence speed of several proximal-based algorithms

Phase retrieval [Repetti *et al.* - ICIP 2014]

Challenges:

- Only the modulus of the observed data is available
- Non-Fourier measurements
- Nonconvex data fidelity term

Results:



SparseFienup Proposed method

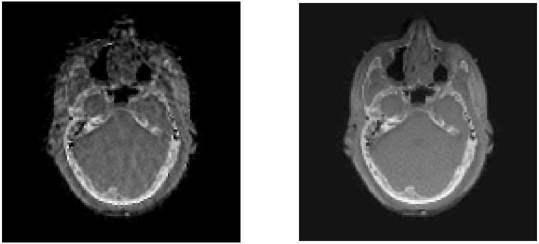
Phase retrieval [Repetti *et al.* - ICIP 2014]

Challenges:

- Only the modulus of the observed data is available
- Non-Fourier measurements
- Nonconvex data fidelity term

Results:

imaginary part



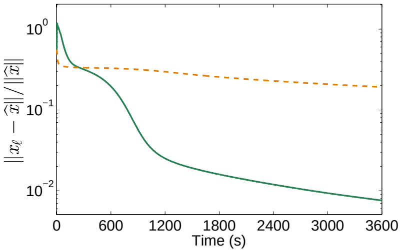
SparseFienup Proposed method

Phase retrieval [Repetti *et al.* - ICIP 2014]

Challenges:

- Only the modulus of the observed data is available
- Non-Fourier measurements
- Nonconvex data fidelity term

Results:



Influence of the variable metric strategy

Blind deconvolution of video [Abboud *et al.* - EUSIPCO 2014]

Challenges:

- The degradation blur operator is unknown
- Nonconvex data fidelity term

Results:

Observed

Restored

Blind deconvolution of video [Abboud *et al.* - EUSIPCO 2014]

Challenges:

- The degradation blur operator is unknown
- Nonconvex data fidelity term

Results:

Observed

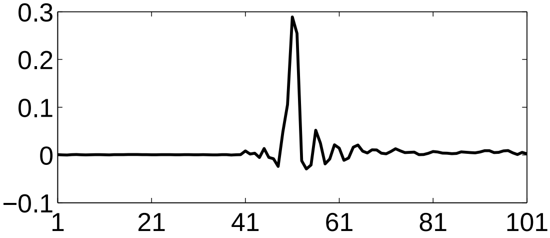
Restored

Blind deconvolution of video [Abboud *et al.* - EUSIPCO 2014]

Challenges:

- The degradation blur operator is unknown
- Nonconvex data fidelity term

Results:

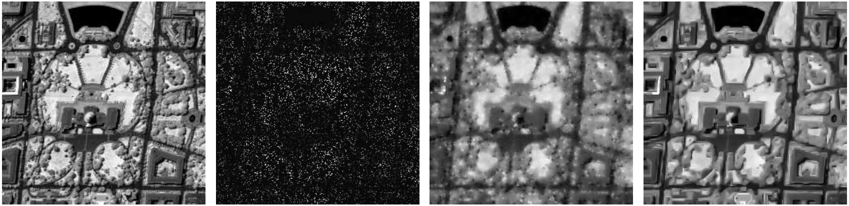


Estimated blur kernel

Multi-channel image restoration [Chierchia *et al.* - 2014]

Challenges:

- Deal with images having a large number of components
- Circumvent the choice of regularization parameters by introducing suitable nonlocal constraints
- Develop epigraphical techniques to address these constraints efficiently

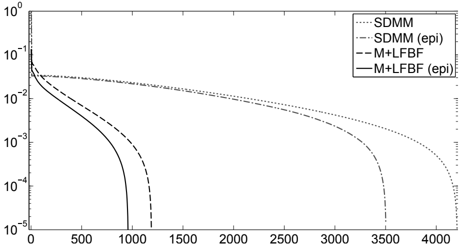


Original Observed H-TV Proposed

Multi-channel image restoration [Chierchia *et al.* - 2014]

Challenges:

- Deal with images having a large number of components
- Circumvent the choice of regularization parameters by introducing suitable nonlocal constraints
- Develop epigraphical techniques to address these constraints efficiently



Time (s.)

Constrained formulation VS Variational formulation

Conclusion

✓ Proximal-based algorithms: An efficient tool for solving large scale possibly difficult optimization problem;

Conclusion

✓ Proximal-based algorithms: An efficient tool for solving large scale possibly difficult optimization problem;

✓ Two recipes for accelerating the algorithms:

- Majoration-Minimization strategy
- Block alternation

Conclusion

✓ Proximal-based algorithms: An efficient tool for solving large scale possibly difficult optimization problem;

✓ Two recipes for accelerating the algorithms:

- Majoration-Minimization strategy
- Block alternation

✓ No need to invert large size matrices through primal-dual forward-backward based methods;

Conclusion

✓ Proximal-based algorithms: An efficient tool for solving large scale possibly difficult optimization problem;

✓ Two recipes for accelerating the algorithms:

- Majoration-Minimization strategy
- Block alternation

✓ No need to invert large size matrices through primal-dual forward-backward based methods;

✓ Parallel implementations possible thanks to splitting techniques.

Conclusion

✓ Proximal-based algorithms: An efficient tool for solving large scale possibly difficult optimization problem;

✓ Two recipes for accelerating the algorithms:

- Majoration-Minimization strategy
- Block alternation

✓ No need to invert large size matrices through primal-dual forward-backward based methods;

✓ Parallel implementations possible thanks to splitting techniques.

Future challenges: Find more tricks!

Thank you ! Questions ?

E. Chouzenoux, J.-C. Pesquet and A. Repetti.
Variable Metric Forward-Backward Algortihm for Minimizing the Sum of a Differentiable Function and a Convex Function.
To appear in J. Optim. Theory Appl, 2013.

E. Chouzenoux, J.-C. Pesquet and A. Repetti.
A Block Coordinate Variable Metric Forward-Backward Algorithm.
Tech. Rep., 2013. Available on
http://www.optimization-online.org/DB_HTML/2013/12/4178.html.

A. Florescu, E. Chouzenoux, J.-C. Pesquet, P. Ciuciu and S. Ciochina.
A Majorize-Minimize Memory Gradient Method for Complex-Valued Inverse Problems.
Signal Processing, Vol. 103, pages 285-295, 2014.

F. Abboud, E. Chouzenoux, J.-C. Pesquet, J.-H. Chenot and L. Laborelli
A Hybrid Alternating Proximal Method for Blind Video Restoration.
Accepted to EUSIPCO 2014.