Extended Forward-Backward Algorithm with n "proximable" functions

Hugo Raguet*
Jalal Fadili†, Gabriel Peyré‡

May 25, 2011

*CEREMADE, hugo.raguet@gmail.com
†ENSICAEN
‡CEREMADE

A Common Problem...

A Common Problem...

coefficients

x

A Common Problem...

coefficients

x

Φ

$\rightarrow$

$\Phi x + w = Y$

A Common Problem...

coefficients

x

Φ

$\rightarrow$

data-fidelity

$\Phi x + w = Y$

$\frac{1}{2}\|Y - \Phi x\|^2$

A Common Problem...

coefficients *data-fidelity*

Ψx

$\Phi\Psi x + w = Y$

Φ
$\rightarrow$

$\uparrow \ \Psi$

$\frac{1}{2}\|Y - \Phi\Psi x\|^2$

x

A Common Problem...

coefficients

Ψx

$\uparrow \ \Psi$

x

Φ

$\rightarrow$

data-fidelity

$\Phi \Psi x + w = Y$

$\frac{1}{2}\|Y - \Phi \Psi x\|^2$

A Common Problem...

coefficients

Ψx

$\uparrow \ \Psi$

x

Φ

$\rightarrow$

data-fidelity

$\Phi \Psi x + w = Y$

$\frac{1}{2} \| Y - \Phi \Psi x \|^2$

penalization

L_1-norm $\sum_i |x_i|$

A Common Problem...

Find x minimizing $\frac{1}{2}\|Y - \Phi\Psi x\|^2 + \lambda\|x\|_1$

coefficients *data-fidelity* *penalization*

Ψx

$\uparrow \Psi$

x

Φ
$\rightarrow$

$\Phi\Psi x + w = Y$

$\frac{1}{2}\|Y - \Phi\Psi x\|^2$

L_1-norm $\sum_i |x_i|$

DAUPHINE UNIVERSITÉ PARIS NatImages

A Common Problem...

Find	x	minimizing	$F(x)$	$+$	$G(x)$
Find	x	minimizing	$\frac{1}{2}\|Y - \Phi\Psi x\|^2$	$+$	$\lambda\|x\|_1$

x *coefficients* F *data-fidelity* G *penalization*

Ψx

$\uparrow \ \Psi$

x

Φ
$\rightarrow$

$\Phi\Psi x + w = Y$

$\frac{1}{2}\|Y - \Phi\Psi x\|^2$

L_1-norm $\sum_i |x_i|$

DAUPHINE UNIVERSITÉ PARIS NatImages

Towards More Complex Penalization

Towards More Complex Penalization

$\|x\|_1 = \sum_i |x_i|$

Towards More Complex Penalization

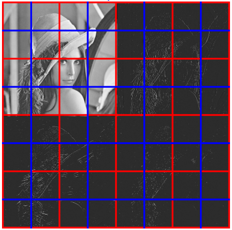
$\|x\|_1 = \sum_i |x_i|$ $\sum_{b \in \mathcal{B}} \sqrt{\sum_{i \in b} x_i^2}$

Towards More Complex Penalization

$\|x\|_1 = \sum_i |x_i|$

$\sum_{b \in \mathcal{B}} \sqrt{\sum_{i \in b} x_i^2}$

$\sum_{b \in \mathcal{B}_1} \sqrt{\sum_{i \in b} x_i^2}$
$+$
$\sum_{b \in \mathcal{B}_2} \sqrt{\sum_{i \in b} x_i^2}$



Towards More Complex Penalization

$\|x\|_1 = \sum_i |x_i|$

$\sum_{b \in \mathcal{B}} \sqrt{\sum_{i \in b} x_i^2}$

$\sum_{b \in \mathcal{B}_1} \sqrt{\sum_{i \in b} x_i^2}$ + $\sum_{b \in \mathcal{B}_2} \sqrt{\sum_{i \in b} x_i^2}$

Decomposition $G = \sum_k G_k$

Conditions

To Deal with High Dimensionnality

Conditions

Interesting Properties

F convex smooth

To Deal with High Dimensionnality

Conditions

Interesting Properties

F convex smooth

$$F(x) = \frac{1}{2}\|Y - Lx\|^2$$
$$\Rightarrow \quad \nabla F(x) = L^\star (Lx - Y)$$

To Deal with High Dimensionnality

Conditions

Interesting Properties

F convex smooth

G convex "proximable"

$$\mathrm{prox}_G(x) \stackrel{\mathrm{def}}{=} \operatorname*{argmin}_y \frac{1}{2}\|x-y\|^2 + G(y)$$

To Deal with High Dimensionnality

Conditions

Interesting Properties

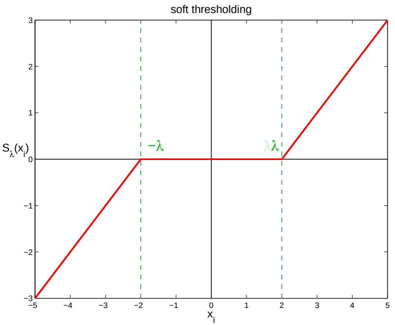
F convex smooth

G convex “proximable”

$$\mathrm{prox}_G(x) \stackrel{\text{def}}{=} \underset{y}{\operatorname{argmin}} \frac{1}{2}\|x-y\|^2 + G(y)$$

$$G(x) = \lambda \textstyle\sum_i |x_i|$$

$$\Rightarrow \quad \mathrm{prox}_G(x) = S_\lambda(x)$$



To Deal with High Dimensionnality

Conditions

Interesting Properties

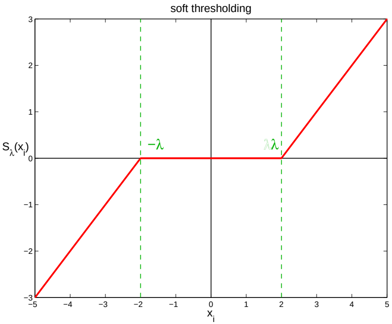
F convex smooth

G convex "proximable"

$$\mathrm{prox}_G(x) \overset{\mathrm{def}}{=} \underset{y}{\mathrm{argmin}}\ \frac{1}{2}\|x-y\|^2 + G(y)$$

$$G(x) = \lambda \sum_{b\in\mathcal{B}} \sqrt{\sum_i x_i^2}$$

$$\Rightarrow \quad \mathrm{prox}_G(x) = S_\lambda{}^{\mathcal{B}}(x)$$



To Deal with High Dimensionnality

Conditions

Interesting Properties

F convex smooth

G convex "proximable"

$$\text{prox}_G(x) \stackrel{\text{def}}{=} \underset{y}{\text{argmin}} \ \tfrac{1}{2}\|x - y\|^2 + G(y)$$

$$G(x) = \lambda \sum_k \sum_{b \in \mathcal{B}_k} \sqrt{\sum_i x_i^2}$$

$$\Rightarrow \quad \text{prox}_G(x) = \quad \text{No closed form expression}$$

To Deal with High Dimensionnality

Conditions

Interesting Properties

F convex smooth

G convex "proximable"

$$\mathrm{prox}_G(x) \stackrel{\text{def}}{=} \underset{y}{\mathrm{argmin}}\ \frac{1}{2}\|x-y\|^2 + G(y)$$

$$G(x) = \lambda \sum_k \sum_{b \in \mathcal{B}_k} \sqrt{\sum_i x_i^2}$$

$$\Rightarrow \quad \mathrm{prox}_G(x) = \quad \text{No closed form expression}$$

To Deal with High Dimensionnality

need for quick and easy access to the operators

Conditions

Interesting Properties

F convex smooth

G convex "proximable"

$$\operatorname{prox}_G(x) \stackrel{\text{def}}{=} \underset{y}{\operatorname{argmin}} \frac{1}{2}\|x-y\|^2 + G(y)$$

$$G(x) = \lambda \sum_k \sum_{b \in \mathcal{B}_k} \sqrt{\sum_i x_i^2}$$

$$\Rightarrow \quad \operatorname{prox}_G(x) = \quad \text{No closed form expression}$$

To Deal with High Dimensionnality

need for quick and easy access to the operators

avoid nested algorithms

State-of-the-Art

Forward-Backward $F(x) + G(x)$

$$\boxed{x \leftarrow \text{prox}_{\gamma G}\left(x - \gamma \nabla F(x)\right)}$$

Peaceman-Rachford $\sum_k G_k(x)$

Preconditionned ADMM $\sum_k G_k(L_k x)$

State-of-the-Art

Forward-Backward $F(x) + G(x)$

$$\boxed{x \leftarrow \text{prox}_{\gamma G}\left(x - \gamma \nabla F(x)\right)}$$

Peaceman-Rachford $\sum_k G_k(x)$

$$F(x) = \frac{1}{2}\|Y - Lx\|^2 \Rightarrow \text{prox}_F(x) = \left(\text{Id} + L^\star L\right)^{-1}\left(x + L^\star Y\right);$$

Preconditionned ADMM $\sum_k G_k(L_k x)$

State-of-the-Art

Forward-Backward $F(x) + G(x)$

$$\boxed{x \leftarrow \mathrm{prox}_{\gamma G}\left(x - \gamma \nabla F(x)\right)}$$

Peaceman-Rachford $\sum_k G_k(x)$

$$F(x) = \frac{1}{2}\|Y - Lx\|^2 \Rightarrow \mathrm{prox}_F(x) = (\mathrm{Id} + L^\star L)^{-1}(x + L^\star Y);$$

Preconditionned ADMM $\sum_k G_k(L_k x)$

Using ∇F ?

DAUPHINE UNIVERSITÉ PARIS NatImages

State-of-the-Art

Forward-Backward $F(x) + G(x)$

$$x \leftarrow \text{prox}_{\gamma G}\left(x - \gamma \nabla F(x)\right)$$

Peaceman-Rachford $\sum_k G_k(x)$

$$F(x) = \frac{1}{2}\|Y - Lx\|^2 \Rightarrow \text{prox}_F(x) = \left(\text{Id} + L^\star L\right)^{-1}\left(x + L^\star Y\right);$$

Preconditionned ADMM $\sum_k G_k(L_k x)$

Using ∇F ?

prox_F not necessary available

State-of-the-Art

Forward-Backward $F(x) + G(x)$

$$\boxed{x \leftarrow \text{prox}_{\gamma G}\left(x - \gamma \nabla F(x)\right)}$$

Peaceman-Rachford $\sum_k G_k(x)$

$$F(x) = \frac{1}{2}\|Y - Lx\|^2 \Rightarrow \text{prox}_F(x) = \left(\text{Id} + L^\star L\right)^{-1}(x + L^\star Y);$$

Preconditionned ADMM $\sum_k G_k(L_k x)$

Using ∇F ?

- prox_F not necessary available
- convergence speed and accelarated schemes

DAUPHINE UNIVERSITÉ PARIS NatImages

Extended Forward-Backward

Algorithm for $\min_x F(x) + \sum_{k=1}^{n} G_k(x)$

$$\forall\, k \in [\![1, n]\!],\; z_k \leftarrow z_k + \left(\mathrm{prox}_{\gamma G_k} \left(2x - z_k - \frac{\gamma}{n} \nabla F(x)\right) - x \right)$$

$$x \leftarrow \frac{1}{n} \sum_k z_k$$

DAUPHINE UNIVERSITÉ PARIS NatImages

Extended Forward-Backward

Algorithm for $\min_x F(x) + \sum_{k=1}^n G_k(x)$

$$\forall\, k \in [\![1, n]\!],\ z_k \leftarrow z_k + \left(\operatorname{prox}_{\gamma G_k} \left(2x - z_k - \frac{\gamma}{n} \nabla F(x)\right) - x \right)$$

$$x \leftarrow \frac{1}{n} \sum_k z_k$$

Theorem: Convergence with Robustness to Errors

if	*then*
∇F $1/\beta$-Lipschitz continuous; $\gamma \in]0, 2n\beta[$; Set of minimizers non-empty;	x converges towards a minimizer.

DAUPHINE UNIVERSITÉ PARIS NatImages

Extended Forward-Backward

Algorithm for $\min_x F(x) + \sum_{k=1}^{n} G_k(x)$

$$\forall\, k \in [\![1, n]\!],\; z_k \leftarrow z_k + \left(\operatorname{prox}_{\gamma G_k} \left(2x - z_k - \frac{\gamma}{n} \nabla F(x) + \epsilon_0 \right) + \epsilon_k - x \right)$$

$$x \leftarrow \frac{1}{n} \sum_k z_k$$

Theorem: Convergence with Robustness to Errors

if

- ∇F $1/\beta$-Lipschitz continuous;
- $\gamma \in]0, 2n\beta[$;
- Set of minimizers non-empty;
- $\forall\, k,\; \sum_{\text{iter}}^{\infty} \|\epsilon_k\| < \infty$.

then

x converges towards a minimizer.

DAUPHINE UNIVERSITÉ PARIS

NatImages

Extended Forward-Backward

Algorithm for $\min_x F(x) + \sum_{k=1}^{n} G_k(x)$

$$\forall\, k \in [\![1, n]\!],\ z_k \leftarrow z_k + \left(\operatorname{prox}_{\gamma G_k} \left(2x - z_k - \frac{\gamma}{n} \nabla F(x)\right) - x \right)$$

$$x \leftarrow \frac{1}{n} \sum_k z_k$$

Hybrid Forward-Backward/Peaceman-Rachford

If $n \equiv 1$, $x = z = z_1$,

DAUPHINE UNIVERSITÉ PARIS NatImages

Extended Forward-Backward

Algorithm for $\min_x F(x) + \sum_{k=1}^n G_k(x)$

$$\forall k \in [\![1, n]\!],\ z_k \leftarrow z_k + \left(\operatorname{prox}_{\gamma G_k} \left(2x - z_k - \frac{\gamma}{n} \nabla F(x)\right) - x \right)$$

$$x \leftarrow \frac{1}{n} \sum_k z_k$$

Hybrid Forward-Backward/Peaceman-Rachford

If $n \equiv 1$, $x = z = z_1$, this is forward-backward.

Extended Forward-Backward

Algorithm for $\min_x F(x) + \sum_{k=1}^n G_k(x)$

$$\forall\, k \in [\![1, n]\!],\ z_k \leftarrow z_k + \left(\mathrm{prox}_{\gamma G_k} \left(2x - z_k - \frac{\gamma}{n}\nabla F(x)\right) - x \right)$$

$$x \leftarrow \frac{1}{n} \sum_k z_k$$

Hybrid Forward-Backward/Peaceman-Rachford

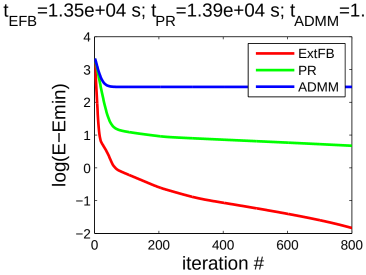
If $n \equiv 1$, $x = z = z_1$, this is forward-backward.

If $F \equiv 0$, this is a special case of Peaceman-Rachford.

DAUPHINE UNIVERSITÉ PARIS NatImages

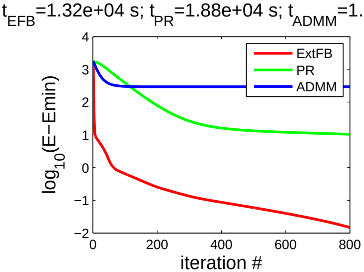
Numerical Experiments

Deconvolution $\min_x \frac{1}{2}\|Y - K * \Psi x\| + \sum_{k=1}^{16} \|x\|_{1,2}^{\mathcal{B}_k}$



Numerical Experiments

Deconvolution + Inpainting $\min_x \frac{1}{2}\|Y - P_\Omega K * \Psi x\| + \sum_{k=1}^{16} \|x\|_{1,2}^{\mathcal{B}_k}$



Conclusion

Extended Forward-Backward

- Enlarges the class of convex optimization problems which can be solved in a "reasonable amount of time";
- Robust to errors.

Future Work

- Convergence speed;
- Accelarated scheme.

Aknowledgement

Gabriel Peyré, (CEREMADE)

for believing that such an extended forward-backward was possible

Jalal Fadili, (ENSICAEN)

for giving the impression to know everything about convex optimization

Nicolas Schmidt, (CEREMADE)

for mathematical rigor and friendship