Optimization problem

Jessy Khafif

September 7, 2023

Summary

Généralités

Présentation de la problématique

"Comment pouvons nous retrouver une donnée originale sachant qu'elle a été détérioré ?"

Vers un problème formel

$$\begin{aligned} g &= S(h \circledast f) + n \\ &= SHf + n \end{aligned}$$

Avec,

- g l'image observée avec une basse résolution
- S est un opérateur de décimation
- h un kernel de convolution
- H la matrice BCCB obtenue à partir du kernel h
- f l'image désirée avec une haute résolution
- n un bruit additif
- $\circledast$ l'opération de la convolution circulaire

Vers un problème formel

$$
\begin{aligned}
g &= S(h \circledast f) + n \\
&= SHf + n
\end{aligned}
$$

L'énoncé formalisé se lit comme :

g est le résultat de f ayant subit un flou "H" suivi d'une décimation "S", puis bruité par "n". L'objectif est donc de retrouver f.

Les problèmes inverses

- C'est un problème inverse
- Dans la littérature, ce genre de problème est abordé avec des modèles mathématiques
- Il existe des algorithmes itératifs permettant de résoudre des problèmes inverses en minimisant un modèle mathématique

Modèle mathématique générique

$$\hat{f} = \arg\min_{f} \{ \frac{\alpha}{2} \underbrace{\|g - SHf\|_2^2}_{\text{Attache aux données}} + \underbrace{\sum_{i \in \Omega} \beta_i \times R_i(f)}_{\text{combinaison de régularisation}} \}$$

Avec,

- g : la donnée observée avec détérioration
- S : un opérateur de décimation
- H : la matrice BCCB obtenue à partir du kernel h
- f : la donnée désirée sans détérioration
- α : un hyper-paramètre
- Ω : l'ensemble des indices des régularisations
- β_i : l'hyper-paramètre associé au terme de régularisation R_i
- R_i : régurisation d'indice i

Modèle mathématique générique

$$
\begin{aligned}
g &= S(h \circledast f) + n \\
&= SHf + n \iff n = g - SHf
\end{aligned}
$$

$$\hat{f} = \arg\min_f \{ \frac{\alpha}{2} \|g - SHf\|_2^2 + \sum_{i \in \Omega} \beta_i \times R_i(f) \}$$

Modèle mathématique générique

$$\hat{f} = \arg\min_{f}\{\frac{\alpha}{2}\|g - SHf\|_2^2 + \sum_{i\in\Omega}\beta_i \times R_i(f)\}$$

Dans un langage plus naturel :

Nous recherchons f tel qu'il minimise le bruit n sachant que nous avons un ensemble de connaissances/régularisations certaines sur f.

Tikhonov : $R(x) = \|x\|_2^2$

$$\hat{x} = \arg\min_x \{\frac{\alpha}{2}\|y - Hx\|_2^2 + \beta_0\|x\|_2^2\}$$

$$\frac{d}{dx}\{\frac{\alpha}{2}\|y - Hx\|_2^2 + \beta_0\|x\|_2^2\} = 0$$

$$\iff -\alpha H^T(y - Hx) + 2\beta_0 x = 0$$

$$\iff -\alpha H^T y + \alpha H^T H x + 2\beta_0 x = 0$$

$$\iff -\alpha H^T y + (\alpha H^T H + 2\beta_0 I)x = 0$$

$$\iff x = (\alpha H^T H + 2\beta_0 I)^{-1}\alpha H^T y$$

ADMM

Definition

Résolution avec l'ADMM:

- Équation : $\min_{x,z}\{f(x)+g(z)\}$
- Contrainte : $Ax+Bz=c$

Lagrangien augmenté :

$$\begin{aligned} L(x,z,u) &= f(x)+g(z)+u^T(Ax+Bz-c)+\frac{\rho}{2}\|Ax+Bz-c\|_2^2 \\ &= f(x)+g(z)+\frac{\rho}{2}\|Ax+Bz-c+\frac{u}{\rho}\|_2^2-\frac{\|u\|_2^2}{2\rho} \end{aligned}$$

Algorithm

Algorithm 1 ADMM

1: **procedure** ADMM

Input: y; α; $\forall i \in \Omega, \beta_i$, ρ, $nb\_iterations$

2: **Output:** $x_{nb\_iterations}$

3: **for** $k \in 0...nb\_iterations$ **do**

4: $x_{k+1} = \arg\min_x L(x_k, z_k, u_k)$

5: $z_{k+1} = \arg\min_z L(x_{k+1}, z, u_k)$

6: $u_{k+1} = u_k + \rho(z_{k+1} - x_{k+1})$

7: **end for**

8: **end procedure**

$L_1 : R(x) = \|x\|_1$

Résolution avec l'ADMM:

- Équation : $\min\limits_{x,z}\{\frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1\}$
- Contrainte : $x = z$

Lagrangien augmenté :

$$\begin{aligned} L(x,z,u) &= \frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + u^T(x-z) + \frac{\rho}{2}\|x-z\|_2^2 \\ &= \frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + \frac{\rho}{2}\|x - z + \frac{u}{\rho}\|_2^2 - \frac{\|u\|_2^2}{2\rho} \end{aligned}$$

$L_1 : R(x) = \|x\|_1$

$$\frac{d}{dx}\{\|y - Hx\|_2^2\} = -2H^T(y - Hx) = -2H^T y + 2H^T Hx$$
$$\frac{d}{dx}\{\|Hx - y\|_2^2\} = 2H^T(Hx - y) = 2H^T Hx - 2H^T y$$

$\arg\min_x L(x, z, u)$

$$\frac{d}{dx}\{\frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + \frac{\rho}{2}\|x - z + \frac{u}{\rho}\|_2^2 - \frac{\|u\|_2^2}{2\rho}\} = 0$$

$$\iff -H^T(y - Hx) + \rho(x - z + \frac{u}{\rho}) = 0$$

$$\iff -H^T y + H^T Hx + \rho x - \rho z + u = 0$$

$$\iff -H^T y + (H^T H + \rho I)x - \rho z + u = 0$$

$$\iff (H^T H + \rho I)x = H^T y + \rho z - u$$

$$\iff x = (H^T H + \rho I)^{-1}(H^T y + \rho z - u)$$

$\arg\min_{z} L(x, z, u)$

$$\frac{d}{dz}\{\frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + \frac{\rho}{2}\|x - z + \frac{u}{\rho}\|_2^2 - \frac{\|u\|_2^2}{2\rho}\} = 0$$

$$\iff \lambda sign(z) - \rho(x - z + \frac{u}{\rho}) = 0$$

$$\iff \lambda sign(z) - \rho x + \rho z - u = 0$$

$$\iff \lambda sign(z) + \rho z = \rho x + u$$

$$\iff z + \frac{\lambda}{\rho} sign(z) = x + \frac{u}{\rho}$$

$$\iff z = sign(x + \frac{u}{\rho}) \times max(|x + \frac{u}{\rho}| - \frac{\lambda}{\rho}, 0)$$

$$\iff z = Soft_{\frac{\lambda}{\rho}}(x + \frac{u}{\rho})$$

Algorithm

Algorithm 2 ADMM-L_1

procedure ADMM

Input: y, λ, ρ, $nb\_iterations$

Output: $x_{nb\_iterations}$

for $k \in 0...nb\_iterations$ **do**

$x_{k+1} = (H^T H + \rho I)^{-1}(H^T y_k + \rho z_k - u_k)$

$z_{k+1} = Soft_{\frac{\lambda}{\rho}}(x_{k+1} + \frac{u_k}{\rho})$

$u_{k+1} = u_k + \rho(z_{k+1} - x_{k+1})$

end for

end procedure

Total Variation : $R(x) = \|\nabla x\|_1$

Résolution avec l'ADMM:

- Équation : $\min\limits_{x,z}\{\frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1\}$
- Contrainte : $\nabla x = z$

Lagrangien augmenté :

$$\begin{aligned} L(x,z,u) &= \frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + u^T(\nabla x - z) + \frac{\rho}{2}\|\nabla x - z\|_2^2 \\ &= \frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + \frac{\rho}{2}\|\nabla x - z + \frac{u}{\rho}\|_2^2 - \frac{\|u\|_2^2}{2\rho} \end{aligned}$$

$\arg\min_x L(x, z, u)$

$$\frac{d}{dx}\{\frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + \frac{\rho}{2}\|\nabla x - z + \frac{u}{\rho}\|_2^2 - \frac{\|u\|_2^2}{2\rho}\} = 0$$

$$\iff -H^T(y - Hx) + \rho\nabla^T(\nabla x - z + \frac{u}{\rho}) = 0$$

$$\iff -H^Ty + H^THx + \rho\nabla^T\nabla x - \rho\nabla^T z + \nabla^T u = 0$$

$$\iff -H^Ty + (H^TH + \rho\nabla^T\nabla)x - \rho\nabla^T z + \nabla^T u = 0$$

$$\iff (H^TH + \rho\nabla^T\nabla)x = \rho\nabla^T z - \nabla^T u + H^Ty$$

$$\iff (H^TH + \rho\nabla^T\nabla)x = \rho\nabla^T(z - \frac{u}{\rho}) + H^Ty$$

$$\iff x = (H^TH + \rho\nabla^T\nabla)^{-1}[\rho\nabla^T(z - \frac{u}{\rho}) + H^Ty]$$

$\arg\min_z L(x, z, u)$

$$
\begin{aligned}
&\frac{d}{dz}\{\frac{1}{2}\|y - Hx\|_2^2 + \lambda\|z\|_1 + \frac{\rho}{2}\|\nabla x - z + \frac{u}{\rho}\|_2^2 - \frac{\|u\|_2^2}{2\rho}\} = 0 \\
&\iff \lambda sign(z) - \rho(\nabla x - z + \frac{u}{\rho}) = 0 \\
&\iff \lambda sign(z) - \rho\nabla x + \rho z - u = 0 \\
&\iff \lambda sign(z) + \rho z = \rho\nabla x + u \\
&\iff z + \frac{\lambda}{\rho} sign(z) = \nabla x + \frac{u}{\rho} \\
&\iff z = sign(\nabla x + \frac{u}{\rho}) \times max(|\nabla x + \frac{u}{\rho}| - \frac{\lambda}{\rho}, 0) \\
&\iff z = Soft_{\frac{\lambda}{\rho}}(\nabla x + \frac{u}{\rho})
\end{aligned}
$$

Algorithm

Algorithm 3 ADMM-L_1

procedure ADMM

Input: y, λ, ρ, $nb\_iterations$

Output: $x_{nb\_iterations}$

for $k \in 0 ... nb\_iterations$ **do**

$$x_{k+1} = (H^T H + \rho \nabla^T \nabla)^{-1} [\rho \nabla^T (z_k - \frac{u_k}{\rho}) + H^T y]$$

$$z_{k+1} = Soft_{\frac{\lambda}{\rho}}(\nabla x_{k+1} + \frac{u_k}{\rho})$$

$$u_{k+1} = u_k + \rho(z_{k+1} - x_{k+1})$$

end for

end procedure

Split Bregman

Definition

Problème de minimisation

$$\hat{f} = \arg\min_{f}\{\frac{1}{2}\|g - Hf\|_2^2 + \lambda\|\nabla f\|_1\}$$

Réécriture de $\|\nabla f\|_1$

En posant,

$$\begin{cases} d_x = \nabla_x f = \dfrac{df}{dx} \\ d_y = \nabla_y f = \dfrac{df}{dy} \\ \|\nabla f\|_1 = \|(d_x, d_y)\|_1 = \sum_i \sqrt{(d_x)_i^2 + (d_y)_i^2} \end{cases}$$

nous avons:

$$\hat{f} = \arg\min_f \{\frac{1}{2}\|g - Hf\|_2^2 + \lambda\|(d_x, d_y)\|_1\}$$

Réécriture de $\|\nabla f\|_1$ (Relaxation)

nous avons,

$$(f, d_x, d_y) = \arg\min_{f,d_x,d_y} \{\frac{1}{2}\|g - Hf\|_2^2 + \lambda\|(d_x, d_y)\|_1 + \frac{\sigma}{2}\|d_x - \nabla_x f\|_2^2 + \frac{\sigma}{2}\|d_y - \nabla_y f\|_2^2\}$$

Introduction du processus itératif

$$(f^{k+1}, d_x^{k+1}, d_y^{k+1}) = \arg\min_{f,d_x,d_y} \{\frac{1}{2}\|g - Hf^k\|_2^2 \lambda \|(d_x^k, d_y^k)\|_1 + \frac{\sigma}{2}\|d_x^k - \nabla_x f^k - b_x^k\|_2^2 + \frac{\sigma}{2}\|d_y^k - \nabla_y f^k - b_y^k\|_2^2\}$$

Avec,

$$\begin{cases} b_x^{k+1} = b_x^k + (\nabla_x f^{k+1} - d_x^{k+1}) \\ b_y^{k+1} = b_y^k + (\nabla_y f^{k+1} - d_y^{k+1}) \end{cases}$$

Division en deux sous-problème

1. Calcul de f
2. Calcul de (d_x, d_y)

Calcul de f

$$
\begin{aligned}
f^{k+1} &= \arg\min_f \{ \frac{1}{2} \|g - Hf^k\|_2^2 + \lambda \|(d_x^k, d_y^k)\|_1 + \frac{\sigma}{2} \|d_x^k - \nabla_x f^k - b_x^k\|_2^2 + \frac{\sigma}{2} \|d_y^k - \nabla_y f^k - b_y^k\|_2^2 \} \\
&= \arg\min_f \{ \frac{1}{2} \|g - SHf^k\|_2^2 + \frac{\sigma}{2} \|d_x^k - \nabla_x f^k - b_x^k\|_2^2 + \frac{\sigma}{2} \|d_y^k - \nabla_y f^k - b_y^k\|_2^2 \}
\end{aligned}
$$

Calcul de f

$$\frac{d}{df}\{\frac{1}{2}\|g - Hf\|_2^2 + \frac{\sigma}{2}\|d_x - \nabla_x f - b_x\|_2^2 + \frac{\sigma}{2}\|d_y - \nabla_y f - b_y\|_2^2\} = 0$$

$$\Longrightarrow \ -H^T(g - Hf) - \sigma\nabla_x^T(d_x - \nabla_x f - b_x) - \sigma\nabla_y^T(d_y - \nabla_y f - b_y) = 0$$
$$\Longleftrightarrow \ -H^T g + H^T H f - \sigma\nabla_x^T d_x + \sigma\nabla_x^T\nabla_x f + \sigma\nabla_x^T b_x - \sigma\nabla_y^T d_y + \sigma\nabla_y^T\nabla_y f + \sigma\nabla_y^T b_y = 0$$
$$\Longleftrightarrow \ H^T H f + \sigma\nabla_x^T\nabla_x f + \sigma\nabla_y^T\nabla_y f = \sigma\nabla_x^T d_x - \sigma\nabla_x^T b_x + \sigma\nabla_y^T d_y - \sigma\nabla_y^T b_y + H^T g$$

$$\begin{aligned} &H^T H f + \sigma\nabla_x^T\nabla_x f + \sigma\nabla_y^T\nabla_y f \\ &= [H^T H + \sigma(\nabla_x^T\nabla_x + \nabla_y^T\nabla_y)]f \\ &= [H^T H + \sigma\Delta]f \end{aligned}$$

$$\begin{aligned} &\sigma\nabla_x^T d_x - \sigma\nabla_x^T b_x + \sigma\nabla_y^T d_y - \sigma\nabla_y^T b_y + H^T g \\ &= \sigma\nabla_x^T(d_x - b_x) + \sigma\nabla_y^T(d_y - b_y) + H^T g \\ &= \sigma[\nabla_x^T(d_x - b_x) + \nabla_y^T(d_y - b_y)] + H^T g \end{aligned}$$

Calcul de f

Ainsi, dans le domaine spatial, nous avons :

$$f = [H^T H + \sigma\Delta]^{-1}[\sigma(\nabla_x^T(d_x - b_x) + \nabla_y^T(d_y - b_y)) + H^T g]$$

Calcul de (d_x, d_y)

$$
\begin{aligned}
(d_x^{k+1}, d_y^{k+1}) &= \arg\min_{d_x,d_y}\{\frac{1}{2}\|g - Hf^k\|_2^2 + \lambda\|(d_x^k, d_y^k)\|_1 + \frac{\sigma}{2}\|d_x^k - \nabla_x f^k - b_x^k\|_2^2 + \frac{\sigma}{2}\|d_y^k - \nabla_y f^k - b_y^k\|_2^2\} \\
&= \arg\min_{d_x,d_y}\{\lambda\|(d_x^k, d_y^k)\|_1 + \frac{\sigma}{2}\|d_x^k - \nabla_x f^k - b_x^k\|_2^2 + \frac{\sigma}{2}\|d_y^k - \nabla_y f^k - b_y^k\|_2^2\} \\
&= \arg\min_{d_x,d_y}\{\lambda\|(d_x^k, d_y^k)\|_1 + \frac{\sigma}{2}\|d_x^k - (\nabla_x f^k + b_x^k)\|_2^2 + \frac{\sigma}{2}\|d_y^k - (\nabla_y f^k + b_y^k)\|_2^2\}
\end{aligned}
$$

$$
\begin{cases}
d_x^{k+1} = \max(s^k - \dfrac{\lambda}{\sigma}, 0)\dfrac{s_x^k}{s^k} \\
d_y^{k+1} = \max(s^k - \dfrac{\lambda}{\sigma}, 0)\dfrac{s_y^k}{s^k}
\end{cases}
$$

Avec,

$$
\begin{cases}
s_x^k = \nabla_x f^k + b_x^k \\
s_y^k = \nabla_y f^k + b_y^k \\
s^k = \sqrt{(s_x^k)^2 + (s_y^k)^2}
\end{cases}
$$

The End