

Exploration de différentes pénalisations sur la fonction coût utilisée pour optimiser les niveaux de gris.

1.1 Régularisation de Moreau-Yosida

1.1.1 Présentation de la régularisation et propriétés.

Définition 1.1.1. Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction. La transformée (ou régularisée) de Moreau-Yosida de J est définie à l'aide d'un paramètre $c > 0$ par

$$J_c(u) := \min_{v \in \mathbb{R}^n} \left(J(v) + \frac{1}{2c} \|v - u\|^2 \right). \quad (1.1)$$

En prenant $v = u$ dans le minimum, on obtiens que :

Propriété 1.1.2. $\forall c > 0, \quad \forall u \in \mathbb{R}^n, \quad J_c(u) \leq J(u).$

Si J est une fonction à minimiser, alors la propriété suivante est utile :

Propriété 1.1.3. Soit a le minimum de J . Soit u l'un des antécédants de a . Alors u minimise aussi J_c et on a $J_c(u) = J(u) = a$ quelque soit $c > 0$.

Démonstration. Notons $f_u(v) := J(v) + \frac{1}{2c} \|v - u\|^2$.

\Rightarrow) Soit u_0 , un point en lequel J atteint son minimum.

On a alors $f_{u_0}(u_0) = J(u_0) \leq J(u) \leq J(u) + \frac{1}{2c} \|u - u_0\|^2 = f_{u_0}(u), \quad \forall u,$

d'où $J_c(u_0) = \min_u f_{u_0}(u) = J(u_0)$. On a donc égalité des fonctions J et J_c en u_0 . Reste à montrer que u_0 est bien un point de minimum pour J_c .

Comme u_0 minimise J , on a $J(u_0) \leq J(v), \quad \forall v.$

d'où $J(u_0) + \frac{1}{2c} \|v - u\|^2 \leq J(v) + \frac{1}{2c} \|v - u\|^2, \quad \forall u, \forall v.$

Ainsi $\min_v (J(u_0) + \frac{1}{2c} \|v - u\|^2) \leq \min_v (J(v) + \frac{1}{2c} \|v - u\|^2) = J_c(u) \quad \forall u.$

Or $\min_v (J(u_0) + \frac{1}{2c} \|v - u\|^2) = J(u_0)$ et $J(u_0) = J_c(u_0)$

donc $J_c(u_0) \leq J(u), \quad \forall u.$ et ainsi u_0 minimise J_c .

\Leftarrow) Supposons que u_0 minimise J_c . On a alors

$$J_c(u_0) \leq J_c(u) = \min_v f_u(v) \leq f_u(u) = J(u) \quad \forall u.$$

Mais $J_c(u_0) = J(u_0)$. Ainsi $J(u_0) \leq J(u)$, $\forall u$ et donc u_0 minimise J . ■

Ainsi minimiser la fonction J est équivalent à minimiser la fonction J_c . L'avantage de la fonction J_c , c'est qu'elle est construite de sorte à être différentiable aux abords du point de minimum même si J ne l'est pas (J peut même être discontinue). Le principal désavantage, c'est que le problème $\min_u J_c(u)$ est infiniment plus complexe à résoudre que le problème $\min_u J(u)$ puisque la fonction J_c nécessite le calcul d'un minimum à chaque évaluation. Qu'à cela ne tienne. Dans notre cas, nous connaissons explicitement J . La transformée de Moreau-Yosida n'est appliquée que pour gagner en régularité. Ainsi, on peut calculer explicitement J_c . Avant de fournir un exemple, remarquons que

Propriété 1.1.4. Si la fonction J est paire, alors sa régularisée J_c l'est aussi.

Démonstration.

$$\begin{aligned} J_c(-u) &= \min_{v \in \mathbb{R}} \left(J(v) + \frac{1}{2c} \|v + u\|^2 \right) \\ &= \min_{s \in \mathbb{R}} \left(J(-s) + \frac{1}{2c} \|-s + u\|^2 \right) \quad \text{en posant } s = -v \\ &= \min_{s \in \mathbb{R}} \left(J(s) + \frac{1}{2c} \|s - u\|^2 \right) = J_c(u) \quad (\text{car } J \text{ et } \|\cdot\|^2 \text{ sont paires}). \end{aligned}$$

■

Ceci va nous permettre d'alléger les calculs. Regardons en exemple, le calcul de la régularisée d'une fonction créneau.

Exemple : Considérons la fonction créneau :

$$J(u) = \begin{cases} 0 & \text{si } u \in [-a; a], \\ 1 & \text{sinon.} \end{cases}$$

On pose $f_u(v) = J(v) + \|v - u\|^2/2c$. On a $f'_u(v) = (v - u)/c$, $\forall u \in \mathbb{R} \setminus \{-a, a\}$. En $\pm a$, la dérivée n'est pas définie. Puisque J est paire, nous pouvons limiter l'étude de sa régularisée aux u positifs. La fonction étant discontinue en a , il faut distinguer 2 cas :

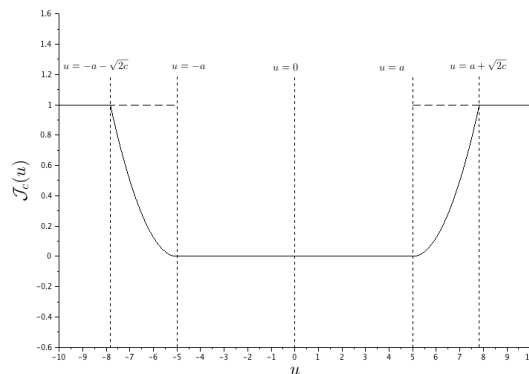


FIGURE 1.1 – Régularisation de Moreau-Yosida d'une fonction créneau avec $c = 4$.

- Cas 1 : $a < u$

v	0	a	u	$+\infty$
$f'_u(v)$	—	—	0	+
$f_u(v)$	$u^2/2c$	$(a-u)^2/2c$	1	$+\infty$

Ici, $u \leq a \Rightarrow \min_v f_u(v) = \min(1, (a-u)^2/2c)$. Ainsi J_c vaut 1 si $(a-u)^2/2c \geq 1$, i.e. si $u \leq a + \sqrt{2c}$; si $u > a + \sqrt{2c}$ alors $J_c(u) = (a-u)^2/2c$.

- Cas 2 : $a \geq u$. Ici tout les points u tel que $0 \leq u \leq a$ minimise J . Donc $J_c(u) = J(u) = 0$ sur cet intervalle. On peut s'en rassurer avec le tableau de variation suivant :

v	0	u	a	$+\infty$
$f'_u(v)$	—	0	—	+
$f_u(v)$	$u^2/2c$	0	$(a-u)^2/2c$	$+\infty$

qui nous donne bien $J_c(u) = \min_v f_u(v) = 0$.

Bilan : On a une expression explicite de la régularisée (que l'on complète par parité de la fonction) :

$$J_c(u) = \begin{cases} 1 & \text{si } u \in]-\infty; -a - \sqrt{2c}] \cup [a + \sqrt{2c}; +\infty[\\ (a-u)^2/2c & \text{si } u \in]-a - \sqrt{2c}; -a[\cup]a; a + \sqrt{2c}[\\ 0 & \text{si } u \in [-a; a] \end{cases}$$

L'allure de la fonction créneau J et de sa régularisée J_c , sont présentées Figure 1.1.

Notons que la fonction n'est régulière qu'aux abords des minimums. La valeur 0 est raccordé de manière dérivable, mais pas de dérivabilité pour le raccord en 1.

Remarquons également que plus le paramètre c est grand, plus l'intervalle sur lequel agit la régularisation est large.

On pourrait aller beaucoup plus loin dans l'étude des propriétés de cette régularisation. De nombreuses publications ont été faite notamment dans des cas où l'on ne connaît pas analytiquement la fonction J et où l'on s'intéresse notamment au problème adjoint... Nous ne nous étalerons pas plus sur ce vaste sujet : là n'est pas l'objet de ce manuscrit.

1.1.2 Régularisation de Moreau-Yosida appliquée à une parabole tronquée.

On considère dans cette section, la parabole tronquée suivante

$$J(u) = [u^2 - a^2]^+ \quad (1.2)$$

où a^2 est le minimum de la parabole d'origine, et où $[x]^+ = \max(x, 0)$. De la même manière que dans l'exemple de la section précédente, nous allons construire explicitement la régularisée de cette fonction. On a ici :

$$f_u(v) = [v^2 - a^2]^+ + \frac{1}{2c} \|v - u\|^2. \quad (1.3)$$

De plus, comme la fonction J est paire, on restreint l'étude à l'ensemble \mathbb{R}_+ . Ainsi la dérivée est caractérisée sur \mathbb{R}_+ par

$$f'_u(v) = \begin{cases} \frac{v - u}{c} & \text{si } v < a, \\ \text{non définie} & \text{si } v = a, \\ 2v + \frac{v - u}{c} & \text{si } v > a. \end{cases} \quad (1.4)$$

Ainsi

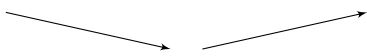
$$f'_u(v) = 0 \iff \begin{cases} v = u & \text{si } v < a, \\ v = \frac{u}{2c + 1} & \text{si } v > a. \end{cases} \quad (1.5)$$

Les cas suivants sont donc à considérer (remarquer que comme $c > 0$ et $u > 0$ on a toujours $\frac{u}{2c+1} < u$) :

- Cas 1 : $0 < \frac{u}{2c+1} < u \leq a$

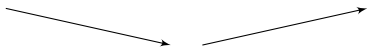
Dans ce cas là u minimise J . Donc $J_c(u) = J(u) = 0, \forall u \leq a$.

- Cas 2 : $0 < \frac{u}{2c+1} < a < u$

v	0	$\frac{u}{2c+1}$	a	u	$+\infty$
$f'_u(v)$	—	—		+	+
$f_u(v)$					

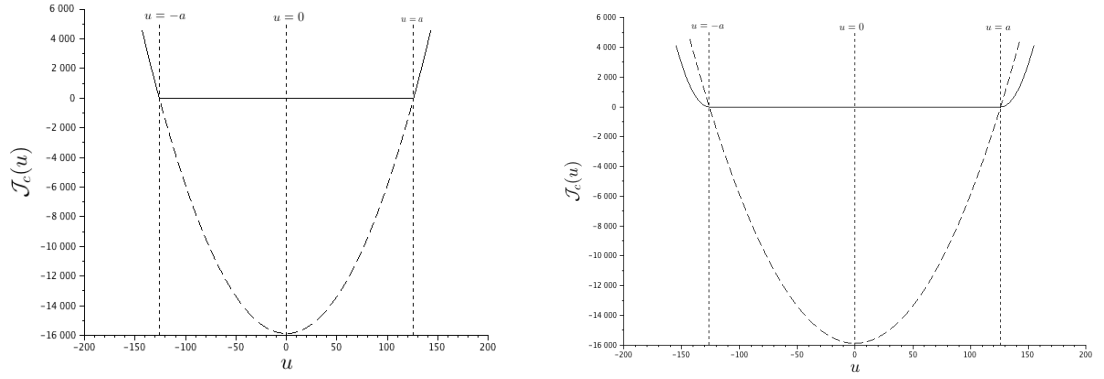
Ici, $\min_v f_u(v) = f(a) = \frac{1}{2c} \|a - u\|^2, \forall u \in]a, a(2c + 1)[$.

- Cas 2 : $0 < a < \frac{u}{2c+1} < u$

v	0	a	$\frac{u}{2c+1}$	u	$+\infty$
$f'_u(v)$	—		— 0	+	+
$f_u(v)$					

$$\begin{aligned} \text{Ici, } \min_v f_u(v) &= f\left(\frac{u}{2c+1}\right) = \left[\left(\frac{u}{2c+1}\right)^2 - a^2\right]^+ + \frac{1}{2c} \left\|\frac{u}{2c+1} - u\right\|^2 \\ &= \left(\frac{u}{2c+1}\right)^2 - a^2 + \frac{1}{2c} \left(\frac{-2cu}{2c+1}\right)^2 \\ &= \frac{u^2}{2c+1} - a^2, \quad \forall u \in]a(2c+1), +\infty[. \end{aligned}$$

1.2. Impact du choix de la pénalisation sur l'optimisation des niveaux de gris



(a) Parabole tronquée
 $J(u) = [u^2 - a^2]^+$ (avec $a = 127.5$).

(b) Régularisée de Moreau-Yosida d'une parabole tronquée ($c = 0.1$).

FIGURE 1.2 – Exemple d'une régularisation de Moreau-Yosida.

Ainsi l'expression de la régularisée de la parabole tronquée 1.2 est donnée par

$$J_c(u) = \begin{cases} \frac{u^2}{1+2c} - a^2 & \text{si } u \in]-\infty; -a(2c+1)] \cup [a(2c+1); +\infty[\\ \frac{1}{2c} \|a - u\|^2 & \text{si } u \in]-a(2c+1); -a[\cup]a; a(2c+1)[\\ 0 & \text{si } u \in [-a; a], \end{cases} \quad (1.6)$$

et son allure est présentée sur la Figure 1.2b.

1.2 Impact du choix de la pénalisation sur l'optimisation des niveaux de gris

Comme montré dans le chapitre ?? (cf. notamment Table ?? et surtout dans la Table ??), les algorithmes d'optimisations sur les niveaux de gris tendent, dans certaines configurations, vers des jeux de paramètres optimaux qui s'approchent du bord 0 ou du bord 255, voire même qui sont négatifs (*i.e.* non convergence de l'algorithme d'optimisation). Ce phénomène pourrait être dû notamment au fait que la pénalisation en créneau (??) présente une discontinuité. Les algorithmes de descente notamment fonctionnant sur une approximation du gradient peuvent ainsi être perturbé par cette discontinuité. Examinons ceci de plus près. Essayons une pénalisation continue : une parabole coupée

$$\mathcal{P}(\tau) = \left[(E(\tau) - 127.5)^2 - 126^2 \right]^+, \quad (1.7)$$

où $[\cdot]^+ = \max(0, \cdot)$ désigne la partie positive et où $E(\tau)$ est la composante de τ la plus éloignée du centre de l'intervalle autorisé (127.5 milieu de $[0; 255]$) :

$$E(\tau) = \arg \max_{i=1,2,3} (|\tau_i - 127.5|). \quad (1.8)$$

Scanners choisis pour l'optimisation	Algorithme d'optimisation			
	SLSQP	GC	Nelder-Mead	BFGS
	τ_N, τ_P	τ_N, τ_P	τ_N, τ_P	τ_N, τ_P
[2, 3]	25.49, 150.89 Err : 1.9e-02	31.85, 146.9 Err : 1.9e-02	253.5, 7.25 Err : 1.4e-02	243.58, 12.08 Err : 1.6e-02
[2, 3, 4]	23.64, 147.84 Err : 4.4e-02	32.94, 142.01 Err : 4.3e-02	253.5, 3.7 Err : 3.6e-02	251.64, 7.61 Err : 4.0e-02
[2, 3, 5]	22.75, 146.5 Err : 6.9e-02	27.06, 143.79 Err : 6.9e-02	253.3, 1.5 Err : 6.6e-02	223.15, 18.18 Err : 6.7e-02
[2, 3, 4, 5]	21.9, 145.05 Err : 6.5e-02	27.6, 141.48 Err : 6.5e-02	250.96, 1.5 Err : 6.0e-02	240.24, 10.71 Err : 6.3e-02
[2, 5, 7]	15.96, 135.58 Err : 2.0e-01	24.3, 130.33 Err : 2.0e-01	229.3, 1.5 Err : 2.0e-01	235.36, 2.59 Err : 2.0e-01
[1, 7, 11]	1.5, 139.07 Err : 1.1e-01	1.37, 139.29 Err : 3.3e+01	1.5, 139.51 Err : 1.1e-01	6.14, 132.04 Err : 1.2e-01
[1, 9, 11]	31.17, 131.88 Err : 2.7e-01	1.53, 133.91 Err : 2.6e-01	1.5, 134.01 Err : 2.6e-01	2.14, 140.56 Err : 2.7e-01
[1, 2, 7, 11]	28.41, 138.98 Err : 1.3e-01	6.75, 143.27 Err : 1.2e-01	6.3, 143.36 Err : 1.2e-01	6.13, 143.4 Err : 1.2e-01
[1, 3, 7, 11]	27.94, 137.75 Err : 1.2e-01	1.44, 142.73 Err : 1.5e+01	1.5, 142.8 Err : 1.1e-01	1.6, 145.64 Err : 1.1e-01
[3, 5, 7]	13.92, 132.24 Err : 1.9e-01	14.03, 132.18 Err : 1.9e-01	223.42, 1.5 Err : 1.9e-01	13.91, 132.23 Err : 1.9e-01
[1, 3, 5, 7]	23.41, 137.16 Err : 1.4e-01	1.33, 144.12 Err : 4.3e+01	1.5, 144.36 Err : 1.2e-01	3.0, 148.52 Err : 1.3e-01
[2, 3, 5, 7]	18.39, 139.46 Err : 1.7e-01	23.89, 136.02 Err : 1.7e-01	238.19, 1.5 Err : 1.7e-01	235.76, 6.99 Err : 1.7e-01
[1, 2, 5, 7]	24.17, 138.77 Err : 1.4e-01	1.53, 146.1 Err : 1.3e-01	1.5, 146.21 Err : 1.3e-01	2.29, 142.22 Err : 1.4e-01

TABLE 1.1 – Tableau récapitulatif des optimisations réalisées sur 2 niveaux de gris, τ_S fixé à 197, avec pénalisation quadratique (1.7) (*cf.* Figure 1.2a), pour les combinaisons défavorables de la Table ??, page ??.

1.2. Impact du choix de la pénalisation sur l'optimisation des niveaux de gris

L'aspect de cette pénalisation est présenté sur la Figure 1.2a (avec $a = 126; u = E(\tau) - 127.5$). Il s'agit d'une parabole dont on ignore la partie négative. Ici la pénalisation intervient sur un intervalle légèrement plus court que $[0; 255]$, car de toute façon les valeurs de τ n'ont pas à s'approcher de ces bornes (la pénalisation est non nulle en dehors de $[1.5; 253.5]$).

Les résultats des optimisations de niveaux de gris faites avec la pénalisation (1.7) sont présentés dans la Table 1.1. Ici plus de valeurs négatives, cependant la borne 1.5 est atteinte à plusieurs reprises. De même les résultats entre les algorithmes sont très variables. Ainsi, ici, il est difficile de dire s'il y a eu une amélioration ou non par rapport à la pénalisation en créneau. La pénalisation considérée ici n'est que \mathcal{C}^0 car il y a 2 points anguleux. Peut-être que cette régularité n'est pas suffisante encore.

Scanners choisis pour l'optimisation	Algorithme d'optimisation			
	SLSQP	GC	Nelder-Mead	BFGS
	τ_N, τ_P	τ_N, τ_P	τ_N, τ_P	τ_N, τ_P
[2, 3]	25.49, 150.89 Err : 1.9e-02	31.85, 146.9 Err : 1.9e-02	253.5, 7.25 Err : 1.4e-02	250.64, 7.62 Err : 1.6e-02
[2, 3, 4]	23.64, 147.84 Err : 4.4e-02	32.94, 142.01 Err : 4.3e-02	253.5, 3.7 Err : 3.6e-02	219.06, 27.85 Err : 4.0e-02
[2, 3, 5]	22.75, 146.5 Err : 6.9e-02	27.06, 143.79 Err : 6.9e-02	253.3, 1.5 Err : 6.6e-02	223.15, 18.18 Err : 6.7e-02
[2, 3, 4, 5]	21.9, 145.05 Err : 6.5e-02	27.6, 141.48 Err : 6.5e-02	250.96, 1.5 Err : 6.0e-02	218.18, 24.41 Err : 6.3e-02
[2, 5, 7]	15.96, 135.58 Err : 2.0e-01	24.3, 130.33 Err : 2.0e-01	229.3, 1.5 Err : 2.0e-01	220.88, 11.53 Err : 2.0e-01
[1, 7, 11]	1.5, 139.48 Err : 1.1e-01	1.37, 139.29 Err : 1.9e-01	1.5, 139.5 Err : 1.1e-01	6.14, 132.04 Err : 1.2e-01
[1, 9, 11]	31.17, 131.88 Err : 2.7e-01	1.5, 134.0 Err : 2.6e-01	1.5, 134.01 Err : 2.6e-01	2.14, 140.56 Err : 2.7e-01
[1, 2, 7, 11]	28.41, 138.98 Err : 1.3e-01	6.75, 143.27 Err : 1.2e-01	6.3, 143.36 Err : 1.2e-01	6.13, 143.4 Err : 1.2e-01
[1, 3, 7, 11]	27.94, 137.75 Err : 1.2e-01	1.44, 142.73 Err : 1.3e-01	1.5, 142.8 Err : 1.1e-01	1.5, 142.8 Err : 1.1e-01
[3, 5, 7]	13.92, 132.24 Err : 1.9e-01	14.03, 132.18 Err : 1.9e-01	223.42, 1.5 Err : 1.9e-01	13.91, 132.23 Err : 1.9e-01
[1, 3, 5, 7]	23.41, 137.16 Err : 1.4e-01	1.33, 144.12 Err : 2.7e-01	1.5, 144.36 Err : 1.2e-01	3.0, 148.52 Err : 1.3e-01
[2, 3, 5, 7]	18.39, 139.46 Err : 1.7e-01	23.89, 136.02 Err : 1.7e-01	238.19, 1.5 Err : 1.7e-01	216.14, 19.13 Err : 1.7e-01
[1, 2, 5, 7]	24.17, 138.77 Err : 1.4e-01	1.5, 146.2 Err : 1.3e-01	1.5, 146.21 Err : 1.3e-01	1.5, 146.21 Err : 1.3e-01

TABLE 1.2 – Tableau récapitulatif des optimisations réalisées sur 2 niveaux de gris, τ_S fixé à 197, avec pour pénalisation une parabole tronquée régularisée (cf. Figure 1.2b), pour les configurations défavorables de la Table ??, page ??.

Essayons donc une troisième fonction de pénalisation qui soit \mathcal{C}^1 . Considérons la régularisation de Moreau-Yosida de la pénalisation précédente. La courbe de la régularisée d'une parabole tronquée est représentée sur la Figure 1.2b (on prend ici $a = 126; u = E(\tau) - 127.5$). Les niveaux de gris optimaux obtenus avec cette nouvelle pénalisation plus régulière sont présentés dans la Table 1.2.

Visiblement rien ne semble y faire : il y a toujours des valeurs de τ_N qui s'approchent des bornes autorisées et les cas où l'on considère les images $[2, 5, 7]$ et $[3, 7, 9]$ par exemple fournissent des résultats très différents selon les algorithmes d'optimisation utilisés, qui de plus font apparaître $\tau_N \gg \tau_S$ ce qui est aberrant. La régularité de la pénalisation ne semblent donc pas en cause ici. Le problème semble donc venir d'ailleurs (*cf.* fin de la section ?? traitant des problèmes de mauvais conditionnement de nos systèmes).