Introduction à l'optimisation continue
Contrôle
(9 janvier 2018)

Exercice I : algorithme de Dykstra et variantes

On considère X un espace Euclidien (de dimension finie). On rappelle que pour une fonction convexe, semi-continue inférieurement (sci) $f : X \to \mathbb{R} \cup \{+\infty\}$,

$$\operatorname{prox}_f(\bar{x}) = \arg\min_x f(x) + \frac{1}{2}\|x - \bar{x}\|^2.$$

On considère f, g convexes, sci et on suppose qu'on sait facilement calculer prox_f et prox_g. On cherche un algorithme pour calculer

$$\operatorname{prox}_{f+g}(\bar{x}) = \arg\min_x f(x) + g(x) + \frac{1}{2}\|x - \bar{x}\|^2. \tag{A}$$

1. On introduit f^*, g^* les conjuguées de Legendre-Fenchel de f, g. Rappeler comment on exprime $\operatorname{prox}_{f^*}$ en fonction de prox_f.

L'identité de Moreau dit que $\operatorname{prox}_{f^*}(\bar{x}) = \bar{x} - \operatorname{prox}_f(\bar{x})$.

2. On suppose qu'il existe un point dans l'intersection des intérieurs relatifs de $\operatorname{dom} f$ et $\operatorname{dom} g$. Déduire de **1.** que si $\hat{x}$ minimise (A), $\hat{x} = \bar{x} - \hat{y} - \hat{z}$ où $(\hat{y}, \hat{z})$ minimise le problème

$$\min_{y,z} f^*(y) + g^*(z) + \frac{1}{2}\|y + z - \bar{x}\|^2. \tag{B}$$

Les hypothèses [Lemme 4.18 du cours] garantissent que l'inf-convolution

$$x \mapsto f^* \square g^*(x) := \min_{y+z=x} f^*(y) + g^*(z)$$

est convexe, sci et que le min est atteint. De plus, on a [Corollaire 4.19] $f^* \square g^* = (f+g)^*$. D'après l'identité de Moreau, $\hat{x} = \bar{x} - \operatorname{prox}_{(f+g)^*}(\bar{x})$. Donc pour calculer $\hat{x}$ il suffit de savoir calculer un minimiseur ξ de

$$\min_\xi \left(\min_{y+z=\xi} f^*(y) + g^*(z) \right) + \frac{1}{2}\|\xi - \bar{x}\|^2$$

ce qui (du fait que le min sur y, z est bien atteint [Lemme 4.18]) est équivalent à (B).

3. Montrer que l'on sait trouver, à y fixé, le point z_y qui minimise

$$\min_z g^*(z) + \frac{1}{2}\|y + z - \bar{x}\|^2. \tag{B_y}$$

et de même, à z fixé, le point y_z qui minimise

$$\min_y f^*(y) + \frac{1}{2}\|y + z - \bar{x}\|^2. \qquad (B_z).$$

Donner une expression de y_z et z_y en fonction de y, z et en utilisant les fonctions prox_f et prox_g.

A nouveau, c'est parce qu'on sait calculer prox_f et prox_g, et donc, par l'identité de Moreau, prox_{f^*} et prox_{g^*}. On a plus particulièrement $z_y = \mathrm{prox}_{g^*}(\bar{x} - y) = \bar{x} - y - \mathrm{prox}_g(\bar{x} - y)$ et de même $y_z = \mathrm{prox}_{f^*}(\bar{x} - z) = \bar{x} - z - \mathrm{prox}_f(\bar{x} - z)$.

Algorithme de minimisations alternées. On introduit pour (B) un algorithme de minimisations alternées : étant donné y^0, on définit pour $k \geq 0$ z^k solution de (B_{y^k}) puis y^{k+1} solution de (B_{z^k}). On pose $x^k = \bar{x} - y^k - z^k$ pour $k \geq 0$, et on veut montrer que $x^k \to \hat{x}$, solution de (A).

On introduit, pour y fixé, la fonction

$$G^*(y) := \min_z g^*(z) + \frac{1}{2}\|y + z - \bar{x}\|^2 = g^*(z_y) + \frac{1}{2}\|y + z_y - \bar{x}\|^2.$$

4. Montrer que G^* est convexe, C^1 à gradient 1-Lipschitz, et que $\nabla G^*(y) = z_y + y - \bar{x}$. En déduire que $y^{k+1} = \mathrm{prox}_{f^*}(y^k - \nabla G^*(y^k))$.

$G^*(y)$ est la régularisée de Moreau-Yosida de g^* au point $\bar{x} - y$ (notée $(g^*)_1(\bar{x} - y)$ dans le cours). Son gradient est donc bien 1-Lipschitz, voir par exemple la section 4.3 du cours. On a $\nabla(g^*)_1(\xi) = \xi - \mathrm{prox}_{g^*}(\xi)$, donc $\nabla G^*(y) = z_y + y - \bar{x}$.

5. En notant $(\hat{y}, \hat{z})$ une solution optimale de (B) et $\hat{x} = \bar{x} - \hat{y} - \hat{z}$, montrer que pour tout (y, z),

$$f^*(y) + g^*(z) + \frac{1}{2}\|y + z - \bar{x}\|^2 - \left(f^*(\hat{y}) + g^*(\hat{z}) + \frac{1}{2}\|\hat{y} + \hat{z} - \bar{x}\|^2\right) \geq \frac{1}{2}\|(y + z) - (\hat{y} + \hat{z})\|^2$$

et en déduire que pour tout k,

$$f^*(y^k) + G^*(y^k) - (f^*(\hat{y}) + G^*(\hat{y})) \geq \frac{1}{2}\|x^k - \hat{x}\|^2.$$

La première formule est une conséquence du fait que le problème est fortement convexe par rapport à $(y + z)$. On peut par exemple écrire que si on pose $\xi = y + z$, $\hat{\xi} = \hat{y} + \hat{z}$,

$$\begin{aligned} f^*(y) + g^*(z) + \frac{1}{2}\|y + z - \bar{x}\|^2 &\geq f^*\Box g^*(\xi) + \frac{1}{2}\|\xi - \bar{x}\|^2 \\ &\geq f^*\Box g^*(\hat{\xi}) + \frac{1}{2}\|\hat{\xi} - \bar{x}\|^2 + \frac{1}{2}\|\xi - \hat{\xi}\|^2 \end{aligned}$$

en remarquant que $\hat{\xi}$ est le minimiseur (unique) de la fonction fortement convexe $f^*\Box g^*(\xi)+\frac{1}{2}\|\xi-\bar{x}\|^2$.

La deuxième formule s'obtient en remarquant que $\hat{z}$ minimise $(B_{\hat{y}})$, de sorte que $g^*(\hat{z})+\frac{1}{2}\|\hat{y}+\hat{z}-\bar{x}\|^2 = G^*(\hat{y})$. Par ailleurs, par définition $G^*(y^k) = g^*(z^k)+\frac{1}{2}\|y^k+z^k-\bar{x}\|^2$. On déduit alors la seconde formule de la première.

6. En déduire que $\|x^k-\hat{x}\| \le \|y^0-\hat{y}\|/\sqrt{k}$. Proposer une variante de l'algorithme améliorant ce taux et vérifiant $\|x^k-\hat{x}\|\le 2\|y^0-\hat{y}\|/(k+1)$.

La formule (42) du cours (avec $\tau=1$) montre que $f^*(y^k)+G^*(y^k)-(f^*(\hat{y})+G^*(\hat{y}))\le \|y^0-\hat{y}\|^2/(2k)$, par conséquent $\|x^k-\bar{x}\|^2 \le \|y^0-\hat{y}\|^2/k$. Pour accélerer, on peut ici appliquer l'algorithme « FISTA ». Ainsi, on va poser plutôt pour tout k : $\bar{y}^k = y^k+\beta_k(y^k-y^{k-1})$ avec $\beta_k=(t_k-1)/t_{k+1}$, $t_k=(k+1)/2$, et $y^{k+1}=(I+\partial f^*)^{-1}(\bar{y}^k-\nabla G^*(\bar{y}^k))$; en pratique cela implique qu'on trouve z^k en minimisant $(B_{\bar{y}^k})$ puis y^{k+1} en minimisant (B_{z^k}). D'après le cours, on a alors $\|x^k-\bar{x}\|^2\le 4/(k+1)^2\|y^0-\hat{y}\|^2$.

7 - Algorithme de Dykstra. Montrer (utiliser les expressions demandées en **3.**) qu'on peut réécrire l'algorithme de minimisations alternées (pour $y^0=0$) sous la forme

1. $y^0=0$, $z^{-1}=0$, $x^{-\frac{1}{2}}=\bar{x}$; et pour $k\ge 0$:
2. $x^k=\mathrm{prox}_g(x^{k-\frac{1}{2}}+z^{k-1})$, $z^k=z^{k-1}+(x^{k-\frac{1}{2}}-x^k)$;
3. $x^{k+\frac{1}{2}}=\mathrm{prox}_f(x^k+y^k)$, $y^{k+1}=y^k+(x^k-x^{k+\frac{1}{2}})$.

On a vu que $z^k=\mathrm{prox}_{g^*}(\bar{x}-y^k)=\bar{x}-y^k-\mathrm{prox}_g(\bar{x}-y^k)$ et de même $y^{k+1}=\mathrm{prox}_{f^*}(\bar{x}-z^k)=\bar{x}-z^k-\mathrm{prox}_f(\bar{x}-z^k)$. Donc

$$\bar{x}-y^k-z^k=\mathrm{prox}_g(\bar{x}-y^k)=\mathrm{prox}_g(\bar{x}-y^k-z^{k-1}+z^{k-1}),$$

$$\bar{x}-y^{k+1}-z^k=\mathrm{prox}_f(\bar{x}-z^k)=\mathrm{prox}_f(\bar{x}-y^k-z^k+y^k).$$

En posant $x^{k-\frac{1}{2}}:=\bar{x}-y^k-z^{k-1}$, on obtient bien que $x^k=\mathrm{prox}_g(x^{k-\frac{1}{2}}+z^{k-1})$ et $x^{k+\frac{1}{2}}=\mathrm{prox}_f(x^k+y^k)$. Les formules pour calculer z^k puis y^{k+1} sont aisément déduites.

8 - Un cas particulier. On suppose maintenant que $f=\delta_E$ et $g=\delta_F$ où E et F sont des ensembles convexes fermés, d'intersection non vide i.e. : $f(x)=0$ si $x\in E$, $+\infty$ sinon). Que valent $\mathrm{prox}_f(x)$, $\mathrm{prox}_g(x)$? Quelle est la limite de l'algorithme de Dykstra ?

On suppose de plus que E et F sont des espaces affines. Montrer que pour tout k, $z^k\perp F$ et $y^k\perp E$. En déduire que l'algorithme de projections alternées : $x^0=\bar{x}$, $x^{k+\frac{1}{2}}=\Pi_E(x^k)$, $x^{k+1}=\Pi_F(x^{k+\frac{1}{2}})$ converge vers $\Pi_{E\cap F}(\bar{x})$, où Π_A est la projection orthogonale sur l'ensemble A.

Bien sûr, $\mathrm{prox}_f(x)=\arg\min_{\xi\in E}\frac{1}{2}\|\xi-x\|^2$ est $\Pi_E(x)$, la projection (orthogonale) de x sur E. Et l'algorithme de Dykstra, qui converge vers $\mathrm{prox}_{f+g}(\bar{x})$, converge vers le projeté de x sur $E\cap F$ puisque $\delta_E+\delta_F=\delta_{E\cap F}$.

Dans le cas où E est affine, on vérifie par récurrence que $y^k\perp E$. En effet, $y^0=0$, et si on suppose $y^k\perp E$, alors $x^{k+\frac{1}{2}}=\mathrm{prox}_f(x^k+y^k)=\Pi_E(x^k+y^k)=\Pi_E(x^k)$, tandis que $y^{k+1} = y^k + (x^k - x^{k+\frac{1}{2}}) \perp E$ puisque $x^k - \Pi_E(x^k) \in E^\perp$. Donc on n'a pas besoin d'utiliser y^k (ni de le garder en mémoire) au cours des itérations. De même si F est affine, on peut oublier z^k. Dans ce cas l'algorithme de Dykstra est équivalent à un algorithme de projections alternées.

Exercice II

Calculer les « prox » des fonctions suivantes :

1. τf, $\tau > 0$, $f(x) = |x|^3/3$ ($x \in \mathbb{R}$) ;
2. la fonction de « Huber » $h(x) = x^2/2$ si $|x| \leq 1$, $|x| - 1/2$ si $|x| \geq 1$.
3. Sa conjuguée $h^*(y) = y^2/2$ si $|y| \leq 1$, $+\infty$ sinon.
4. $f(x) = \max_{i=1}^N x_i$, $x \in \mathbb{R}^N$ (on écrira le problème sous la forme
$$\min_x \left(\max_{i=1,\ldots,N} x_i + \tfrac{1}{2}\|x-y\|^2 \right) = \min_t \left(t + \min_{x_i \leq t\, \forall i} \tfrac{1}{2}\|x-y\|^2 \right),$$
pourquoi ?) Sans restreindre la généralité, on pourra supposer qu'on calcule $\mathrm{prox}_f(\bar{x})$ pour $\bar{x}$ tel que $\bar{x}_1 \geq \bar{x}_2 \geq \cdots \bar{x}_N$.
5. $f(u) = \int_{-\infty}^{\infty} (u')^2/2dx$, $u \in L^2(\mathbb{R})$ ($+\infty$ si $u \notin H^1(\mathbb{R})$) : on montrera que $v = \mathrm{prox}_f u$ est donné par $\rho * u$, où $\rho(t) = e^{-|t|}/2$.

Exercice III

Suggérer des algorithmes pour résoudre :

1. Pour $x \in \mathbb{R}^N$, $N \geq 1$:
$$\min_{|x_i| \leq 1, i=1,\ldots,N} \|Ax - b\|^2$$
(norme Euclidienne, $A \in \mathbb{R}^{d\times N}$, $b \in \mathbb{R}^d$).

2. Pour $x, \bar{x} \in \mathbb{R}^N$, $N \geq 1$:
$$\min \left\{ \|x - \bar{x}\|^2 : |x_i| \leq 1, i = 1, \ldots, N, \quad \sum_{i=1}^N x_i^2 \leq \delta^2 \right\}$$
($\delta > 0$ donné).

Lorsque c'est possible, donnez les taux de convergence.

Exercice IV : splitting de trois opérateurs

On veut résoudre
$$A\xi + B\xi + C\xi \ni 0 \tag{Zer}$$
où A, B, C sont des opérateurs maximaux monotones dans un espace de Hilbert X et C est co-coercif : $\langle Cx - Cy, x - y\rangle \geq \gamma\|Cx - Cy\|^2$ pour tout x, y (notamment donc, $1/\gamma$-Lipschitz).

On introduit le Splitting de Davis et Yin (2015), pour $\tau > 0$:
$$T_\tau := I - J_{\tau B} + J_{\tau A} \circ (2J_{\tau_B} - I - \tau C \circ J_{\tau_B}),$$
où comme toujours $J_{\tau A} := (I + \tau A)^{-1}$. On introduit alors l'algorithme : $x^0 \in X$, $x^{k+1} = T_\tau x^k$ pour $k \geq 0$.

1. Quels sont les points fixes de T_τ ?

On a $x = T_\tau x$ si et seulement si
$$\begin{aligned}
& J_{\tau B}x = J_{\tau A}(2J_{\tau B}x - x - \tau C J_{\tau B}x) \\
\Leftrightarrow\ & J_{\tau B}x + \tau A J_{\tau B}x = 2J_{\tau B}x - x - \tau C J_{\tau B}x \\
\Leftrightarrow\ & x - J_{\tau B}x + \tau A J_{\tau B}x + \tau C J_{\tau B}x \ni 0 \\
\Leftrightarrow\ & B J_{\tau B}x + A J_{\tau B}x + C J_{\tau B}x \ni 0
\end{aligned}$$
en utilisant que $x - J_{\tau B}x = \tau B J_{\tau B}x$. Donc x est point fixe si et seulement si $J_{\tau B}x$ résout (Zer).

2. Que devient l'algorithme si $A = 0$? si $B = 0$? si $C = 0$?

Pour $A = 0$, $J_{\tau A} = I$, et on trouve $T_\tau = (I - \tau C) \circ J_{\tau_B}$. Il s'agit d'un algorithme implicite-explicite (ou explicite-implicite décalé d'une demi-itération).

Pour $B = 0$, $T_\tau = J_{\tau A} \circ (I - \tau C)$, il s'agit d'un algorithme explicite-implicite.

Pour $C = 0$, il s'agit du splitting de Douglas-Rachford introduit en cours.

3. Montrer que $(I - \tau C)$ est un opérateur θ-moyenné pour $0 < \tau < 2\gamma$, où $\theta = \tau/(2\gamma)$. On peut donc écrire $I - \tau C = (1 - \theta)I + \theta S$ pour un opérateur 1-Lipschitz S.

C'est une question de cours. On a
$$\begin{aligned}
&\|(I - \tau C)x - (I - \tau C)y\|^2 + \|\tau Cx - \tau Cy\|^2 \\
&\qquad = \|x - y\|^2 - 2\tau\langle x - y, Cx - Cy\rangle + 2\tau^2\|Cx - Cy\|^2 \\
&\qquad\qquad \leq \|x - y\|^2 + 2\tau(\tau - \gamma)\|Cx - Cy\|^2 \leq \|x - y\|^2
\end{aligned}$$
lorsque $\tau \leq \gamma$. Ainsi, pour $\tau \leq \gamma$, $I - \tau C$ est « firmly non expansive » et donc (1/2)-moyenné (on peut donc l'écrire $(I + S)/2$ avec S non expansif, i.e., 1-Lipschitz).

Si $0 < \tau < 2\gamma$, en posant $\theta = \tau/(2\gamma) \in (0, 1)$, on a
$$I - \tau C = I - 2\gamma\theta C = 2\theta(I - \gamma C) + (1 - 2\theta)I = \theta S + (1 - \theta)I$$
qui est θ-moyenné.

4. Pour $\tau < 2\gamma$ on a donc :

$$T_\tau = I - J_{\tau B} + J_{\tau A} \circ (J_{\tau_B} - I + ((1-\theta)I + \theta S) \circ J_{\tau_B}),$$

et l'on rappelle que $J_{\tau B} = (I + R_{\tau B})/2$ et $J_{\tau_A} = (I + R_{\tau A})/2$, où $R_{\tau A}$, $R_{\tau B}$ sont non-expansifs. On peut donc écrire (avec $S_\theta = (1-\theta)I + \theta S$) :

$$T_\tau = \frac{1}{2}(I - J_{\tau B} + S_\theta J_{\tau B}) + \frac{1}{2}R_{\tau A} \circ (J_{\tau B} - I + S_\theta J_{\tau B}). \tag{1}$$

Notamment, par convexité et en utilisant que $R_{\tau A}$ est 1-Lipschitz,

$$\begin{aligned}\|T_\tau x - T_\tau y\|^2 \le\ & \frac{1}{2}\|(I - J_{\tau B} + S_\theta J_{\tau B})x - (I - J_{\tau B} + S_\theta J_{\tau B})y\|^2 \\ & +\ \frac{1}{2}\|R_{\tau A}(J_{\tau B} - I + S_\theta J_{\tau B})x - R_{\tau A}(J_{\tau B} - I + S_\theta J_{\tau B})y\|^2 \\ \le\ & \frac{1}{2}\|(I - J_{\tau B} + S_\theta J_{\tau B})x - (I - J_{\tau B} + S_\theta J_{\tau B})y\|^2 \\ & +\ \frac{1}{2}\|(J_{\tau B} - I + S_\theta J_{\tau B})x - (J_{\tau B} - I + S_\theta J_{\tau B})y\|^2.\end{aligned} \tag{2}$$

En déduire que T_τ est non-expansif. Peut-on en déduire que x^k converge ? Proposer une modification de l'algorithme qui garantit la convergence.

L'identité du parallélogramme entraine que

$$\begin{aligned}&\frac{1}{2}\|(I - J_{\tau B} + S_\theta J_{\tau B})x - (I - J_{\tau B} + S_\theta J_{\tau B})y\|^2 \\ &\qquad\qquad + \frac{1}{2}\|(J_{\tau B} - I + S_\theta J_{\tau B})x - (J_{\tau B} - I + S_\theta J_{\tau B})y\|^2 \\ &\qquad\quad = \|(I - J_{\tau B})x - (I - J_{\tau B})y\|^2 + \|S_\theta J_{\tau B}x - S_\theta J_{\tau B}y\|^2 \le \|x-y\|^2\end{aligned} \tag{3}$$

en utilisant que S_θ est non-expansif et $J_{\tau B}$ « firmly non expansive ».

Mais on n'a pas montré que T_τ était moyenné : on n'est donc pas sûr de la convergence. Une méthode convergente consisterait à poser $x^{k+1} = \alpha x^k + (1-\alpha)T_\tau x^k$ pour $\alpha \in (0,1)$.

5. On peut montrer (difficile) que T_τ est effectivement moyenné (on n'a pas utilisé jusqu'ici le fait que S_θ l'était). On remarque que les expressions qui apparaissent dans (1) s'écrivent :

$$\begin{aligned}I - J_{\tau B} + S_\theta J_{\tau B} =\ & (1-\theta)I + \theta(I - J_{\tau B} + SJ_{\tau B}), \\ J_{\tau B} - I + S_\theta J_{\tau B} =\ & (2-\theta)J_{\tau B} - (1-\theta)I + \theta(-I + SJ_{\tau B}) \\ =\ & (1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B}).\end{aligned}$$

L'équation (1) se réécrit :

$$\begin{aligned}T_\tau = \tfrac{1-\theta}{2}I + \tfrac{\theta}{2}(I - J_{\tau B} + SJ_{\tau B}) + \tfrac{1}{2}R_{\tau A}((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B})) \\ = (1 - \tfrac{1+\theta}{2})I + \tfrac{1+\theta}{2}\tilde{T}\end{aligned}$$

où
$$\tilde{T} = \frac{\theta}{1+\theta}(I - J_{\tau B} + SJ_{\tau B}) + \frac{1}{1+\theta}R_{\tau A}\left((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B})\right).$$
En s'inspirant de la question **4.** (en particulier (2)), montrer que $\tilde{T}$ est non-expansif. Conclure. (Remarque : le papier original de Davis et Yin montre un taux de moyennage légèrement meilleur $(2/(2-\theta))$).

On écrit à nouveau

$$\begin{aligned}\|\tilde{T}x - \tilde{T}y\|^2 &\le \tfrac{\theta}{1+\theta}\|(I - J_{\tau B} + SJ_{\tau B})x - (I - J_{\tau B} + SJ_{\tau B})y\|^2\\ &+\tfrac{1}{1+\theta}\|R_{\tau A}((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B}))x - R_{\tau A}((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B}))y\|^2\\ &\le \tfrac{\theta}{1+\theta}\|(I - J_{\tau B} + SJ_{\tau B})x - (I - J_{\tau B} + SJ_{\tau B})y\|^2\\ &+ \tfrac{1}{1+\theta}\|((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B}))x - ((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B}))y\|^2\end{aligned}$$

puisque $R_{\tau A}$ est 1-Lipschitz. De plus,

$$\begin{aligned}&\|((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B}))x - ((1-\theta)R_{\tau B} + \theta(J_{\tau B} - I + SJ_{\tau B}))y\|^2\\ &\quad\le (1-\theta)\|R_{\tau B}x - R_{\tau B}y\|^2 + \theta\|(J_{\tau B} - I + SJ_{\tau B})x - (J_{\tau B} - I + SJ_{\tau B})y\|^2\\ &\quad\le (1-\theta)\|x-y\|^2 + \theta\|(J_{\tau B} - I + SJ_{\tau B})x - (J_{\tau B} - I + SJ_{\tau B})y\|^2\end{aligned}$$

puisque $R_{\tau B}$ est 1-Lipschitz. Donc :

$$\begin{aligned}\|\tilde{T}x - \tilde{T}y\|^2 &\le \tfrac{\theta}{1+\theta}\|(I - J_{\tau B} + SJ_{\tau B})x - (I - J_{\tau B} + SJ_{\tau B})y\|^2\\ &\quad+ \tfrac{\theta}{1+\theta}\|(J_{\tau B} - I + SJ_{\tau B})x - (J_{\tau B} - I + SJ_{\tau B})y\|^2 + \tfrac{1-\theta}{1+\theta}\|x-y\|^2\end{aligned}$$

Comme à la question **5.**, l'identité du parallélogramme donne

$$\begin{aligned}&\|(I - J_{\tau B} + SJ_{\tau B})x - (I - J_{\tau B} + SJ_{\tau B})y\|^2\\ &\qquad+ \|(J_{\tau B} - I + SJ_{\tau B})x - (J_{\tau B} - I + SJ_{\tau B})y\|^2\\ &\quad= 2\left(\|(I - J_{\tau B})x - (I - J_{\tau B})y\|^2 + \|SJ_{\tau B}x - SJ_{\tau B}y\|^2\right)\\ &\quad\le 2\left(\|(I - J_{\tau B})x - (I - J_{\tau B})y\|^2 + \|J_{\tau B}x - J_{\tau B}y\|^2\right) \le 2\|x-y\|^2\end{aligned}$$

puisque S est 1-Lipschitz et $J_{\tau B}$ est « firmly non expansive ». On en déduit que

$$\|\tilde{T}x - \tilde{T}y\|^2 \le \tfrac{2\theta}{1+\theta}\|x-y\|^2 + \tfrac{1-\theta}{1+\theta}\|x-y\|^2 = \|x-y\|^2.$$

Donc T_τ est moyenné (de facteur $(1+\theta)/2$). Il en résulte, d'après le théorème de Krasnoselskii et Mann, que x^k converge (faiblement) vers un point fixe x, et donc tel que $J_{\tau B}x$ résout (Zer) (on suppose qu'il y a des solutions).