A primal-dual dynamical approach to a nonsmooth convex minimization

Szilárd Csaba László

Department of Mathematics

Technical University of Cluj-Napoca

Joint work with Radu Ioan Boț and Ernö Robert Csetnek

GDO, 2019

Cluj-Napoca

Introduction

Let $\mathcal{H}$ and $\mathcal{G}$ be real Hilbert spaces and consider the optimization problem

$$\inf_{x \in \mathcal{H}} (f(x) + h(x) + g(Ax)), \tag{1}$$

where, $f : \mathcal{H} \longrightarrow \overline{\mathbb{R}}$, $g : \mathcal{G} \longrightarrow \overline{\mathbb{R}}$ are proper, convex and lower semicontinuous functions, $h : \mathcal{H} \longrightarrow \mathbb{R}$ is a convex and Frèchet differentiable function with L_h Lipschitz continuous gradient, i.e. there exists $L_h \geq 0$ such that $\|\nabla h(x) - \nabla h(y)\| \leq L_h \|x - y\|$ for all $x, y \in \mathcal{H}$, and $A : \mathcal{H} \longrightarrow \mathcal{G}$ is a continuous linear map.

If $L_h = 0$ obviously h is constant and will not contribute to problem (1), therefore we will assume in this case that $h \equiv 0$.

Note, that problem (1) can be rewritten as

$$\inf_{\substack{(x,z)\in\mathcal{H}\times\mathcal{G}\\ Ax-z=0}} (f(x)+h(x)+g(z)), \tag{2}$$

hence $x^* \in \mathcal{H}$ is an optimal solution of (1), if and only if $(x^*, z^*) \in \mathcal{H} \times \mathcal{G}$ is an optimal solution of (2), and $Ax^* = z^*$.

Note, that problem (1) can be rewritten as

$$\inf_{\substack{(x,z)\in\mathcal{H}\times\mathcal{G}\\ Ax-z=0}} (f(x)+h(x)+g(z)), \tag{2}$$

hence $x^* \in \mathcal{H}$ is an optimal solution of (1), if and only if $(x^*, z^*) \in \mathcal{H} \times \mathcal{G}$ is an optimal solution of (2), and $Ax^* = z^*$.

The usefulness of the above formulation consists in the fact that the Lagrangian,

$$l : \mathcal{H} \times \mathcal{G} \times \mathcal{G} \longrightarrow \overline{\mathbb{R}},\ l(x, z, y) = f(x) + h(x) + g(z) + \langle y, Ax - z\rangle,$$

can be introduced.

We emphasize that $(x^*, z^*, y^*) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ is a saddle point of l, that is

$$l(x^*, z^*, y) \leq l(x^*, z^*, y^*) \leq l(x, z, y^*), \forall (x, z, y) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G},$$

if and only if $z^* = Ax^*$, x^* is an optimal solution of (1), hence, (x^*, z^*) is an optimal solution of (2), and y^* is an optimal solution of the Fenchel dual to problem (1), i.e.

$$\sup_{y \in \mathcal{G}} (-(f^* \Box h^*)(-A^* y) - g^*(y)) \tag{3}$$

and the optimal values of (1) and (3) coincide.

We emphasize that $(x^*, z^*, y^*) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ is a saddle point of l, that is

$$l(x^*, z^*, y) \leq l(x^*, z^*, y^*) \leq l(x, z, y^*), \forall (x, z, y) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G},$$

if and only if $z^* = Ax^*$, x^* is an optimal solution of (1), hence, (x^*, z^*) is an optimal solution of (2), and y^* is an optimal solution of the Fenchel dual to problem (1), i.e.

$$\sup_{y \in \mathcal{G}} (-(f^* \Box h^*)(-A^* y) - g^*(y)) \tag{3}$$

and the optimal values of (1) and (3) coincide.

The existence of a saddle point is guaranteed whenever the Attouch-Brézis regularity condition

$$0 \in \operatorname{sqri}(\operatorname{dom} g - A(\operatorname{dom} f))$$

holds, where sqri denotes the strong quasi-relative interior of a set.

Let us denote by $S_+(\mathcal{H})$ the family of continuous linear operators $U : \mathcal{H} \longrightarrow \mathcal{H}$ which are self-adjoint and positive semidefinite and for $U \in S_+(\mathcal{H})$ we introduce the following seminorm:

$$\|x\|_U^2 = \langle x, Ux\rangle, \forall x \in \mathcal{H}.$$

In $S_+(\mathcal{H})$ can be introduced a partial ordering as follows: for $U_1, U_2 \in S_+(\mathcal{H})$

$$U_1 \succcurlyeq U_2 \Leftrightarrow \|x\|_{U_1}^2 \geq \|x\|_{U_2}^2 \ \forall x \in \mathcal{H}.$$

For $\alpha > 0$ we denote

$$P_\alpha(\mathcal{H}) = \{U \in S_+(\mathcal{H}) : U \succcurlyeq \alpha I\}.$$

Here $I : \mathcal{H} \longrightarrow \mathcal{H}$, $I(x) = x$ denotes the identity operator.

Consider the mappings $M_1 : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ and $M_2 : [0, +\infty) \longrightarrow S_+(\mathcal{G})$ and the parameters $c > 0$, $\gamma \geq 0$.

We define the functions $F : [0, +\infty) \times \mathcal{H} \longrightarrow \overline{\mathbb{R}}$

$$F(t,x) = f(x) + \frac{c}{2}(\|Ax\|^2 - \|x\|^2) + \frac{1}{2}\|x\|^2_{M_1(t)}$$

and $G : [0, +\infty) \times \mathcal{G} \longrightarrow \overline{\mathbb{R}}$

$$G(t,x) = g(x) + \frac{1}{2}\|x\|^2_{M_2(t)}.$$

The dynamical system related to the problems (1)-(3) is

$$\begin{cases} \dot{x}(t) + x(t) \in \operatorname{argmin}_{x \in \mathcal{H}} \left(F(t,x) + \frac{c}{2} \left\| x - \left(\frac{M_1(t)}{c} x(t) + A^* z(t) - \frac{A^*}{c} y(t) - \frac{1}{c} \nabla h(x(t)) \right) \right\|^2 \right) \\ \dot{z}(t) + z(t) = \operatorname{argmin}_{x \in \mathcal{G}} \left(G(t,x) + \frac{c}{2} \left\| x - \left(\frac{M_2(t)}{c} z(t) + A(\gamma \dot{x}(t) + x(t)) + \frac{1}{c} y(t) \right) \right\|^2 \right) \\ \dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\ t \in [0, +\infty),\ x(0) = x_0 \in \mathcal{H},\ y(0) = y_0 \in \mathcal{G},\ z(0) = z_0 \in \mathcal{G},\ c > 0,\ \gamma \in [0,1]. \end{cases} \tag{4}$$

Remark

Meanwhile, for every $t \in [0, +\infty)$ the convexity of the function $G(t, \cdot)$ is obvious, hence the equality in the second equation of (4) is assured by the strong convexity of the function $x \to G(t, x) + \frac{c}{2}\|x - u\|^2$ for all $u \in \mathcal{G}$, observe that the positive semidefiniteness of the operator $M_1(t) + c(A^*A - I)$ for all $t \in [0, +\infty)$, ensures the convexity of the function $F(t, \cdot)$ and implicitly the equality in the first equation of (4).

In this case the dynamical system (4) becomes

$$
\begin{cases}
\dot{x}(t) = \mathrm{prox}_{\frac{1}{c}F(t,\cdot)}\left(\frac{M_1(t)}{c}x(t) + A^*z(t) - \frac{A^*}{c}y(t) - \frac{1}{c}\nabla h(x(t))\right) - x(t) \\
\dot{z}(t) = \mathrm{prox}_{\frac{1}{c}G(t,\cdot)}\left(\frac{M_2(t)}{c}z(t) + A(\gamma\dot{x}(t) + x(t)) + \frac{1}{c}y(t)\right) - z(t) \\
\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\
x(0) = x_0 \in \mathcal{H},\ y(0) = y_0 \in \mathcal{G},\ z(0) = z_0 \in \mathcal{G},\ c > 0,\ \gamma \in [0,1].
\end{cases}
\tag{5}
$$

Here

$$
\mathrm{prox}_{\lambda f} : \mathcal{H} \to \mathcal{H}, \quad \mathrm{prox}_{\lambda f}(x) = \mathrm{argmin}_{y\in\mathcal{H}}\left\{f(y) + \frac{1}{2\lambda}\|y - x\|^2\right\},
$$

denotes the proximal point operator of the convex function λf.

Remark

Nevertheless, the strong convexity of the function $x \to F(t,x) + \frac{c}{2}\|x-u\|^2$ for all $u \in \mathcal{H}$ is also assured if one of the following assumptions holds:

(C1) for every $t \in [0,+\infty)$ there exists $\alpha_1(t) > 0$ such that $M_1(t) \in P_{\alpha_1(t)}(\mathcal{H})$,

(C2) there exists $\alpha > 0$ such that $A^*A \in P_\alpha(\mathcal{H})$,

(C3) for every $t \in [0,+\infty)$ there exists $\alpha(t) > 0$ such that $cA^*A + M_1(t) \in P_{\alpha(t)}(\mathcal{H})$.

Indeed, one has

$$\partial_x \left(F(t,x) + \frac{c}{2}\|x-u\|^2 \right) = \partial f(x) + cA^*Ax + M_1(t)x - cu, \quad (6)$$

which is obviously $\alpha_1(t)-$strongly monotone, $c\alpha-$strongly monotone or $\alpha(t)-$strongly monotone, if (C1), (C2) or (C3) holds. Moreover, taking into account that $A^*A \in S_+(\mathcal{H})$ and $M_1(t) \in S_+(\mathcal{H})$ for all $t \in [0,+\infty)$ we conclude that

$$(C1) \Rightarrow (C3)$$

and

$$(C2) \Rightarrow (C3).$$

Remark

Let $S = \{x \in \mathcal{H} : \|x\| = 1\}$ the unit sphere of $\mathcal{H}$.

(A1) Observe that for every $t \geq 0$ we have

$$\alpha_1(t) \leq \inf_{x \in S} \langle x, M_1(t)x \rangle = \inf_{x \in S} \|x\|^2_{M_1(t)}.$$

Consequently (C1) holds, if and only if

$$\inf_{x \in S} \|x\|^2_{M_1(t)} > 0, \ \forall t \in [0, +\infty) \text{ and in this case one can take}$$

$$\alpha_1(t) = \inf_{x \in S} \|x\|^2_{M_1(t)}.$$

Note that $\inf_{t \in [0,+\infty)} \alpha_1(t) > 0$, if and only if, there exists $\alpha > 0$ such that $M_1(t) \in P_\alpha(\mathcal{H})$ for all $t \in [0, +\infty)$.

(A2) Similarly, (C2) holds, if and only if

$$\inf_{x\in S}\|Ax\|>0, \text{ and in this case one can take } \alpha=(\inf_{x\in S}\|Ax\|)^2.$$

(A3) Finally, (C3) holds, if and only if for all $t\in[0,+\infty)$ one has

$$\inf_{x\in S}(c\|Ax\|^2+\|x\|^2_{M_1(t)})>0, \text{ and in this case one can take}$$

$$\alpha(t)=\inf_{x\in S}(c\|Ax\|^2+\|x\|^2_{M_1(t)})$$

that is

$$\alpha(t)=\inf_{x\in S}\|x\|^2_{cA^*A+M_1(t)}.$$

Note that $\inf_{t\in[0,+\infty)}\alpha(t)>0$, if and only if, there exists $\alpha>0$ such that $cA^*A+M_1(t)\in P_\alpha(\mathcal{H})$ for all $t\in[0,+\infty)$.

Let us show that time discretization of the dynamical system (4) leads to the proximal ADMM algorithm from the literature, see M. Fazel, T.K. Pong, D. Sun, P. Tseng [1], R. Shefi, M. Teboulle[2], S. Banert, R.I. Boț, E.R. Csetnek[3] .

---

[1] Hankel matrix rank minimization with applications in system identification and realization, SIAM Journal on Matrix Analysis and Applications 34, 946-977, 2013

[2] Rate of convergence analysis of decomposition methods based on the proximal method of multipliers for convex optimization, Siam Journal on Optimization 24(1), 269-297, 2014,

[3] Fixing and extending some recent results on the ADMM algorithm, arXiv:1612.05057, 2016

Indeed, the first equation of (4) can be written as

$$0 \in \partial f(\dot{x}(t) + x(t)) + cA^*A(\dot{x}(t) + x(t)) + M_1(t)\dot{x}(t) - (cA^*z(t) - A^*y(t) - \nabla h(x(t))).$$

Consequently, by the explicit discretization of the above inclusion with respect to the time variable t, constant step size $h_k \equiv 1$ and initial points $x^0 = x_0, y^0 = y_0, z^0 = z_0$ yields the iterative scheme

$$0 \in \frac{1}{c}\partial f(x^{k+1}) + A^*Ax^{k+1} + \frac{M_1^k}{c}(x^{k+1} - x^k) - A^*z^k + \frac{A^*}{c}y^k + \frac{1}{c}\nabla h(x^k).$$

Hence,

$$0 \in \partial \left(f(x) + \langle x - x^k, \nabla h(x^k) \rangle + \frac{c}{2} \left\| Ax - z^k + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|x - x^k\|^2_{M_1^k} \right) \Bigg|_{x = x^{k+1}}$$

in other words

$$x^{k+1} \in \operatorname{argmin}_{x \in \mathcal{H}} \left(f(x) + \langle x - x^k, \nabla h(x^k) \rangle + \frac{c}{2} \left\| Ax - z^k + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|x - x^k\|^2_{M_1^k} \right). \tag{7}$$

Similarly, the second equation (4) leads to

$$z^{k+1} = \left(I + \tfrac{1}{c}\partial_x G(t_k,\cdot)\right)^{-1}\left(\tfrac{M_2^k}{c} z^k + A(\gamma x^{k+1} + (1-\gamma)x^k) + \tfrac{1}{c} y^k\right),$$

hence,

$$0 = \partial\left(g(z) + \tfrac{c}{2}\left\|A(\gamma x^{k+1} + (1-\gamma)x^k) - z + \tfrac{y^k}{c}\right\|^2 + \tfrac{1}{2}\|z - z^k\|^2_{M_2^k}\right)\Big|_{z=z^{k+1}}.$$

Consequently,

$$z^{k+1} = \operatorname{argmin}_{z\in\mathcal{G}}\left(g(z) + \tfrac{c}{2}\left\|A(\gamma x^{k+1} + (1-\gamma)x^k) - z + \tfrac{y^k}{c}\right\|^2 + \tfrac{1}{2}\|z - z^k\|^2_{M_2^k}\right). \tag{8}$$

Taking into account the equations (7) and (8), our dynamical system (4) leads through explicit discretization to

$$
\begin{cases}
x^{k+1} \in \operatorname{argmin}_{x\in\mathcal{H}} \left(f(x) + \langle x - x^k, \nabla h(x^k)\rangle + \frac{c}{2}\left\| Ax - z^k + \frac{y^k}{c}\right\|^2 + \frac{1}{2}\|x - x^k\|^2_{M_1^k} \right) \\
z^{k+1} = \operatorname{argmin}_{z\in\mathcal{G}} \left(g(z) + \frac{c}{2}\left\| A(\gamma x^{k+1} + (1-\gamma)x^k) - z + \frac{y^k}{c}\right\|^2 + \frac{1}{2}\|z - z^k\|^2_{M_2^k} \right) \\
y^{k+1} = y^k + c(Ax^{k+1} - z^{k+1}) \\
x^0 \in \mathcal{H},\ y^0, z^0 \in \mathcal{G}.
\end{cases}
\tag{9}
$$

Let us notice that in case $\gamma = 1$, $h = 0$ and M_1 and M_2 are constant in each iteration, this is nothing else than the proximal ADMM method from the literature.

Furthermore, the situation $\gamma = 0$ leads to an extension of the linearized proximal method of multipliers of Chen-Teboulle.[4]

Let us consider now the particular case

$$M_1(t) = \frac{1}{\tau(t)} I - cA^*A, \; M_2(t) = 0, \; \forall t \in [0, +\infty),$$

where $\tau(t) > 0$ for all $t \geq 0$ and $c\tau(t)\|A\|^2 \leq 1$.

In this particular case (4) becomes

$$\begin{cases} \dot{x}(t) + x(t) = \operatorname{prox}_{\tau(t)f}\left((I - c\tau(t)A^*A)x(t) + c\tau(t)A^*z(t) - \tau(t)A^*y(t) - \tau(t)\nabla h(x(t))\right) \\ \dot{y}(t) + y(t) + c(\gamma - 1)A\dot{x}(t) = \operatorname{prox}_{cg^*}(cA(\gamma\dot{x}(t) + x(t)) + y(t)) \\ \dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\ t \in [0, +\infty), \; x(0) = x_0 \in \mathcal{H}, \; y(0) = y_0 \in \mathcal{G}, \; z(0) = z_0 \in \mathcal{G}, \; c > 0, \; \gamma \in [0, 1]. \end{cases} \tag{10}$$

[4] A proximal-based decomposition method for convex minimization problems, Mathematical Programming 64, 81-101, 1994

The discretization of (10) in case $h \equiv 0$ and $\gamma = 1$ leads to

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\tau_k f}\left(x^k - \tau_k A^*(2y^k - y^{k-1})\right) \\ \\ y^{k+1} = \operatorname{prox}_{cg^*}(y^k + cAx^{k+1}). \end{cases} \tag{11}$$

When $\tau_k = \tau > 0$ for all $k \geq 1$, this iterative schemes becomes the primal-dual algorithm proposed by Chambolle and Pock.[5]

[5] A first-order primal-dual algorithm for convex problems with applications to imaging, Journal of Mathematical Imaging and Vision 40(1), 120-145, 2011

Existence and uniqueness

In what follows everywhere we assume that one of the conditions (C1)-(C3) stated in Remark 2 holds. Now we are able to specify which type of solutions are we considering in the analysis of the dynamical system (4).

Definition

We say that the vector function $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ is a strong global solutions of (4), if the following properties are satisfied:

(i) the functions x, z, y are locally absolutely continuous;

(ii)

$$\dot{x}(t) + x(t) = \operatorname{argmin}_{x \in \mathcal{H}} \left(F(t,x) + \frac{c}{2} \left\| x - \left(\frac{M_1(t)}{c} x(t) + A^* z(t) - \frac{A^*}{c} y(t) - \frac{1}{c} \nabla h(x(t)) \right) \right\|^2 \right),$$

$$\dot{z}(t) + z(t) = \operatorname{argmin}_{x \in \mathcal{G}} \left(G(t,x) + \frac{c}{2} \left\| x - \left(\frac{M_2(t)}{c} z(t) + A(\gamma \dot{x}(t) + x(t)) + \frac{1}{c} y(t) \right) \right\|^2 \right),$$

and

$$\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t))$$

for almost every $t \geq 0$;

(iii)

$$x(0) = x_0,\ y(0) = y_0,\ \text{and } z(0) = z_0.$$

We prove existence and uniqueness of a strong global solution of (4) by making use of the Cauchy-Lipschitz-Picard Theorem for absolutely continues trajectories. The key argument is that one can rewrite (4) as a particular first order dynamical system in a suitably chosen product space.

Theorem

Assume that one of the conditions (C1), (C2) or (C3) holds. Assume further that for every $T > 0$ *the functions*

$$t \longrightarrow \|M_1(t)\|, \; t \longrightarrow \|M_2(t)\|$$

are integrable on $[0, T]$*, that is,* $\|M_1(\cdot)\|, \; \|M_2(\cdot)\| \in L^1_{loc}([0, +\infty))$*. Then, for every starting points* $(x_0, z_0, y_0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$*, the dynamical system* (4) *has a unique strong global solution* $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$.

In order to continue our analysis we need the following derivative concept. We say that the map $M : [0, +\infty) \longrightarrow \mathcal{L}(\mathcal{H})$, $t \longrightarrow M(t)$ is derivable at $t \in [0, +\infty)$ if there exists the limit

$$\lim_{h \longrightarrow 0} \frac{M(t+h) - M(t)}{h}$$

taken after the topology induced by the norm of $\mathcal{L}(\mathcal{H})$.

Let us denote by $\dot{M}(t)$ the value of the above limit. Obviously, $\dot{M}(t) \in \mathcal{L}(\mathcal{H})$. If M is locally absolutely continuous then $\dot{M}(t)$ exists at almost every $t \in [0, +\infty)$. It is straightforward that, whenever $\dot{M}(t)$ exists, one has

$$\dot{M}(t)x = \lim_{h \longrightarrow 0} \frac{M(t+h)x - M(t)x}{h}, \text{ for every } x \in \mathcal{H}.$$

Assume now that $M(t) \in \mathcal{L}(\mathcal{H})$ is self adjoint for every $t \in [0, +\infty)$ and that is derivable at $t_0 \in [0, +\infty)$. Then, $\dot{M}(t_0)$ is also self adjoint.

Further we will need the following derivation formula used when we prove the convergence of the trajectories of (4).

Consider the maps $x, y : [0, +\infty) \longrightarrow \mathcal{H}$ and assume that x and y are derivable at t_0. Then, the real function $t \longrightarrow \langle M(t)x(t), y(t)\rangle$ is also derivable at t_0 and one has

$$\frac{d}{dt}\langle M(t)x(t), y(t)\rangle\Big|_{t=t_0} = \tag{12}$$

$$\langle \dot{M}(t_0)x(t_0), y(t_0)\rangle + \langle M(t_0)\dot{x}(t_0), y(t_0)\rangle + \langle M(t_0)x(t_0), \dot{y}(t_0)\rangle.$$

Assume that the mappings $t \longrightarrow M_1(t)$, $t \longrightarrow M_2(t)$ are locally absolutely continuous on $[0, +\infty)$, and for the starting points $(x_0, z_0, y_0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, let $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be the unique strong global solution of the dynamical system (4). Then, $t \longrightarrow (\dot{x}(t), \dot{z}(t), \dot{y}(t))$ is locally absolutely continuous, hence $(\ddot{x}(t), \ddot{z}(t), \ddot{y}(t))$ exists for almost every $t \geq 0$. Moreover, if $\sup_{t\geq 0} \|M_1(t)\| < +\infty$ and $\sup_{t\geq 0} \|M_2(t)\| < +\infty$, then there exists $L > 0$ such that

$$\|\ddot{x}(t)\| + \|\ddot{z}(t)\| + \|\ddot{y}(t)\| \leq$$

$$L(\|\dot{x}(t)\| + \|\dot{z}(t)\| + \|\dot{y}(t)\| + \|\dot{M}_1(t)\|\|\dot{x}(t)\| + \|\dot{M}_2(t)\|\|\dot{z}(t)\|),$$

for almost every $t \in [0, +\infty)$.

Next, we state a version of continuous Opial Lemma that will be used for showing the convergence of the trajectories generated by the dynamical system (4). It can be seen as the continuous counter-part of the Opial Lemma formulated in the setting of variable metrics by Combettes and Vũ.[6]

Lemma

Let $\mathcal{C} \subseteq \mathcal{H}$ be a nonempty set and let $x : [0, +\infty) \to \mathcal{H}$ be a continuous map. Let $M : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ and assume that there exists $\alpha > 0$ such that $M(t) \in P_\alpha(\mathcal{H})$ for all $t \in [0, +\infty)$. Assume further that $M(t_1) \succcurlyeq M(t_2)$ for all $t_1 \leq t_2$ and the following conditions hold.

[6] Variable metric quasi-Fejér monotonicity, Nonlinear Analysis 78, 17-31, 2013

(i) for every $z \in \mathcal{C}$, $\lim_{t\to+\infty} \|x(t) - z\|_{M(t)}$ exists;

(ii) every weak sequential cluster point of the map x belongs to $\mathcal{C}$.

Then there exists $x_\infty \in \mathcal{C}$ such that $w - \lim_{t\to+\infty} x(t) = x_\infty$.

Remark
If a map $M : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ satisfies $M(t_1) \succcurlyeq M(t_2)$ for all $t_1 \leq t_2$, $t_1, t_2 \in [0, +\infty)$ we will say that M is monotone decreasing. Note, that in case M is monotone decreasing and locally absolutely continuous then $\dot{M}(t)$ exists for almost every $t \in [0, +\infty)$ and, by making abuse of notation, $\|x\|^2_{\dot{M}(t)} = \langle \dot{M}(t)x, x\rangle \leq 0$ for almost every $t \in [0, +\infty)$.

Our convergence result is the following.

Theorem

Consider the Problem (1) *and assume that* $\mathcal{C}$, *the set of saddle points of the Lagrangian* l *is nonempty. Assume further that the maps* $M_1(t) + \frac{c(1-\gamma)}{4}A^*A - \frac{L_h}{4}I$, $M_1(t) \in S_+(\mathcal{H})$, $M_2(t) \in S_+(\mathcal{G})$ *for all* $t \geq 0$ *are locally absolutely continuous and monotone decreasing and* $\sup_{t\geq 0} \|\dot{M}_1(t)\| < +\infty$ *and* $\sup_{t\geq 0} \|\dot{M}_2(t)\| < +\infty$. *Moreover, assume that one of the following assumptions hold.*

(I) $M_1(t) + \frac{c(1-\gamma)}{4}A^*A - \frac{L_h}{4}I \in P_{\alpha_1}(\mathcal{H})$ *for all* $t \geq 0$ *and for some* $\alpha_1 > 0$.

(II) $\gamma \in [0, 1)$ *and* $A^*A \in P_\alpha(\mathcal{H})$ *for some* $\alpha > 0$.

For a starting point $(x_0, z_0, y_0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, let $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be the unique strong global solution of the dynamical system (4). Then, the vector function $t \longrightarrow (x(t), z(t), y(t))$ converges weakly to a saddle point of l as $t \longrightarrow +\infty$.

Our proof, beside the previously stated continuous Opial lemma, is based on a result of A. Alotaibi, P. L. Combettes and N. Shahzad.[7]

[7] Solving Coupled Composite Monotone Inclusions by Successive Fejér Approximations of their Kuhn-Tucker Set, SIAM J. Optim., 24(4), 2076-2095, (2014)

Lemma ACS

In the setting of Problem (1), let $(a_n, a_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{Gr}\partial(f+h)$, let $(b_n, b_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{Gr}\partial g$. Suppose that a_n converges weakly to $\overline{x}\in\mathcal{H}$, b_n^* converges weakly to $\overline{v}\in\mathcal{G}$ $a_n^* + A^*b_n^* \longrightarrow 0$, and $Aa_n - b_n \longrightarrow 0$. Then,

$$\langle a_n, a_n^*\rangle + \langle b_n, b_n^*\rangle \longrightarrow 0$$

and

$$\overline{v}\in\partial g(A\overline{x}),\ -A^*\overline{v} - \nabla h(\overline{x}) \in \partial f(\overline{x}).$$

Further, we derive the following key inequality.

For almost every $t \in [0, +\infty)$ one has

$$\frac{1}{2}\frac{d}{dt}\left(\|x(t) - x^*\|^2_{M_1(t)+c(1-\gamma)A^*A} + \|z(t) - Ax^*\|^2_{M_2(t)+cI} + \frac{1}{c}\|y(t) - y^*\|^2\right) +$$

$$\|\dot{x}(t)\|^2_{M_1(t)+\frac{(1-\gamma)c}{4}A^*A-\frac{L_h}{4}I} + \|\dot{z}(t)\|^2_{M_2(t)+\frac{c}{4}I} + \frac{\gamma+1}{3c}\|\dot{y}(t)\|^2 +$$

$$\left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + (1-\gamma)\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 +$$

$$-\frac{1}{2}\|x(t) - x^*\|^2_{\dot{M}_1(t)} - \frac{1}{2}\|z(t) - Ax^*\|^2_{\dot{M}_2(t)} +$$

$$\frac{1}{L_h}\left\|\nabla h(x(t)) - \nabla h(x^*) + \frac{L_h}{2}\dot{x}(t)\right\|^2 \leq 0.$$

Analogously, if $L_h = 0$, i.e. $h \equiv 0$, we obtain the same inequality without the last term.

From here almost immediately we obtain that

$\lim_{t \longrightarrow +\infty}(\|x(t) - x^*\|^2_{M_1(t)+c(1-\gamma)A^*A} + \|z(t) - Ax^*\|^2_{M_2(t)+cI} + \frac{1}{c}\|y(t) - y^*\|^2) \in \mathbb{R}.$

which is nothing else but the first assumption of our continuous Opial Lemma applied in the product space $\mathcal{H} \times \mathcal{G} \times \mathcal{G}$ for the function $t \longrightarrow (x(t), z(t), y(t))$, for the map

$$W(t) = \left(M_1(t) + c(1-\gamma)A^*A, M_2(t) + cI, \frac{1}{c}I\right)$$

and $\mathcal{C}$ the set of saddle points of the Lagrangian l.

Further, $\dot{x}(t) \in L^2([0, +\infty), \mathcal{H})$, $\dot{z}(t)$, $\dot{y}(t) \in L^2([0, +\infty), \mathcal{G})$.

From here we get

$$\lim_{t \longrightarrow +\infty} \dot{x}(t) = \lim_{t \longrightarrow +\infty} \dot{z}(t) = \lim_{t \longrightarrow +\infty} \dot{y}(t) = 0.$$

It remained to show that every weak sequential cluster point of $t \longrightarrow (x(t), z(t), y(t))$ belongs to S.

Let $(\overline{x}, \overline{z}, \overline{y})$ a weak sequentially cluster point of the vector function $t \longrightarrow (x(t), z(t), y(t))$. Then, there exists a sequence $(s_n)_{n\geq 0}$ with $s_n \longrightarrow +\infty$ such that $(x(s_n), z(s_n), y(s_n))$ converges to $(\overline{x}, \overline{z}, \overline{y})$ in the weak topology of $\mathcal{H} \times \mathcal{G} \times \mathcal{G}$ as $n \longrightarrow +\infty$.

We apply Lemma ACS with

$$a_n = \dot{x}(s_n) + x(s_n),$$

$$a_n^* = -cA^*A(\dot{x}(s_n) + x(s_n)) - M_1(s_n)\dot{x}(s_n) +$$

$$cA^*z(s_n) - A^*y(s_n) - \nabla h(x(s_n)) + \nabla h(\dot{x}(s_n) + x(s_n))$$

and

$$b_n = \dot{z}(s_n) + z(s_n),$$

$$b_n^* = -c(\dot{z}(s_n) + z(s_n)) + cA(\gamma\dot{x}(s_n) + x(s_n)) - M_2(s_n)\dot{z}(s_n) + y(s_n)$$

We get

$$a_n \rightharpoonup \overline{x}$$

and

$$b_n^* \rightharpoonup \overline{y}.$$

$$-A^*\overline{y} - \nabla h(\overline{x}) \in \partial f(\overline{x}) \tag{13}$$

and

$$\overline{y} \in \partial g(A\overline{x}). \tag{14}$$

Further, since $Aa_n - b_n \rightharpoonup 0$ and $a_n \rightharpoonup \overline{x}$, $b_n \rightharpoonup \overline{z}$ we have

$$A\overline{x} = \overline{z}. \tag{15}$$

Consequently, $(\overline{x}, \overline{z}, \overline{y})$ is a saddle point of l.

Consequences

In case $M_1(t) = M_2(t) = 0$ for all $t \geq 0$ we have

Theorem

Consider the Problem (1) *and assume that* $\mathcal{C}$*, the set of saddle points of the Lagrangian* l *is nonempty. Assume further, that* $\gamma \in [0, 1)$ *and* $A^*A - \frac{L_h}{c(1-\gamma)} I \in S_+(\mathcal{H})$ *(or, if* $h \equiv 0$*,* $A^*A \in P_\alpha(\mathcal{H})$ *for some* $\alpha > 0$*).*

For a starting point $(x_0, z_0, y_0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$*, let* $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ *be the unique strong global solution of the dynamical system obtained from* (4)*. Then, the vector function* $t \longrightarrow (x(t), z(t), y(t))$ *converges weakly to a saddle point of* l *as* $t \longrightarrow +\infty$*.*

Let us consider now the particular case

$$M_1(t) = \frac{1}{\tau(t)} I - cA^*A,\ M_2(t) = 0,\ \forall t \in [0, +\infty),$$

where $\tau(t) > 0$ for all $t \geq 0$ and $c\tau(t)\|A\|^2 \leq 1$.

Theorem

Consider the Problem (1) *and assume that $\mathcal{C}$, the set of saddle points of the Lagrangian l is nonempty. Assume further that the map $\frac{4-\tau(t)L_h}{4\tau(t)} I + \frac{5c-c\gamma}{4} A^*A \in S_+(\mathcal{H})$ for all $t \geq 0$ and τ is locally absolutely continuous, $\tau'(t) \geq 0$ for almost every $t \geq 0$ and $\sup_{t\geq 0} \frac{\tau'(t)}{\tau^2(t)} < +\infty$. Moreover, assume that one of the following assumptions hold.*

(I) $\frac{4-\tau(t)L_h}{4\tau(t)} I + \frac{5c-c\gamma}{4} A^*A \in P_{\alpha_1}(\mathcal{H})$ for all $t \geq 0$ and for some $\alpha_1 > 0$.

(II) $\gamma \in [0,1)$ and $A^*A \in P_\alpha(\mathcal{H})$ for some $\alpha > 0$.

For a starting point $(x_0, z_0, y_0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, let $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be the unique strong global solution of the dynamical system (10). Then, the vector function $t \longrightarrow (x(t), z(t), y(t))$ converges weakly to a saddle point of l as $t \longrightarrow +\infty$.

Thank you for your attention.