Relaxed-inertial proximal point algorithms for problems involving strongly quasiconvex functions

Sorin-Mihai Grad

ENSTA Paris / Polytechnic Institute of Paris
based on joint work with Felipe Lara & Raúl Marcavillaca

GdR MOA Annual Days 2022
November 11–14, 2022

- $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ proper lsc convex

$$\Rightarrow \mathrm{Prox}_h : \mathbb{R}^n \to \mathbb{R}^n$$

- h proper lsc but not convex

$$\Rightarrow \begin{cases} \mathrm{Prox}_h : \mathbb{R}^n \rightrightarrows \mathbb{R}^n \\ \text{no formulae for } \mathrm{Prox}_h \\ \text{PPA usually fails to converge (to a minimum of } h) \end{cases}$$

- PPA for quasiconvex problems (by means of Bregman distances): [Kaplan & Tichatschke, JoGO, 1998], [Langenberg & Tichatschke, JoGO, 2012], [Papa Quiroz, Mallma Ramirez & Oliveira, EJOR, 2015] etc
- convergence of the iterates towards stationary points

- PPA for minimizing strongly quasiconvex functions: [Lara, JOTA, 2022]

- $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ with a convex domain is
  - *strongly convex*: $\exists \gamma \in]0, +\infty[$ s.t. $\forall x, y \in \mathrm{dom}\ h\ \forall \lambda \in [0,1]$
    $$h(\lambda y + (1-\lambda)x) \leq \lambda h(y) + (1-\lambda)h(x) - \lambda(1-\lambda)\frac{\gamma}{2}\|x-y\|^2$$
  - *strongly quasiconvex*: $\exists \gamma \in]0, +\infty[$ s.t. $\forall x, y \in \mathrm{dom}\ h$ $\forall \lambda \in [0,1]$
    $$h(\lambda y + (1-\lambda)x) \leq \max\{h(y), h(x)\} - \lambda(1-\lambda)\frac{\gamma}{2}\|x-y\|^2$$
- both properties can be considered on a set $U \subseteq \mathbb{R}^n$, too
- strongly convex $\Rightarrow$ strongly quasiconvex
- $\|\cdot\|$ is strongly quasiconvex on any bounded convex $U \subseteq \mathbb{R}^n$, but *not* strongly convex
- $\sqrt{\|\cdot\|}$ is strongly quasiconvex on any bounded convex $U \subseteq \mathbb{R}^n$, but *not* convex
- any constant function is convex but *not* strongly quasiconvex

- [Bauschke & Combettes] a strongly quasiconvex function has at most one minimizer on a convex set that touches its domain
- [Lara, JOTA, 2022] a proper lsc strongly quasiconvex function has one minimizer on a closed convex subset of its domain
- the maximum of finitely many strongly quasiconvex functions with moduli $t_i > 0$ is strongly quasiconvex with modulus $\min\{t_i\} > 0$
- [Lara, JOTA, 2022] $t \mapsto \sqrt[4]{t^2 + k^2}$ $(k \in \mathbb{R})$ is strongly quasiconvex on any interval $[-c, c] \subseteq \mathbb{R}$

$$\min_{x \in K} h(x)$$

- $K \subseteq \mathbb{R}^n$ linear subspace
- $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ proper lsc, $K \subseteq \operatorname{dom} h$ and

  (A) h strongly quasiconvex on K

  $(B1)$ h quasiconvex on K

  $(B2)$ h is 2-weakly coercive on K:

$$\liminf_{x \in K,\, \|x\| \to +\infty} \frac{h(x)}{\|x\|^2} \geq 0$$

- $(A) \Rightarrow (B1)\&(B2)$, however a constant function satisfies $(B1)\&(B2)$, but not (A)

Step 0. let $x^0 = x^{-1} \in K$, $\alpha \in [0, 1[$, $0 < \rho' \leq \rho'' < 2$, $\{c_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}_{++}$, $k = 0$

Step 1. choose $\alpha_k \in [0, \alpha]$, set

$$y^k = x^k + \alpha_k(x^k - x^{k-1})$$

and compute

$$z^k \in \mathrm{Prox}_{c_k(h+\delta_K)}(y^k)$$

Step 2. if $z^k = y^k$: STOP $\Rightarrow y^k \in \arg\min_K h$

Step 3. choose $\rho_k \in [\rho', \rho'']$ and update

$$x^{k+1} = (1 - \rho_k)y^k + \rho_k z^k$$

Step 4. $k = k + 1$ and go to Step 1

- the algorithm was proposed in the convex framework: [Attouch & Cabot, Optimization, 2020]
- if $\alpha = 0$ & $\rho_k = 1$, $k \geq 0$ the algorithm collapses to PPA ([Lara, JOTA, 2022] for the (strongly) quasiconvex framework)
- it is necessary to take K linear subspace to guarantee that y^k is feasible (inertial step) and (if $\rho'' > 1$) that x^{k+1} is feasible (relaxation step)
- if $\alpha = 0$ & $\rho'' \leq 1$: K can be taken closed convex
- the proximity operator restricted to a set has already been considered: [Boț & Csetnek, Opt, 2017], [Gribonval & Nikolova, JMIV, 2020], [Yen & Muu, arXiv, 2021] etc

- h strongly quasiconvex on K
- $0 < \rho' \leq \rho'' < 2$, $\{\rho_k\}_k \subseteq [\rho', \rho'']$, $\alpha \in [0, 1[$, $\{\alpha_k\}_k \subseteq [0, \alpha]$
- $\sum_{k=0}^{\infty} \alpha_k \|x^k - x^{k-1}\|^2 < +\infty$
- $\Omega := \{x \in K : h(x) \leq h(z^k) \ \forall k \in \mathbb{N}\}$

$$\Rightarrow$$

- $\forall x^* \in \Omega \ \exists \lim_{k\to\infty} \|x^k - x^*\|$ and

$$\lim_{k\to+\infty} \|x^{k+1} - y^k\| = \lim_{k\to+\infty} \|z^k - y^k\| = 0$$

- if in addition $c_k \geq c' > 0 \ \forall k \geq 0$

$$\Rightarrow \begin{cases} x^k \to \overline{x} = \arg\min_K \ h \\ \lim\limits_{k\to+\infty} h(x^k) = \min_K h \end{cases}$$

- $\sum_{k=0}^{\infty} \alpha_k \|x^k - x^{k-1}\|^2 < +\infty$ is fulfilled when
  - $\{\alpha_k\}_k$ is nondecreasing satisfying (for a $\beta < 1$)

    $$0 \leq \alpha_k \leq \alpha_{k+1} \leq \alpha < \beta \ \forall \ k \geq 0$$

    and

    $$\rho'' = \rho''(\beta, \rho') := \frac{2\rho'(\beta^2 - \beta + 1)}{2\rho'\beta^2 + (2 - \rho')\beta + \rho'}$$

  - $\{\alpha_k\}_k$ is nondecreasing satisfying

    $$0 \leq \alpha_k \leq \alpha_{k+1} \leq \alpha < \frac{1}{3} \ \forall \ k \geq 0$$

- h quasiconvex and 2-weakly coercive on K
- $\Omega \neq \emptyset$
- $+$ previous hypotheses on sequences

$$\Rightarrow$$

- $\forall x^* \in \Omega \ \exists \lim_{k\to\infty} \|x^k - x^*\|$ and

$$\lim_{k\to+\infty} \|x^{k+1} - y^k\| = \lim_{k\to+\infty} \|z^k - y^k\| = 0$$

- if in addition h is bounded from below and $c_k \geq c' > 0 \ \forall k \geq 0$

$\Rightarrow \{h(x^k)\}_k$ is convergent

- $q \in \mathbb{N}$
- $K = \mathbb{R}^n$
- $h_1, h_2 : \mathbb{R}^n \to \mathbb{R}$, $h_1(x) = \sqrt{\|x\|}$ and $h_2(x) = \|x\|^2 - q$
- $h : \mathbb{R}^n \to \mathbb{R}$, $h(x) := \max\{h_1(x), h_2(x)\}$ is continuous and strongly quasiconvex (but not convex)
- $\arg\min_{\mathbb{R}^n} h = \{0\}$

- $n = 5$: $\varepsilon = 10^{-6}$, $q = 133$, $x^0 = [7, -8, 5, 2, 55]^\top$, $\rho' = 0.9$, $\rho'' = 1.5$, $\alpha = 0.125$, $c_1 = 1$, $c_{k+1} = 100/k^2 + c_k$, $\alpha_{k+1} = \alpha_k + 1/(900(k+1)^2)$, $\rho_k = (1 - 1/k)\rho' + (1/k)\rho''$, $k \geq 0$
- to reach $\bar{x}$ with error ε: 11 iterations / 0.9306 seconds
- PPA: 43 iterations / 0.9885 seconds

- $n = 50$: x^0 random, similar constellation
- to reach $\bar{x}$ with error ε: 13 iterations / 1.1360 seconds
- PPA: 46 iterations / 1.5866 seconds

	$n=5,\ \varepsilon=10^{-6}$		$n=50,\ \varepsilon=10^{-6}$		$n=500,\ \varepsilon=10^{-3}$	
	Alg	PPA	Alg	PPA	Alg	PPA
s	0.9306	0.9885	1.1360	1.5866	4.9289	9.4276
it	11	43	13	46	23	43

[running time (in seconds) and number of iterations performed by our algorithm and PPA to reach $\|z^k - y^k\| < \varepsilon$]

	$\varepsilon = 10^{-4}$		$\varepsilon = 10^{-5}$		$\varepsilon = 10^{-4} (q = 25)$	
	Alg	PPA	Alg	PPA	Alg	PPA
s	15.5450	24.7519	16.3472	294.0988	7.8103	14.6288
it	63	93	68	1105	14	59

[performance evaluation of our algorithm and PPA to reach $\|z^k - y^k\| < \varepsilon$ for $n = 500$]

$$\text{find } \overline{x} \in K : \quad f(\overline{x}, y) \geq 0, \ \forall \ y \in K$$

- solution set: $S(K, f)$
- $K \subseteq \mathbb{R}^n$ linear subspace
- $f : K \times K \to \mathbb{R}$
- $f(x, \cdot)$ strongly quasiconvex $\forall x \in K$
- $f(\cdot, y)$ usc $\forall y \in K$
- f (jointly) lsc and pseudomonotone on K, i.e.

$$f(x, y) \geq 0 \implies f(y, x) \leq 0 \ \forall x, y \in K$$

- f satisfies the Lipschitz type condition $\exists \eta > 0$ s.t.

$$f(x, z) - f(x, y) - f(y, z) \leq \eta \left(\|x - y\|^2 + \|y - z\|^2 \right) \ \forall \ x, y, z \in K$$

- $(f(x, x) = 0 \ \forall x \in K)$

- existence & uniqueness results for $f(x, \cdot)$ lsc (strongly) quasiconvex $\forall x \in K$ [Iusem & Lara, JOTA, 2019 / 2022]

Step 0. let $x^0, x^{-1} \in K$, $\alpha, \rho \in [0,1[$ and $\{\beta_k\}$, $k=0$

Step 1. choose $\alpha_k \in [0,\alpha]$, set

$$y^k = x^k + \alpha_k(x^k - x^{k-1})$$

and compute

$$z^k \in \arg\min_{x \in K} \left\{ f(y^k, x) + \frac{1}{2\beta_k} \|y^k - x\|^2 \right\}$$

Step 2. if $z^k = y^k$: STOP $\Rightarrow S(K,f) = \{y^k\}$

Step 3. choose $\rho_k \in [1-\rho, 1+\rho]$ and update

$$x^{k+1} = (1-\rho_k)y^k + \rho_k z^k$$

Step 4. $k = k+1$ and go to Step 1

- the algorithm was proposed in the convex framework: [Hieu, Duong & Thai, JANO, 2021], [Van Vinh, Tran & Vuong, NA, 2022]
- if $\alpha = 0$ & $\rho_k = 1$, $k \geq 0$ the algorithm collapses to PPA ([Iusem & Lara, JOTA, 2022] for the (strongly) quasiconvex framework)
- it is necessary to take K linear subspace to guarantee that y^k is feasible (in the inertial step) and (if $\rho_k > 1$) that x^{k+1} is feasible (relaxation step)
- the hypotheses guarantee that the solution set $S(K, f)$ is a singleton
- if $\alpha = 0$ & $\rho_k \leq 1$: K can be taken closed convex

- $\alpha \in [0, 1[$, $\{\alpha_k\}_k \subseteq [0, \alpha]$
- $\sum_{k=0}^{\infty} \alpha_k \|x^k - x^{k-1}\|^2 < +\infty$
- $\frac{1}{\gamma - 8\eta} < \beta_k < \epsilon \leq \frac{1}{4\eta}\ \forall k \geq 0$
- $0 < 1 - \rho \leq \rho_k \leq 1 + \rho$ with $0 \leq \rho \leq 1 - 4\eta\epsilon\ \forall k \geq 0$

$$\Rightarrow$$

- if $\{\overline{x}\} = S(K, f)\ \exists \lim_{k\to\infty} \|x^k - \overline{x}\|$ and

$$\lim_{k\to+\infty} \|x^{k+1} - y^k\|^2 = \lim_{k\to+\infty} \|z^k - y^k\|^2 = 0$$

- $\{x^k\}_k$, $\{y^k\}_k$, $\{z^k\}_k$ converge all to $\bar{x} \in S(K, f)$

- $q \in \mathbb{N}$
- $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$f(x, y) := \max\{\sqrt{|y|}, y^2 - q\} - \max\{\sqrt{|x|}, x^2 - q\} + x(y - x)$$

- $K = \mathbb{R}$ or $K = [0, +\infty[$
- f monotone fulfilling the hypotheses
- $f(x, \cdot)$ lsc strongly quasiconvex $\forall x \in K$
- $S([0, +\infty[, f) = \{0\}$
- $S(\mathbb{R}, f)$ not easy to determine "by hand"

- $K = \mathbb{R}$, $q = 99$, $\epsilon = 1/4$, $\eta = 1/2$, $\rho = 2/5$, $\varepsilon = 10^{-7}$, $x^0 = 17$, $x^{-1} = 10$, $\alpha = 1/29 - \varepsilon$, $\alpha_k = \alpha - 1/(29 + k)$, $\beta_k = 1/(k + 4)$, $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$, $k \geq 0$
- to reach $\bar{x} = -10.1084$ with error ε: 139 iterations / 2.5369 seconds
- PPA: 373 iterations / 5.5698 seconds

- $\varepsilon = 10^{-9}$: 179 iterations / 4.1559 seconds
- PPA: 549 iterations / 14.1292 seconds

- $K = [0, +\infty[$, $\varepsilon = 10^{-7}$: 40 iterations / 0.8841 seconds
- PPA: 41 iterations / 0.9481 seconds

- $K = \mathbb{R}$, $q = 50$, $\epsilon = 1/3$, $\eta = 2/3$, $\alpha = 1/200$, $\rho = 1/10$, $\varepsilon = 10^{-7}$, $\alpha_k = \alpha - 1/(200 + k)$, $\beta_k = 1/(k + 3)$, $\rho_k = (1/k)(1 - \rho) + (1 - 1/k)(1 + \rho)$, $k \geq 0$
- to reach $\bar{x} = -7.2591$ with error ε: 84 iterations / 1.7391 seconds
- PPA: 101 iterations / 2.7640 seconds
- $\rho_k = (1 - 1/k)(1 - \rho) + (1/k)(1 + \rho)$, $k \geq 0$: 136 iterations / 2.2756 seconds

- $\varepsilon = 10^{-9}$: 91 iterations / 1.8548 seconds
- PPA: 126 iterations / 2.7514 seconds

- $K = [0, +\infty[$, $\varepsilon = 10^{-7}$: 24 iterations / 0.4643 seconds
- PPA: 20 iterations / 0.5523 seconds
- $\rho_k = (1 - 1/k)(1 - \rho) + (1/k)(1 + \rho)$, $k \geq 0$: 28 iterations / 0.4798 seconds

- S.-M. Grad, F. Lara, R. T. Marcavillaca: *Relaxed-inertial proximal point type algorithms for nonconvex pseudomonotone equilibrium problems*, in revision
- S.-M. Grad, F. Lara, R. T. Marcavillaca: *Relaxed-inertial proximal point type algorithms for quasiconvex minimization*, JoGO, DOI: 10.1007/s10898-022-01226-z
- A.N. Iusem, F. Lara: *Proximal point algorithms for quasiconvex pseudomonotone equilibrium problems*, JOTA 193:443–461, 2022
- F. Lara: *On strongly quasiconvex functions: existence results and proximal point algorithms*, JOTA 192:891–911, 2022