△ Problem setting

$$\min_{x\in X} \hat{p}(x) := p(x) + h(x)$$

where h is a proper lower-semicontinuous convex function

X is a nonempty convex set

$p(x) := \max_{y\in Y} \Phi(x,y)$, $\forall x \in X$

Y is a nonempty compact convex set

Φ, for some scalar $m>0$ and open set $\Omega \supset X$, is such that (i) $\Phi \in C(\Omega \times Y)$ (ii) $-\Phi(x,\cdot): Y \to \mathbb{R}$ is lower-semicontinuous and convex for every $x\in X$ (iii) for every $y \in Y$, $\Phi(\cdot,y) + \frac{m}{2}\|\cdot\|^2$ is convex, differentiable and its gradient is Lipschitz continuous on $X\times Y$

△ Smooth approximation

$$p_\xi(x) := \max_{y\in Y}\left\{\Phi_\xi(x,y) := \Phi(x,y) - \frac{1}{2\xi}\|y-y_0\|^2\right\} \quad \forall x \in X$$

for some $y_0 \in Y$

△ Algorithms

(i) ACG (accelerated composite gradient)

Input: $(\mu, L) \in \mathbb{R}^2_{++}$, a function pair (ψ_n, ψ_s), an initial point $z_0 \in \operatorname{dom}\psi_n$

(0) set $y_0 = z_0$, $A_0 = 0$, $\Gamma_0 \equiv 0$ $\to (\alpha_0 = \beta_0 = 0)$ and $j=0$

(1) $A_{j+1} = A_j + \frac{1}{2L}\left(\mu A_j + 1 + \sqrt{(\mu A_j+1)^2 + 4L(\mu A_j+1)A_j}\right)$

$$\tilde{z}_j = \frac{A_j}{A_{j+1}} z_j + \frac{A_{j+1}-A_j}{A_{j+1}} y_j$$

$$\Gamma_{j+1}(y) = \alpha_{j+1} + \langle y, \beta_{j+1}\rangle, \forall y$$

where
$$\begin{cases} \alpha_{j+1} = \frac{A_j}{A_{j+1}}\alpha_j + \frac{A_{j+1}-A_j}{A_{j+1}}\left[\psi_s(\tilde{z}_j) - \langle \nabla\psi_s(\tilde{z}_j), \tilde{z}_j\rangle\right] \\ \beta_{j+1} = \frac{A_j}{A_{j+1}}\beta_j + \frac{A_{j+1}-A_j}{A_{j+1}}\nabla\psi_s(\tilde{z}_j) \end{cases}$$

$$y_{j+1} = \underset{y}{\arg\min} \left\{ \Gamma_{j+1}(y) + \psi_n(y) + \frac{1}{2A_{j+1}} \|y - y_0\|^2 \right\} \rightarrow \text{needs extra care}$$

$$z_{j+1} = \frac{A_j}{A_{j+1}} z_j + \frac{A_{j+1} - A_j}{A_{j+1}} y_{j+1}$$

(2) $u_{j+1} = (y_0 - y_{j+1}) / A_{j+1}$

$$\varepsilon_{j+1} = \psi(z_{j+1}) - \Gamma_{j+1}(y_{j+1}) - \psi_n(y_{j+1}) - \langle u_{j+1}, z_{j+1} - y_{j+1} \rangle$$

(3) $j = j+1$ and go to (1)

(ii) AIPP (Accelerated inexact proximal point method)

Input: a function pair (f, h), $(m, M) \in \mathbb{R}^2_{++}$ satisfying (P2), $\lambda \in (0, \frac{1}{2m}]$, $\sigma \in (0,1)$, an initial point $x_0 \in \operatorname{dom} h$, a tolerance $\bar{\rho} > 0$

Output: $(\bar{x}, \bar{u}) \in \operatorname{dom} h \times \mathcal{X}$ satisfying $\bar{u} \in \nabla f(\bar{x}) + \partial h(\bar{x})$, $\|\bar{u}\| \le \bar{\rho}$

(0) set $k=1$, $\hat{\rho} := \frac{\bar{\rho}}{4}$, $\hat{\varepsilon} := \bar{\rho}^2 / (32(M + \frac{1}{\lambda}))$, $M_\lambda := M + \frac{1}{\lambda}$

(1) call ACG with inputs $z_0 = x_{k-1}$, $(\mu, L) = (\frac{1}{2}, \lambda M + \frac{1}{2})$, $\psi_s = \lambda f + \frac{1}{4}\|\cdot - x_{k-1}\|^2$ and $\psi_n = \lambda h + \frac{1}{4}\|\cdot - x_{k-1}\|^2$ in order to obtain a triple $(x, u, \varepsilon) \in \mathcal{X} \times \mathcal{X} \times \mathbb{R}_+$ satisfying $u \in \partial_\varepsilon (\lambda \phi + \frac{1}{2}\|\cdot - x_{k-1}\|^2)(x)$, $\underbrace{\|u\|^2 + 2\varepsilon \le \sigma \|x_{k-1} - x + u\|^2}_{\text{stop condition}}$ $(*)$

(2) if $\|x_{k-1} - x + u\| \le \lambda \hat{\rho} / 5$, then go to (3);
otherwise set $(x_k, \tilde{u}_k, \tilde{\varepsilon}_k) = (x, u, \varepsilon)$, $k = k+1$ and go to (1)

(3) restart the previous call to ACG in step 1 to find a triple $(\tilde{x}, \tilde{u}, \tilde{\varepsilon})$ such that $\tilde{\varepsilon} \le \hat{\varepsilon}\lambda$ and $(x, u, \varepsilon) = (\tilde{x}, \tilde{u}, \tilde{\varepsilon})$ satisfies $(*)$

(4) $\bar{x} := \underset{x' \in \mathcal{X}}{\arg\min} \left\{ \langle \nabla f(x), x' - x \rangle + h(x') + \frac{M_\lambda}{2} \|x' - x\|^2 \right\}$

$\bar{u} := M_\lambda (x - \bar{x}) + \nabla f(\bar{x}) - \nabla f(x)$

(iii) AIPP-S

Input: $(m, L_x, L_y) \in \mathbb{R}^3_{++}$ satisfying (A3), a smoothing constant $\xi > 0$
an initial point $(x_0, y_0) \in X \times Y$, a tolerance $\rho > 0$

Output: $(x,u) \in X \times \mathcal{X}$

(0) $L_\xi = L_y Q_\xi + L_x$, where $Q_\xi = \xi L_y + \sqrt{\xi(L_x+m)}$

$\sigma = \frac{1}{2}$, $\lambda = \frac{1}{4m}$

$p_\xi(x) := \max_{y \in Y} \{ \Phi_\xi(x,y) := \Phi(x,y) - \frac{1}{2\xi}\|y - y_0\|^2 \} \quad \forall x \in X$

(1) apply AIPP with inputs (m, L_ξ), (p_ξ, h), λ, σ, x_0 and ρ to obtain a pair (x,u) satisfying $u \in \nabla p_\xi(x) + \partial h(x)$, $\|u\| \le \rho$

(2) output the pair (x,u)