06.12.19 ①

115

AGD

$$\min_{x \in Q} f(x) + h(x) \qquad f(x) - L\text{-smooth}.$$

w.r.t.

$\| \cdot \|, \omega, V$ $\qquad$ $\| \cdot \| - \| \cdot \|_*$

proximal setup.

$$y_0 = u_0 = x_0$$

$$A_{k+1} = A_k + \alpha_{k+1} = L\alpha_{k+1}^2$$

$$y_{k+1} = \frac{\alpha_{k+1}}{A_{k+1}} u_k + \frac{A_k}{A_{k+1}} x_k.$$

$$u_{k+1} = \arg\min_{x \in Q} \Big\{ \alpha_{k+1} \big(f(y_{k+1}) + \langle \nabla f(y_{k+1}), x - y_{k+1} \rangle + h(x) \big) + V(x, u_k) \Big\}$$

$$x_{k+1} = \frac{\alpha_{k+1}}{A_{k+1}} u_{k+1} + \frac{A_k}{A_{k+1}} x_k$$

Lemma 1

$$u_+ = \arg\min_{x \in Q} \{ \psi(x) + V(x, u) \}$$

$$\psi(x) + V(x, u) \geq \psi(u_+) + V(u_+, u) + V(x, u_+) \qquad \forall x \in Q$$

$$\langle \nabla \psi(u_+) + \nabla_1 V(u_+, u), x - u_+ \rangle \geq 0 \quad \forall x \in Q$$

$$\psi(x) - \psi(u_+) \geq \langle \nabla \psi(u_+), x - u_+ \rangle \geq \langle \nabla_1 V(u_+, u), u_+ - x \rangle =$$

sm.

$$= V(u_+, u) + V(x, u_+) - V(x, u)$$

$$\psi(u_+) + V(u_+, u) \leq \psi(x) + V(x, u) - V(x, u_+)$$

06.12.19 (2)

L-smoothness

$$f(x_{k+1}) + h(x_{k+1}) \le f(y_{k+1}) + \langle \nabla f(y_{k+1}), x_{k+1} - y_{k+1} \rangle +$$

$$+ \frac{L_{k+1}}{2} \| x_{k+1} - y_{k+1} \|^2 + h(x_{k+1}) + \frac{\varepsilon \alpha_{k+1}}{2A_{k+1}}$$

$$= f(y_{k+1}) + \left\langle \nabla f(y_{k+1}), \frac{\alpha_{k+1}}{A_{k+1}} u_{k+1} + \frac{A_k}{A_{k+1}} x_k - y_{k+1} \right\rangle +$$

$$+ \frac{L_{k+1}}{2} \left\| \frac{\alpha_{k+1}}{A_{k+1}} u_{k+1} + \cancel{\frac{A_k}{A_{k+1}} x_k} - \frac{\alpha_{k+1}}{A_{k+1}} u_k - \cancel{\frac{A_k}{A_{k+1}} x_k} \right\|^2 \cancel{+ h(x_{k+1}) \le} + h(x_{k+1})$$

$$= f(y_{k+1}) + \frac{\alpha_{k+1}}{A_{k+1}} \langle \nabla f(y_{k+1}), u_{k+1} - y_{k+1} \rangle + \frac{A_k}{A_{k+1}} \langle \nabla f(y_{k+1}), x_k - y_{k+1} \rangle +$$

$$+ \left[\left(\frac{L_{k+1} \alpha_{k+1}^2}{A_{k+1}^2} \right) = \frac{1}{A_{k+1}} \right] \frac{1}{2A_{k+1}} \| u_{k+1} - u_k \|^2 + h(x_{k+1}) \le$$

$$\le \frac{A_k}{A_{k+1}} \left(f(y_{k+1}) + \langle \nabla f(y_{k+1}), x_k - y_{k+1} \rangle + h(x_k) \right) +$$

$$+ \frac{\alpha_{k+1}}{A_{k+1}} \left(f(y_{k+1}) + \langle \nabla f(y_{k+1}), u_{k+1} - y_{k+1} \rangle + \frac{1}{2\alpha_{k+1}} \| u_{k+1} - u_k \|^2 + h(u_{k+1}) \right) \le$$

convexity

$$\le \frac{A_k}{A_{k+1}} \left(f(x_k) + h(x_k) \right)$$

$$+ \frac{\alpha_{k+1}}{A_{k+1}} \left[\frac{1}{\alpha_{k+1}} \left(\underbrace{\alpha_{k+1} \left(f(y_{k+1}) + \langle \nabla f(y_{k+1}), u_{k+1} - y_{k+1} \rangle + h(u_{k+1}) \right)}_{\psi(u_{k+1}) \sim \psi(u_+)} + \underbrace{V(u_{k+1}, u_k)}_{V(u_+, u)} \right) \right] \le$$

By Lemma 1

$$\le \frac{A_k}{A_{k+1}} f(x_k) + \frac{\alpha_{k+1}}{A_{k+1}} \left(f(y_{k+1}) + \langle \nabla f(y_{k+1}), x - y_{k+1} \rangle + h(x) + \frac{1}{\alpha_{k+1}} V(x, u_k) - \frac{1}{\alpha_{k+1}} V(x, u_{k+1}) \right)$$

06.12.19 (3)

convexity

$$\leq \frac{A_k}{A_{k+1}}\left(f(x_k)+h(x_k)\right)+\frac{\alpha_{k+1}}{A_{k+1}}\left(f(x)+h(x)\right)+\frac{1}{A_{k+1}}V(x,u_k)-\frac{1}{A_{k+1}}V(x,u_{k+1})$$

Rearranging, we get:

$$A_{k+1}\left(f(x_{k+1})+h(x_{k+1})\right)-A_k\left(f(x_k)+h(x_k)\right)+V(x,u_{k+1})-V(x,u_k)\leq$$

$$\leq \alpha_{k+1}\left(f(x)+h(x)\right)$$

$$\sum_{k=0}^{N-1} \Rightarrow A_N F(x_N)-A_0 F(x_0)+V(x,u_N)-V(x,u_0)\leq (A_N-A_0)F(x)$$

Set $A_0=0$

$$A_N(\underbrace{F(x_N)-F(x^*)}_{\geq 0})+\cancel{A_0(\underbrace{f(x^*)-f(x_0)}_{\geq 0})}\leq V(x,u_0)-V(x,u_N)\leq$$

$$\leq V(x,u_0)$$

$$F(x_N)-F^*\leq \frac{1}{A_N}V(x,u_0)$$

$$A_k=\sum_{i=0}^{k}\alpha_i \quad = L\alpha_k^2$$

$A_0=0 \Rightarrow \alpha_1=L\alpha_1^2$ $\qquad A_1=\alpha_1=\frac{1}{L}$

$k\geq 2.$ $\qquad A_1\geq \frac{(1+1)^2}{4L}$

$$L\alpha_{k+1}^2=A_{k+1}=A_k+\alpha_{k+1}$$

$$\alpha_{k+1}=\frac{1+\sqrt{1+4LA_k}}{2L}$$

Induction Assumption: $A_k\geq\frac{(k+1)^2}{4L}$

$$\alpha_{k+1}=\frac{1}{2L}+\sqrt{\frac{1}{4L^2}+\frac{A_k}{L}}\geq\frac{1}{2L}+\sqrt{\frac{A_k}{L}}\geq$$

$$\geq\frac{1}{2L}+\frac{1}{\sqrt{L}}\cdot\frac{k+1}{2\sqrt{L}}\geq\frac{k+2}{2L}$$

$$A_{k+1}=A_k+\alpha_{k+1}\geq\frac{(k+1)^2}{4L}+\frac{k+2}{2L}=\frac{k^2+2k+1+2k+4}{4L}=$$

$$=\frac{(k+2)^2+1}{4L}\geq\frac{(k+2)^2}{4L}.$$

$$\boxed{F(x_N)-F^*\leq\frac{4LV(x^*,x_0)}{(k+1)^2}}$$

06.12.19 (4)

restart technique.

conv → str conv.

reduction. str − conv − conv.

$$f(x_N) - f^* \le \frac{L V(x^*, x_0)}{K^2}$$

$$\theta = \arg\min_{x \in Q} d(x) \quad (*)$$

$$d(0) = 0$$

$$d(x) \le \frac{\Omega}{2} \quad \forall x : \|x\| \le 1$$

Assume that we have $x_0 : \|x_0 - x^*\| \le R$

for $p \ge 0$ do.

make $N = 4\sqrt{\frac{L\Omega}{\mu}}$ steps

of AGD from x_p using prox

$d_p(x) = R_p^2 \, d\left(\frac{x - x_p}{R_p}\right)$, where $R_p = R \cdot 2^{-p}$

Set $x_{p+1} = \cancel{\ }\ x_N$

---

1) $\nabla^2 d_p(x) = R_p^2 / R_p^2 \, \nabla^2 d\left(\frac{x - x_p}{R_p}\right) \ge 1$

$$V_p(x, x_p) = d_p(x) - \underbrace{\underbrace{d_p(x_p)}_{=0} - \underbrace{\langle \nabla d_p(x_p), x - x_p \rangle}_{\ge 0}}_{\text{since } (*)} \le d_p(x)$$

$\|x_0 - x^*\| \le R_0$

$$\frac{\mu}{2}\|x_N - x^*\|^2 \le f(x_N) - f^* \le \frac{L V_p(x^*, x_p)}{N^2} \le$$

$x_N = x_{p+1}$

$$\le \frac{L\, d_p(x^*)}{N_p^2} = \frac{L R_p^2}{N_p^2} d\left(\frac{x - x_p}{R_p}\right) \le \frac{L R_p^2 \Omega}{2 N_p^2}$$

$$\|x_{p+1} - x^*\|^2 \le \frac{L R_p^2 \Omega}{\mu N_p^2} \le R_{p+1}^2 = \frac{R_p^2}{4}$$

$$f(x_p) - f^* \le \cancel{\frac{\mu R_p^2}{8}} \quad \frac{\mu R_p^2}{2} = \frac{\mu R^2}{2} \cdot 2^{-2p} \le \varepsilon$$

in $\frac{1}{2}\left[\log_2 \frac{\mu R^2}{2\varepsilon}\right]$

06.12.19(5)

total number of oracle calls.

$$\geq \sqrt{\frac{L\Omega}{\mu}} \cdot \frac{1}{\gamma}\left\lceil \log_2 \frac{\mu R^2}{2\varepsilon} \right\rceil$$ corresponds to lower bound.

Euclidean setup $\Rightarrow \Omega = 1$.

---

f - convex.

We have a method for strongly convex functions.

Say, AGD.

$f_\mu(x) = f(x) + \frac{\mu}{2}\|x - x_0\|^2$ - (L, μ).

in $O\left(\sqrt{\frac{L}{\mu}} \left\lceil \log_2 \frac{\mu}{\varepsilon} \cdot \|x_0 - x_\mu^*\|^2 \right\rceil\right)$

Aplying AGD to f_μ: $f_\mu(\tilde{x}) - f_\mu(x_\mu^*) \leq \frac{\varepsilon}{2}$

$$f(\tilde{x}) + \frac{\mu}{2}\|\tilde{x} - x_0\|^2 - f(x_\mu^*) - \frac{\mu}{2}\|x_0 - x_\mu^*\|^2 \leq \varepsilon$$

$f(x^*) \leq f_\mu(x_\mu^*)$ $\qquad -f(x_\mu^*) \leq -f(x^*)$

$f_\mu(x_\mu^*) \leq f_\mu(x^*)$ $\qquad -f_\mu(x_\mu^*) \geq -f_\mu(x^*)$

$$f_\mu(\tilde{x}) - f_\mu(x^*) \leq f_\mu(\tilde{x}) - f_\mu(x_\mu^*) \leq \frac{\varepsilon}{2}$$

$$f(\tilde{x}) + \frac{\mu}{2}\|\tilde{x} - x_0\|^2 - f(x^*) - \frac{\mu}{2}\|x_0 - x^*\|^2 \geq f(\tilde{x}) - f^* - \frac{\mu}{2}\|x_0 - x^*\|^2$$

$$f(\tilde{x}) - f^* \leq \frac{\varepsilon}{2} + \frac{\mu}{2}\|x_0 - x^*\|^2 \leq \frac{\varepsilon}{2} + \frac{\mu R^2}{2} \qquad \mu = \frac{\varepsilon}{R^2}$$

$$\|x_0 - x_\mu^*\| \leq \|x_0 - x^*\|$$

~~$f(x^*) \leq f(x_\mu^*)$~~

$$f(x^*) + \frac{\mu}{2}\|x_0 - x_\mu^*\|^2 \leq f(x_\mu^*) + \frac{\mu}{2}\|x_0 - x_\mu^*\|^2 = f_\mu(x_\mu^*) \leq f_\mu(x^*) = f(x^*) + \frac{\mu}{2}\|x_0 - x^*\|^2$$

complexity $\sqrt{\frac{LR^2}{\varepsilon}}$

str conv $\rightarrow$ conv $\qquad \mu = \frac{\varepsilon}{R^2}$