# Exercice 1

1. $F$ $\mu$-convexe : $\begin{cases} F(\theta_0) - F(\theta) \geq \langle \nabla F(\theta), \theta_0 - \theta \rangle + \frac{\mu}{2} \|\theta - \theta_0\|^2 \\ F(\theta) - F(\theta_0) \geq \langle \nabla F(\theta_0), \theta - \theta_0 \rangle + \frac{\mu}{2} \|\theta - \theta_0\|^2 \end{cases}$

On combine : $0 \geq - \langle \nabla F(\theta) - \nabla F(\theta_0), \theta - \theta_0 \rangle + \mu \|\theta - \theta_0\|^2$

$\iff \langle \nabla F(\theta) - \nabla F(\theta_0), \theta - \theta_0 \rangle \geq \mu \|\theta - \theta_0\|^2$

2. On pose $G(\theta) := F(\theta) - \frac{\mu}{2} \|\theta - \theta_0\|^2$.

$G$ est :
- convexe
- $\|\nabla G(\theta) - \nabla G(\theta_0)\|_2^2 = \|\nabla F(\theta) - \nabla F(\theta_0) - \mu(\theta - \theta_0)\|_2^2$

$$= \|\nabla F(\theta) - \nabla F(\theta_0)\|_2^2 \underbrace{- 2\mu \langle \nabla F(\theta) - \nabla F(\theta_0), \theta - \theta_0 \rangle}_{\leq -2\mu \frac{1}{L} \|\nabla F(\theta) - \nabla F(\theta_0)\|_2^2 \text{ par cocoercivité de } F} + \mu^2 \|\theta - \theta_0\|_2^2$$

$$\leq \left(1 - 2\frac{\mu}{L}\right) L^2 \|\theta - \theta_0\|^2 + \mu^2 \|\theta - \theta_0\|^2$$

$$= (L^2 - 2\mu L + \mu^2) \|\theta - \theta_0\|^2$$

$$= (L - \mu)^2 \|\theta - \theta_0\|^2$$

$\|\nabla G(\theta) - \nabla G(\theta_0)\|_2^2 \leq (L - \mu) \|\theta - \theta_0\|$

donc $G$ est $(L-\mu)$-smooth, et $\nabla G$ est donc cocoercif.

$$\langle \nabla G(\theta) - \nabla G(\theta_0), \theta - \theta_0 \rangle \geq \frac{1}{L-\mu} \|\nabla G(\theta) - \nabla G(\theta_0)\|_2^2 \quad \text{par cocoercivité de } \nabla G.$$

$$\langle \nabla F(\theta) - \mu(\theta - \theta_0) - \nabla F(\theta_0), \theta - \theta_0 \rangle \geq \frac{1}{L-\mu} \|\nabla F(\theta) - \mu(\theta - \theta_0) - \nabla F(\theta_0)\|_2^2$$

$$\langle \nabla F(\theta) - \nabla F(\theta_0), \theta - \theta_0 \rangle \geq \mu \|\theta - \theta_0\|_2^2 + \frac{1}{L-\mu} \|\nabla F(\theta) - \nabla F(\theta_0)\|_2^2 + \frac{\mu^2}{L-\mu} \|\theta - \theta_0\|_2^2$$

$$- \frac{2\mu}{L-\mu} \langle \nabla F(\theta) - \nabla F(\theta_0), \theta - \theta_0 \rangle$$

$$\frac{L-\mu}{L+\mu} \langle \nabla F(\theta) - \nabla F(\theta_0), \theta - \theta_0 \rangle \geq \frac{\mu L - \mu^2 + \mu^2}{L-\mu} \|\theta - \theta_0\|_2^2 + \frac{1}{L-\mu} \|\nabla F(\theta) - \nabla F(\theta_0)\|_2^2$$

$$\langle \nabla F(\theta) - \nabla F(\theta_0), \theta - \theta_0 \rangle \geq \frac{\mu L}{L+\mu} \|\theta - \theta_0\| + \frac{1}{L+\mu} \|\nabla F(\theta) - \nabla F(\theta_0)\|_2^2$$

3. Recall that $\theta_{k+1} = \theta_k - \gamma \nabla F(\theta_k)$

Ici, $\gamma = \frac{2}{\mu+L}$

$$\|\theta_{k+1} - \theta^*\|_2^2 = \|\theta_k - \gamma \nabla F(\theta_k) - \theta^*\|_2^2$$

$$= \|\theta_k - \theta^*\|^2 + \gamma^2 \|\nabla F(\theta_k)\|^2 - 2\gamma \langle \nabla F(\theta_k), \theta_k - \theta^* \rangle$$

$$= \|\theta_k - \theta^*\|^2 + \frac{4}{(\mu+L)^2} \|\nabla F(\theta_k) - \nabla F(\theta^*)\|^2 \underbrace{- \frac{4}{\mu+L} \langle \nabla F(\theta_k) - \nabla F(\theta^*), \theta_k - \theta^* \rangle}_{\leq - \frac{4\mu L}{(\mu+L)^2} \|\theta_k - \theta^*\|^2 - \frac{4}{(\mu+L)^2} \|\nabla F(\theta_k) - \nabla F(\theta^*)\|^2}$$

$$\leq \frac{\mu^2 + 2\mu L + L^2 - 4\mu L}{(\mu+L)^2} \|\theta_k - \theta^*\|^2$$

$$= \frac{\mu^2 - 2\mu L + L^2}{(\mu+L)^2} \|\theta_k - \theta^*\|^2$$

$$= \left(\frac{L-\mu}{\mu+L}\right)^2 \|\theta_k - \theta^*\|^2$$

$$= \left(\frac{L(1-\kappa)}{L(1+\kappa)}\right)^2 \|\theta_k - \theta^*\|^2$$

$$= \left(\frac{1-\kappa}{1+\kappa}\right)^2 \|\theta_k - \theta^*\|^2$$

4. $\|\theta_{k+1} - \theta^*\|^2 \leq \left(\frac{1-\kappa}{1+\kappa}\right)^2 \|\theta_k - \theta^*\|$

Exercice 2 cvx :

1. $\delta_{t-1} - \delta_t = F(\theta_{t+1}) - F^* - F(\theta_t) + F^*$

$$= \underbrace{F(\theta_t + d_t - \tfrac{1}{L}\nabla F(\beta_t))} - F(\theta_t)$$

$$\leq F(\theta_t + d_t) + \langle \nabla F(\theta_t + d_t), \cancel{\theta_t + d_t} - \tfrac{1}{L}\nabla F(\theta_t + d_t) - \cancel{\theta_t - d_t} \rangle + \tfrac{L}{2}\|\tfrac{1}{L}\nabla F(\theta_t + d_t)\|^2$$

$$\leq -\tfrac{1}{L}\|\nabla F(\theta_t + d_t)\|^2 + \tfrac{1}{2L}\|\nabla F(\theta_t + d_t)\|^2 + F(\theta_t + d_t) - F(\theta_t)$$

$$= \left(\tfrac{1}{2L} - \tfrac{1}{L}\right)\|\nabla F(\theta_t + d_t)\|^2 + F(\theta_t + d_t) - F(\theta_t)$$

$$= \left(\tfrac{L\gamma^2}{2} - \gamma\right)\|\nabla F(\theta_t + d_t)\|^2 + F(\theta_t + d_t) - F(\theta_t)$$

2. $\delta_{t+1} - \delta_t \leq \left(\frac{L\gamma^2}{2} - \gamma\right)\| L g_t \|_2^2 + F(\theta_t + d_t) - F(\theta_t)$

$$= \tfrac{L^2}{2}\left(\tfrac{1}{L} - \tfrac{2}{L}\right)\|g_t\|_2^2 + \underbrace{F(\theta_t + d_t) \underbrace{- F(\theta_t)}_{\leq -F(\theta^*)}}_{\leq -\langle \nabla F(\theta_t + d_t), \theta^* - (\theta_t + d_t) \rangle \text{ (cvx.)}}$$

$$\leq -\tfrac{L}{2}\|g_t\|_2^2 + \langle g_t, \theta_t + d_t - \theta^* \rangle$$

$$= -\tfrac{L}{2}\left(\|g_t\|_2^2 - 2\langle g_t, \theta_t + d_t - \theta^* \rangle\right)$$

3. $\delta_{t+1} = F(\theta_{t+1}) - F^*$

$\leq -\frac{1}{2L}\|g_t\|_2^2 + F(\theta_t + d_t) - F^*$ en reprenant les calculs de (1), (2)

$\leq -\frac{L}{2}\|g_t\|_2^2 + \langle g_t, \theta_t + d_t - \theta^* \rangle$ par les m̂ calculs qu'au (2) dûs à la convexité de F

$= -\frac{L}{2}\left(\|g_t\|_2^2 - 2\langle g_t, \theta_t + d_t - \theta^* \rangle\right)$

4. $(\lambda_t - 1)(\delta_{t+1} - \delta_t) + \delta_{t+1} \leq -(\lambda_t - 1)\frac{L}{2}\left(\|g_t\|_2^2 - 2\langle g_t, \theta_t + d_t - \theta^* \rangle\right) - \frac{L}{2}\left(\|g_t\|_2^2 - 2\langle g_t, \theta_t + d_t - \theta^* \rangle\right)$

$$= -\tfrac{L}{2}\lambda_t\left(\|g_t\|_2^2 - 2\langle g_t, \theta_t + d_t - \theta^* \rangle\right)$$

$$= -\tfrac{L}{2}\lambda_t\left(\tfrac{1}{\lambda_t^2}\|\lambda g_t\|_2^2 - 2\langle \lambda_t g_t, \lambda_t\theta_t + \lambda_t d_t - \lambda_t\theta^* \rangle\right)$$

$=$