Victor-Emmanuel Brunel — *Sigalla & Cloarec*

Notes optimisation avancée

Théorème 0.1. *Soit f une fonction réelle α-fortement convexe et L-lipschitzienne sur E. Soit $x^\star := \arg\min_{x\in E} f(x)$. On a le résultat suivant:*

$$f(\hat{x}) - f(x^\star) \leq \frac{2L^2}{\alpha T} \tag{1}$$

Démonstration. Pour $t = 1$ à T, introduisons le pas:

$$\mu_t := \frac{2}{\alpha t} \tag{2}$$

En utilisant l'identité classique $2\langle x\,,\, y\rangle = \left\|x+y\right\|^2 - \|x\|^2 - \left\|y\right\|^2$, on a pour $t = 0$ à $T-1$:

$$\begin{aligned}
f(\hat{x}) - f(x^\star) &\leq f(x_t) - \left(f(x_t) + u_t^T\left(x^\star - x_t\right) + \frac{\alpha}{2}\left\|x^\star - x_t\right\|^2 \right) \\
&= -u_t^T\left(x^\star - x_t\right) - \frac{\alpha}{2}\left\|x^\star - x_t\right\|^2 \\
&= -\frac{\mu_{t+1}}{\mu_{t+1}} \cdot u_t^T\left(x^\star - x_t\right) - \frac{\alpha}{2}\left\|x^\star - x_t\right\|^2 \\
&= \frac{1}{2\mu_{t+1}}\left(\left\|\mu_{t+1}\cdot u_t\right\|^2 + \left\|x^\star - x_t\right\|^2 - \left\|\mu_{t+1}\cdot u_t + x^\star - x_t\right\|^2 \right) - \frac{\alpha}{2}\left\|x^\star - x_t\right\|^2 \\
&= \frac{1}{2\mu_{t+1}}\left(\left\|\mu_{t+1}\cdot u_t\right\|^2 + \left\|x^\star - x_t\right\|^2 - \left\|x^\star - y_t\right\|^2 \right) - \frac{\alpha}{2}\left\|x^\star - x_t\right\|^2 \\
&\leq \frac{1}{2\mu_{t+1}}\left(\mu_{t+1}^2\cdot L^2 + \left\|x^\star - x_t\right\|^2 - \left\|x^\star - x_{t+1}\right\|^2 \right) - \frac{\alpha}{2}\left\|x^\star - x_t\right\|^2 \\
&= \frac{\mu_{t+1}L^2}{2} + \underbrace{\left(\frac{1}{2\mu_{t+1}} - \frac{\alpha}{2}\right)}_{\frac{\alpha(t-1)}{4}}\left\|x^\star - x_t\right\|^2 - \underbrace{\frac{1}{2\mu_{t+1}}}_{\frac{\alpha(t+1)}{4}}\left\|x^\star - x_{t+1}\right\|^2
\end{aligned}$$

Travaillons sur les deux termes de droites afin de faire apparaître une somme télescopique. □