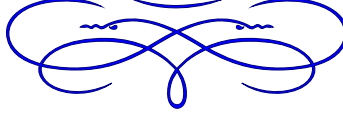


## Notes optimisation avancée



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

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

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

$$\mu_t := \frac{2}{\alpha t} \quad (2)$$

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

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

En utilisant la définition de  $\hat{x}$ , la convexité de  $f$  et l'inégalité d'au-dessous pour tout  $t$ , on a successivement:

$$\begin{aligned} f(\hat{x}) - f(x^\star) &= f\left( \sum_{t=0}^T \frac{2t}{T(T+1)} x_t \right) - f(x^\star) \\ &\leq \sum_{t=0}^T \frac{2t}{T(T+1)} (f(x_t) - f(x^\star)) \\ &\leq \sum_{t=0}^T \frac{2t}{T(T+1)} \left( \frac{\mu_{t+1} L^2}{2} + \frac{\alpha(t-1)}{4} \cdot \|x^\star - x_t\|^2 - \frac{\alpha(t+1)}{4} \cdot \|x^\star - x_{t+1}\|^2 \right) \\ &\leq \underbrace{\sum_{t=0}^T \frac{t}{T(T+1)} \mu_{t+1} L^2}_{:=A} + \underbrace{\sum_{t=0}^T \frac{\alpha t}{2T(T+1)} \left( (t-1) \cdot \|x^\star - x_t\|^2 - (t+1) \cdot \|x^\star - x_{t+1}\|^2 \right)}_{:=B} \end{aligned}$$

En posant  $\delta_t := t(t-1) \cdot \|x^\star - x_t\|^2$ , on peut réécrire la somme de gauche:

$$B = \frac{\alpha}{2T(T+1)} \sum_{t=0}^T (\delta_t - \delta_{t+1}) = \frac{\alpha}{2T(T+1)} (\delta_0 - \delta_{T+1}) \leq 0$$

Montrons maintenant que  $A \leq \frac{2L^2}{\alpha T}$ . En utilisant la définition du pas  $\mu_t$ , on a:

$$A = \frac{L^2}{T(T+1)} \sum_{t=0}^T \frac{2t}{\alpha(t+1)} = \frac{2L^2}{\alpha T} \cdot \frac{1}{(T+1)} \sum_{t=0}^T \underbrace{\frac{t}{t+1}}_{\leq 1} \leq \frac{2L^2}{\alpha T}$$

d'où le résultat. □

Un exemple très important est celui des fonctions  $\alpha$ -fortement convexe et  $\beta$  régulière. Remarquons que nécessairement  $\alpha \leq \beta$ .

## 1 Cas de fonctions $\alpha$ -fortement convexe et $\beta$ régulière

Soit  $f : \mathbb{R}^d \mapsto \mathbb{R}$  une fonction  $\alpha$ -fortement convexe et  $\beta$  régulière. On a donc pour tout  $x, x_0$  appartenant à  $\mathbb{R}^d$ :

**$\beta$ -régularité**  $f(x) \leq f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{\beta}{2} \|x - x_0\|^2$

**$\alpha$ -fortement convexe**  $f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{\alpha}{2} \|x - x_0\|^2 \leq f(x)$