TD3 : algorithme d'optimisation locale sans contraintes

Ex 1, 2, 3 : voir cours chapitre 2

Ex 4) J γ-convexe
∇J Lipschitz

$$x_{k+1} = x_k - \alpha_k \nabla J(x_k) \text{ avec}$$

$$q(\alpha_k) = \min_{\alpha \in \mathbb{R}_+^*} q(\alpha) \text{ où}$$

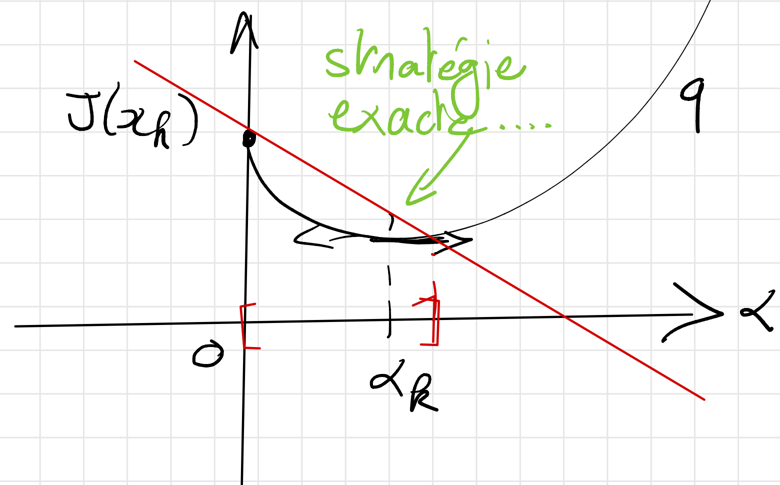
$$q(\alpha) = J(x_k - \alpha \nabla J(x_k))$$

1) Existence de (x_k) ?

$J(x_{k+1}) < J(x_k)$?

$q'(0) < 0$ et $q(\alpha) \underset{\alpha \to +\infty}{\longrightarrow} +\infty$ (J coercive avec γ convexité)

$\Rightarrow$ q admet un minimum et $q(\alpha_k) \leq q(0)$. Si égalité, J ne peut être strictement convexe.



2) $q'(\alpha) = -\langle \nabla J(x_k), \nabla J(x_k - \alpha \nabla J(x_k))\rangle$

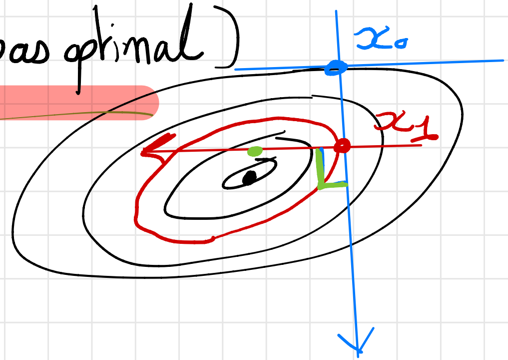
$\langle \nabla J(x_{k+1}), \nabla J(x_k)\rangle = 0$?

(deux directions de descente consécutives sont orthogonales)

* En effet, comme α_k est un minimum de $\alpha \mapsto q(\alpha)$, on a $q'(\alpha_k) = 0$, soit

$$\langle \nabla J(x_k), \nabla J(\underbrace{x_k - \alpha_k \nabla J(x_k)}_{x_{k+1}})\rangle = 0$$

(gradient à pas optimal)



3) On applique la γ-convexité

à $\begin{cases} u = x_{k+1} \\ v = x_k \end{cases}$

En effet, dans ce cas

$$\langle \nabla J(u), v - u\rangle = \langle \nabla J(x_{k+1}), \alpha_k \nabla J(x_k)\rangle = 0$$

On a

$$\frac{\gamma}{2}\|x_{k+1} - x_k\|^2 \leq J(x_k) - J(x_{k+1})$$

Par télescopage,

$$\underbrace{\frac{\gamma}{2}\sum_{k=0}^{N}\|x_{k+1} - x_k\|^2}_{STP} \leq \underbrace{J(x_0) - J(x_N)}_{\text{majoré}}$$

La série est donc convergente. Son terme général tend donc vers 0.

4) (x_k) bornée ?

(si (u_n) réelle tq $u_{n+1} - u_n \to 0$)
A-t-on (u_n) bornée ?

11 personnes : 7 oui / 3 non

oui ? $\|x_n\| \le \underbrace{\sum_{k=1}^{n} \|x_k - x_{k-1}\|}_{\downarrow \atop 0} + \|x_0\|$??

NON ? $u_n = \sqrt{n}$: $u_{n+1} - u_n = \dfrac{1}{\sqrt{n} + \sqrt{n+1}}$

Par l'absurde, il existe une sous suite tq $\|x_{\alpha(k)}\| \to +\infty$

Or, $J(x_{\alpha(k)}) \le J(x_0)$ (descente)

Comme $J(x_{\alpha(k)}) \to +\infty$ par γ convexité, on aboutit à une absurdité.

5) Comme $\langle \nabla J(x_k), \nabla J(x_{k+1}) \rangle$, on a

$$\|\nabla J(x_k) - \nabla J(x_{k+1})\|^2 = \|\nabla J(x_k)\|^2 + \|\nabla J(x_{k+1})\|^2$$

d'où le résultat //

4

Ainsi

$$\|\nabla J(x_k)\|^2 \leq L \|x_{k+1} - x_k\|^2$$

(∇J - L- Lipschitz sur tout borné)

Or $x_{k+1} - x_k \to 0$ (Q.3) d'où

$$\nabla J(x_k) \to 0$$

6) La 2ème relation de γ-convexité donne avec $v = x_k$ et $u = x^*$:

$$\gamma \|x_k - x^*\|^2 \leq \langle \nabla J(x_k) - \underbrace{\nabla J(x^*)}_{= 0}, x_k - x^* \rangle$$

soit le résultat //

Par Cauchy Schwarz,

$$\gamma \|x_k - x^*\|^2 \leq \|\nabla J(x_k)\| . \|x_k - x^*\|$$

soit

$$\|x_k - x^*\| \leq \frac{1}{\gamma} \underbrace{\|\nabla J(x_k)\|}_{\downarrow (Q5)\ 0}$$

et $x_k \to x^*$

Ex 5)

$$X_1 = \begin{pmatrix} -5/6 \\ 5/8 \\ 3/4 \end{pmatrix} = X_0 - \underbrace{Hf(X_0)^{-1}}_{\begin{pmatrix} 12 & 0 & 0 \\ 0 & 24 & 0 \\ 0 & 0 & 12 \end{pmatrix}^{-1}} . \underbrace{\nabla f(X_0)}_{\begin{pmatrix} 2 \\ 9 \\ 3 \end{pmatrix}}$$

Ex 6)

(il s'agit de BFGS)