Victor Niaussat

TD Algorithme SDI

1) $\forall x, y \in \mathbb{R}^N$ $J(y) \geq J(x) + \nabla J(x)(y-x) + \frac{\gamma}{2}|x-y|^2$ (1)
$J(x) \geq J(y) + \nabla J(y)(x-y) + \frac{\gamma}{2}|y-x|^2$ (2)

$\Rightarrow$ (1) − (2) $\boxed{(\nabla J(x) - \nabla J(y)).(x-y) \geq \gamma |x-y|^2}$

J est γ - convexe et on a la définition :

$\forall t \in [0;1];$ $\boxed{J(ty + (1-t)x) \leq (1-t)J(x) + tJ(y) - \frac{\gamma}{2}t(1-t)|y-x|^2}$

2) $x \neq y$ et $t \in]0;1[$ $\Rightarrow$ $-\frac{\gamma}{2}t(1-t)|y-x|^2 < 0$

Comme J γ - convexe ; $J(x_t) < (1-t)J(x) + tJ(y)$ et J est strictement convexe.

De manière équivalente, comme J est C^1 ;

$J(y) \geq J(x) + \nabla J(x).(y-x) + \frac{\gamma}{2}|x-y|^2$

Par Cauchy - Schwartz ; $\nabla J(x)(x-y) \leq \|\nabla J(x)\| |y-x|$
$\Leftrightarrow$ $-\|\nabla J(x)\|.|x-y| \leq \nabla J(x)(y-x)$.

et $J(y) \geq J(x) - \|\nabla J(x)\|.|x-y| + \frac{\gamma}{2}|x-y|^2$
$\geq J(x) + \underbrace{\left[\frac{\gamma}{2}|y-x| - \|J(x)\|\right]|y-x|}_{C_y}$.

À x fixé ; $\exists y_0 \in \mathbb{R}$; $\forall y \geq y_0$; $C_y > 0$

et $J(y) \geq J(x) + C_y |y - x|$

en faisant tendre $|y| \to +\infty$; $|y-x| \to +\infty$
et par inégalité : $J(y) \to +\infty$ donc J est coercive et il existe une et unique solution à (P). (Thm)

3) $\lambda \in \mathbb{R}^m$. Comme J est continue et f_i aussi, il est claire que par somme

$$\mathcal{L} : x \longmapsto J(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \text{ est continue}$$

$\mathcal{L}$ est convexe en x comme somme de fonctions convexes. J est strictement convexe donc $\mathcal{L}$ aussi. (il suffit d'avoir un terme dans la somme pour avoir la strict convexité).

$\forall x, y \in \mathbb{R}^N$ $\mathcal{L}(y;\lambda) = J(y) + \lambda f(y)$

$\geq J(x) + \lambda f(x) + \nabla[J + \lambda f](x)(y-x) + \frac{\gamma}{2}|x-y|^2$

$\geq \mathcal{L}(x;\lambda) - \|\nabla \mathcal{L}(x;\lambda)\| \, |y-x| + \frac{\gamma}{2}|y-x|^2$

En posant $\tilde{C}_y = \frac{\gamma}{2}|y-x| - \|\nabla \mathcal{L}(x;\lambda)\|$ et pour x fixé
$\exists y_0 \in \mathbb{R}$; $\forall y \geq y_0$; $\tilde{C}_y > 0$

$\mathcal{L}(y;\lambda) \geq \mathcal{L}(x;\lambda) + \tilde{C}_y |y - x|$

et $\mathcal{L}(y;\lambda) \xrightarrow[|y| \to +\infty]{} +\infty$.

Ainsi, d'après le théorème du cours ;

$\exists ! \, x^* \in C$ solution de $\mathcal{L}(x^*;\lambda) \leq \mathcal{L}(y;\lambda) \; \forall y \in \mathbb{R}^N$

La solution rend le gradient nul : $\nabla \mathcal{L}(x_\lambda) = 0$

4) $C' = \{ x \in \mathbb{R}^N ; \; f(x_i) < 0 \;\; \forall i \in [\![1;m]\!] \} \neq \emptyset$
qui est la condition de Slater.

De plus, $\nabla \mathcal{L}(x_\lambda) = \nabla J(x_\lambda) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x_\lambda) = 0$
est l'unique solution du problème.

D'après le théorème de Karach Kuhn et Tucker ;

$\exists \lambda^* \in D_+$ tq $(x_{\lambda^*} = x^*; \lambda^*)$ est un point selle.

et $\nabla J(x^*) + \sum \lambda_i^* \nabla f_i(x^*) = 0$ et $\lambda^* . f(x^*) = 0$.
$\nabla J(x^*) = 0$ et on retrouve le x^* de la question 2.

5) Le théorème nous donne le fait que $(x^*; \lambda^*)$ est point selle avec les conditions de Karach-Kahn-Tucker et Slater.

et $\mathcal{L}(x^*; \lambda) \leq \mathcal{L}(x^*; \lambda^*) \leq \mathcal{L}(x; \lambda^*)$

$H(\lambda^k) = \pi_8$

Question 6 : Prenons $\quad ; \lambda_i = \bar{\lambda}_i \quad i \neq i_0$

$$\mathcal{L}(\bar{x};\lambda) \leq \mathcal{L}(\bar{x};\bar{\lambda}) \quad \forall (x;\lambda) \in \mathbb{R}^N \times \mathcal{D}_+ .$$

$$\Leftrightarrow (\lambda - \bar{\lambda}) f(\bar{x}) \leq 0$$

- $\bar{\lambda} = 0 \Rightarrow \underbrace{\lambda}_{>0} f(\bar{x}) \leq 0 \rightsquigarrow f(\bar{x}) \leq 0$
- $\bar{\lambda} \neq 0$; prenons $\lambda_{i_0} = \bar{\lambda}_{i_0} \pm \frac{\bar{\lambda}_{i_0}}{2} > 0 \quad i_0$

$$\pm \frac{\bar{\lambda}_{i_0}}{2} f_i(\bar{x}) \leq 0 \Leftrightarrow \pm f_i(\bar{x}) \leq 0$$
$$\Leftrightarrow f_i(\bar{x}) = 0 \leq 0$$

et $\underbrace{\bar{\lambda} . f(\bar{x}) = 0}$; $f(\bar{x}) \leq 0$ donc $\bar{x} \in C$..

Comme $(\bar{x};\bar{\lambda})$ est un point selle, on a cela d'après le théorème +

$$\nabla \mathcal{L}(\bar{x};\bar{\lambda}) = \nabla J(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla f_i(\bar{x}) . = 0 .$$

On connait la solution à (P_λ) qui est $(x_\lambda ; \lambda)$
Comme $(\bar{x};\bar{\lambda})$ est point selle et que $\bar{x} \in C$;
$\mathcal{L}(\bar{x};\bar{\lambda}) \leq \mathcal{L}(x;\bar{\lambda}) \quad \forall x \in \mathbb{R}^m$; $\quad (P_{\bar{\lambda}})$

$$\boxed{\bar{x} = x_{\bar{\lambda}}} .$$

- Comme on a unicité de ce point ;

$$\boxed{x^* = \bar{x}} .$$

$$\Pi_{\mathcal{D}_+}(\lambda) = \underset{\tilde{\lambda}}{\mathrm{argmin}} \|\lambda - \tilde{\lambda}\|^2 .$$

On sait que $\Pi_{\mathcal{D}_+}(\lambda)$ est l'unique élément de $\mathcal{D}_+$ tq

$$(\tilde{\lambda} - \Pi(\lambda)) \cdot (\lambda - \Pi(\lambda)) \leq 0$$

$\forall \tilde{\lambda} \in \mathcal{D}_+$. Supposons le résultat.

$$\text{et } [\tilde{\lambda} - \Pi(\bar{\lambda} + \tau f(\bar{x}))] \cdot [\bar{\lambda} + \tau f(\bar{x}) - \bar{\lambda}]$$

$$= \tau [\tilde{\lambda} - \bar{\lambda}] \cdot f(\bar{x}).$$

$$= \tau \tilde{\lambda} f(\bar{x}) - \tau \underbrace{\bar{\lambda} . f(\bar{x})}_{=0}$$

$$= \tau \tilde{\lambda} \underbrace{f(\bar{x})}_{\leq 0} \leq 0$$

Et $\bar{\lambda} = \Pi(\bar{\lambda} + \tau f(\bar{x}))$ par unicité de la projection orthogonale sur $\mathcal{D}_+$

Question 7 :

Soit $\lambda \in \mathbb{R}^m$; $H(\lambda) = \inf \{ \mathcal{L}(x ; \lambda) : x \in \mathbb{R}^n \}$.

$$\mathcal{L}(x_\lambda ; \lambda) = H(\lambda) \leq \mathcal{L}(x ; \lambda) \quad \forall x \in \mathbb{R}.$$

- D'après la question deux, l'inf est atteint et

$\mathcal{L}(x^* ; \lambda) = \mathcal{L}(x_\lambda ; \lambda)$ x^* est un point selle d'après la question 5. H concave et $\mathcal{D}_+$ fermée.
Et P' admet bien au moins une solution

Quand on trouve une solution λ à (P'), d'après la question 6, $\boxed{x^* = x_\lambda}$.

Question 8

$\forall \lambda \in \mathbb{R}^m$; $H(\lambda) = \mathcal{L}(x_\lambda ; \lambda)$ est différentiable en tant que composé de fonctions différentiable.

$\forall i \in [1 ; m]$ $\quad \dfrac{\partial H}{\partial \lambda_i}(\lambda) = \dfrac{\partial x_\lambda}{\partial \lambda_i}\left[\nabla_x J(x_\lambda) + \lambda_i \nabla_x f_i(x_\lambda)\right]$

$$\Rightarrow \boxed{\nabla_\lambda H(\lambda) = \nabla_\lambda x_\lambda \cdot \left[\nabla_x J(x_\lambda) + \lambda \cdot \nabla_x f(x_\lambda)\right] + f(x_\lambda)}.$$

Dans l'algorithme du gradient projeté, en cherchant à maximiser $H(\lambda)$ (minimiser $-H$)

$\lambda^0 \in C$ puis pour $K = 0 ; \dots ;$ jusqu'à converger

$$\lambda^{k+1} \leftarrow \Pi\left(\lambda^k + \alpha_k \nabla H(\lambda^k)\right).$$

En ayant chercher en amont x^k tq (existe avec les questions précédentes.

$\nabla J(x^k) + \lambda . \nabla f(x^k) = 0$ pour avoir

$H(\lambda^k) = \inf \{ \mathcal{L}(x^k ; \lambda^k) ; x^k \in \mathbb{R}^N)$

$$\lambda^{k+1} \leftarrow \Pi\Big(\lambda^k + \alpha_k\Big(\nabla_\lambda x_\lambda \cdot \underbrace{\left[\nabla_x J(x^k) + \lambda . \nabla_x f(x_k)\right]}_{=0} + f(x_k)$$

$$\Rightarrow \boxed{\lambda^{k+1} \leftarrow \Pi\left(\lambda^k + \tau f(x^k)\right)}$$

avec $\forall k$; $\alpha_k = \tau$.

② Points intérieurs.

9) Sur le bord ∂C ; f_i s'annule et pour $x \in \mathbb{R}^n$

$$f_i(x) \xrightarrow[x \to y]{} 0^- \Rightarrow -f_i(x) \xrightarrow[x \to y]{} 0^+$$

$$\Rightarrow \ln(-f(x)) \xrightarrow[x \to y]{} -\infty \Rightarrow J_t(x) \xrightarrow[x \to y]{} +\infty.$$

et sur ∂C ; on ne trouve pas de minimiseur : $f_i(x) < 0$.

- J_t est coercive et J_t strictement convexe car

$\Gamma_t : x \longmapsto -\sum_{i=1}^{m} t \ln(-f_i(x))$ est convexe (concavité du log). Pour que Γ_t soit $\mathcal{C}^1$; $f_i(x^t) < 0$

car $\nabla_x \Gamma_t = \dfrac{t \nabla f_i(x)}{f_i(x)}$ et J_t bien définie.

$\forall x ; y \in C$

$$J_t(y) = J(y) + \Gamma(y)$$

$$\geq J(y) + \Gamma_t(x) + \nabla \Gamma_t(x)(y-x)$$

convexité de Γ sur C

$$\geq J(x) + \Gamma_t(x) + \nabla \Gamma_t(x)(y-x) + \nabla J(x)(y-x)$$

γ convexité de J sur C

$$- \frac{\gamma}{2} \|x-y\|^2$$

$$= J_t(x) + \nabla J_t(x)(y-x) - \frac{\gamma}{2} |x-y|^2.$$

et J_t γ convexe donc coercive.

Donc J_t admet un point minimum sur C noté x^t vérifiant $\nabla J_t(x^t) = 0$ ($J_t \in \mathcal{C}^1$).

$$\nabla J_t(x^t) = 0 \iff \nabla J(x^t) - \sum_{i=1}^{m} t \frac{\nabla f_i(x^t)}{f_i(x^t)} = 0$$

en posant $\lambda_i^t = -\dfrac{t}{f_i(x_t)}$

On a $\boxed{\nabla J(x^t) + \sum_{i=1}^{m} \lambda_i^t \nabla f_i(x^t) = 0}$ (E)

Question 10

De (E) $\xrightarrow[t \to 0]{}$ $\nabla J(x') \sum_{i=1}^{m} \lambda_i' \nabla f_i(x') = 0$

par continuité de ∇J et ∇f_i ($\mathcal{C}^1$).

et $\lambda' f(x') = 0$ $\left(\lambda^t f(x^t) \xrightarrow[t \to 0]{} \lambda' f(x') \text{ et } -t \xrightarrow[t \to 0]{} 0 \right.$

D'après le théorème de Karush - Kuhn - Tucker, $(x' ; \lambda')$ est un point selle.

et sur le point selle, on minimise $\mathcal{L}$ en x et maximise H en λ.

$H(\lambda) = \mathcal{L}(x' ; \lambda)$ et $\sup H(\lambda) = H(\lambda')$ et on répond au problème (P')

$t \to 0$; $J_0(x) = J(x)$ et on minimise bien J sur C avec contraintes.

3. Parties numériques.

Question 11 :

$J \in \mathcal{C}^1$; f_1 et $f_2 \in \mathcal{C}^1$ en tant que polynome de $\mathbb{R}^2$

f_1 et f_2 sont même $\mathcal{C}^2$ et les Hessiennes :

$H(f_1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$; $H(f_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ sont positives donc f_1 ; f_2 sont convexe et C convexe.

$$\boxed{\nabla J(x) = \begin{bmatrix} 6x_1 - 2x_2 - 3 \\ 6x_2 - 2x_2 - 7 \end{bmatrix}}.$$

$\forall x;y \in \mathbb{R}^2$; $\langle \nabla J(x) - \nabla J(y) ; x - y \rangle$

$= [(x_1 - y_1)[6(x_1 - y_1) - 2(x_1 - y_1)] + [(x_2 - y_2)(6x_2 - 2x_2)]$

$= 4[[x_1 - y_1]^2 + (x_2 + y_1)^2] = \boxed{4\|x-y\|^2}$

et J et 4 - convexe.

C est non vide car $(0;0)$ est dedans et est aussi dans C'.

Question 12 :

Minimiser J sur C revient à minimiser $\mathcal{L}$ sur $\mathbb{R}^m$.

$(x_1;x_2) \in \mathbb{R}^2$; $\nabla J(x_1;x_2) + \lambda . \nabla f_1(x_1;x_2) = 0$
solution.

$$\Leftrightarrow \begin{bmatrix} 6x_1 - 2x_2 - 3 \\ 6x_2 - 2x_1 - 7 \end{bmatrix} + \begin{bmatrix} 2\lambda_1 x_1 + 2\lambda_2 \\ 2\lambda_1 x_2 \end{bmatrix} = 0 \quad (P)$$

$\nabla f(x_1;x_2) = 0 \Leftrightarrow (x_1;x_2) = (0;0)$
n'appartient pas au borde de $\mathcal{C}$.

$2x_1 - 1 \leq 0 \quad \Leftrightarrow \quad x_1 \leq \frac{1}{2}$
et $x_1^2 + x_2^2 - 1 \leq 0 \quad \Leftrightarrow \quad \frac{1}{4} + x_2^2 \leq x_1^2 + x_2^2 \leq 1$

$$\Leftrightarrow \begin{cases} x_1 \leq \frac{1}{2} \\ x_2^2 \leq \frac{3}{4} \end{cases} \Leftrightarrow \begin{cases} x_1 \leq \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \leq x_2 \leq \frac{\sqrt{3}}{2} \end{cases}.$$

TTK $\begin{cases} \lambda_1 f_1(x) = 0 \\ \lambda_2 f_2(x) = 0 \end{cases} \Leftrightarrow \begin{cases} \lambda_1 (x_1^2 + x_2^2 - 1) = 0 \\ \lambda_2 (2x_1 - 1) = 0 \end{cases}$

- $\lambda_1 = 0$; alors $4(x_1 + x_2) - 10 + 2\lambda_2 = 0$
  - $\lambda_2 = 0$; $x_1 + x_2 = \frac{5}{2}$

et $6x_1 - 2\left(\frac{5}{2} - x_1\right) = 0 \Leftrightarrow x_1 = \frac{5}{8} > \frac{1}{2}$ donc on rejète.

• Si $\lambda_2 \neq 0$; alors $x_1 = \frac{1}{2}$

$$\begin{cases} -2x_2 + \lambda_1 + 2\lambda_2 = 0 \\ 6x_2 - 8 + 2\lambda_1 x_2 = 0 \end{cases}$$

• si $\lambda_1 = 0 \Rightarrow x_2 = \frac{8}{6} = \frac{4}{3} > \frac{\sqrt{3}}{2}$ on rejète.

• si $\lambda_1 \neq 0$ $x_1^2 + x_2^2 = 1 \Rightarrow x_2 = \pm\frac{\sqrt{3}}{2}$

• $x_2 = \frac{\sqrt{3}}{2}$ alors $-2x_2 + \lambda_1 + 2\lambda_2 = 0$

$6x_2 + 2\lambda_1 x_2 = 8$

$$\Rightarrow \lambda_1 = \frac{8}{2x_2} - 3 = \frac{8}{\sqrt{3}} - 3$$

$$-2\frac{\sqrt{3}}{2} + \frac{8}{\sqrt{3}} - 3 + 2\lambda_2 = 0$$

$$\Rightarrow \lambda_2 = \frac{3+\sqrt{3}}{2} - \frac{4}{\sqrt{3}}$$

$$\boxed{x_1 = \frac{1}{2} \quad x_2 = \frac{\sqrt{3}}{2}}$$

Question 13 :

$$\nabla J(x^k) + \lambda_1^k \nabla f_1(x^k) + \lambda_2^k \nabla f_2(x_k) = 0$$

$$\Leftrightarrow \begin{bmatrix} 6+2\lambda_1^k & -2 \\ -2 & 6+2\lambda_1^k \end{bmatrix} \begin{bmatrix} x_1^k \\ x_2^k \end{bmatrix} = \begin{bmatrix} 3-2\lambda_2 \\ 7 \end{bmatrix}$$

sous rèserve d'inversibilité

$$\Rightarrow \begin{bmatrix} x_1^k \\ x_2^k \end{bmatrix} = \begin{bmatrix} 6+2\lambda_1^k & -2 \\ -2 & 6+2\lambda_1^k \end{bmatrix}^{-1} \begin{bmatrix} 3-2\lambda_2 \\ 7 \end{bmatrix} = A^{-1} b$$

inversible car $\det(A) \geq 6^2 - 2^2 = 32$.