Examen du cours d'optimisation différentiable

Durée : 2 heures

Les documents ainsi que les calculatrices ne sont pas autorisés.

****

Exercice 0.1

On note par $\|\cdot\|_2$ la norme Euclidienne sur $\mathbb{R}^n$. Soit $a \in \mathbb{R}^n$, $\gamma > 0$ et la fonction

$$F : \begin{cases} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ x & \longrightarrow & \exp\left(\frac{-\|a-x\|_2^2}{\gamma}\right). \end{cases}$$

Calculer la Hessienne de F.

****

Correction de l'exercice 0.1 On pose $f : x \in \mathbb{R}^n \to -\|a-x\|_2^2/\gamma$. On a $F(x) = \exp(f(x))$ et donc $\nabla F(x) = \nabla f(x) F(x)$. On a pour tout $x \in \mathbb{R}^n$,

$$\nabla F(x) = -\frac{2F(x)}{\gamma}(x-a)$$

et $\nabla^2 F(x) = J(\nabla F)(x)$ alors

$$\nabla^2 F(x) = -\frac{2F(x)}{\gamma} I_n + \frac{4F(x)}{\gamma^2}(x-a)(x-a)^\top.$$

****

Exercice 0.2

Résoudre

$$\min\left((x-1)^2 + (y-1)^2 : x + 2y - 1 \leq 0 \textbf{ et } 2x + y - 1 \leq 0\right) \tag{1}$$

****

Correction de l'exercice 0.2 On pose pour tout $x, y \in \mathbb{R}$

$$f(x,y) = (x-1)^2 + (y-1)^2, h_1(x,y) = x + 2y - 1 \text{ et } h_2(x,y) = 2x + y - 1.$$

On a pour tout $x, y \in \mathbb{R}$

$$\nabla f(x,y) = \begin{pmatrix} 2(x-1) \\ 2(y-1) \end{pmatrix} \text{ et } \nabla^2 f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2I_2$$

donc f est fortement convexe et différentiable. Comme il y a deux contraintes d'inégalités qui sont affines, on conclut que (1) est un problème d'OCD et qu'il admet une unique solution. On détermine cette solution par un argument de dualité Lagrangienne.

La fonction de Lagrange associée à (1) est définie en tout point $x, y \in \mathbb{R}$ et $\mu_1, \mu_2 \geq 0$ par

$$L((x,y),(\mu_1,\mu_2)) = f(x,y) + \mu_1 h(x,y) + \mu_2 h_2(x,y).$$

Le problème dual associé est

$$\max_{\mu_1,\mu_2 \geq 0} \psi(\mu_1,\mu_2) \text{ où } \psi(\mu_1,\mu_2) = \inf_{(x,y)\in\mathbb{R}^2} L((x,y),(\mu_1,\mu_2)).$$

Déterminons explicitement le problème dual. Soit $\mu_1, \mu_2 \geq 0$, on note $\mu = (\mu_1, \mu_2)$. Comme $(x,y) \in \mathbb{R}^2 \to L((x,y),\mu)$ est convexe et deux fois différentiable, $L(\cdot,\mu)$ atteint son minimum sur $\mathbb{R}^2$ en un point (x_μ, y_μ) si et seulement si $\nabla L((x_\mu,y_\mu),\mu) = 0$ càd

$$\begin{cases} x_\mu = 1 - \mu_2 - \mu_1/2 \\ y_\mu = 1 - \mu_1 - \mu_2/2. \end{cases}$$

On a donc

$$\psi(\mu_1,\mu_2) = L((x_\mu,y_\mu),(\mu_1,\mu_2)) = 2\mu_1 + 2\mu_2 - 2\mu_1\mu_2 - \frac{5\mu_1^2}{4} - \frac{5\mu_2^2}{4}.$$

On sait que ψ est concave, on peut aussi vérifiée qu'elle est fortement concave puisque

$$\nabla^2\psi(\mu_1,\mu_2) = \frac{-1}{2}\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \preceq \frac{-1}{2} I_2$$

vu que $-9/2$ et $-1/2$ sont les valeurs propres de cette Hessienne. On sait alors que le problème dual admet une unique solution. On peut regarder le point critique de ψ pour voir s'il est dans la contrainte (cela nous évitera peut-être à analyser le problème sous contrainte donné par le problème dual). On a $\nabla\psi(\mu_1,\mu_2) = 0$ si et seulement si $\mu_1 = \mu_2 = 4/9$. On voit que cette solution est bien dans la contrainte du problème dual (càd $\mu_1 \geq 0$ et $\mu_2 \geq 2$) et c'est donc bien la solution au problème dual.

On pose $\mu^* = (4/9, 4/9) = (\mu_1^*, \mu_2^*)$ et

$$\begin{cases} x^* = x_{\mu^*} = 1 - (4/9) - (4/9)/2 = 1/3 \\ y^* = y_{\mu^*} = 1 - (4/9) - (4/9)/2 = 1/3 \end{cases}$$

On vérifie que (x^*, μ^*) vérifie les conditions KKT :

1) x^* est un minimiseur de $L(\cdot,\mu^*)$ (vu plus haut)
2) x^* est faisable pour le primal car $h_1(x^*,y^*) = 0$ et $h_2(x^*,y^*) = 0$ de même μ^* est faisable pour le dual
3) $\mu_1^* h_1(x^*,y^*) = 0$ et $\mu_2^* h_2(x^*,y^*) = 0$.

On en déduit alors que x^* est solution de (1) et on sait qu'elle est unique par stricte convexité. L'unique solution de (1) est donc $(1/3, 1/3)$.

****

Exercice 0.3
Résoudre le problème d'optimisation suivant :

$$\max\left(\mathrm{Ent}(p) + \mathrm{Ent}(q) : (p,q) \in U, p+q = s\right) \tag{2}$$

où $U =]0,1[^2$**,** $0 < s < 2$ **et** $\mathrm{Ent}(p)$ **est l'entropie d'une loi de Bernoulli de paramètre** p**, càd** $\mathrm{Ent}(p) = -p\log p - (1-p)\log(1-p)$**.**

Astuce : pour résoudre le problème dual $\max_\lambda \psi(\lambda)$ **on pourra poser** $u = \exp(-\lambda)/(1+\exp(-\lambda))$ **et résoudre un problème d'optimisation en** $u \in]0,1[$**.**

****

Correction de l'exercice 0.3 On 'transforme' le problème de maximisation en problème de minimisation. On note $f : (p,q) \in U \to -\mathrm{Ent}(p) - \mathrm{Ent}(q)$ la fonction objectif et $g : (p,q) \in U \to p+q-s$ la contrainte d'égalité. On cherche alors à résoudre $\min(f(p,q) : g(p,q) = 0, (p,q) \in U)$.

On voit que pour tout $(p,q) \in U$, on a

$$\nabla f(p,q) = \begin{pmatrix} \log(p/(1-p)) \\ \log(q/(1-q)) \end{pmatrix} \text{ et } \nabla^2 f(p,q) = \begin{pmatrix} \frac{1}{p(1-p)} & 0 \\ 0 & \frac{1}{q(1-q)} \end{pmatrix}.$$

En particulier, f est fortement convexe en particulier si (2) admet une solution, elle est unique. De plus, il n'y a qu'une seule contrainte qui est une contrainte d'égalité affine, c'est donc un problème d'optimisation convexe différentiable, on peut donc appliquer les résultats en dualité Lagrangienne pour ce problème.

On commence par calculer la fonction duale. On a pour tout $\lambda \in \mathbb{R}$,

$$\psi(\lambda) = \inf_{(p,q)\in U} \mathcal{L}((p,q),\lambda) \text{ où } \mathcal{L}((p,q),\lambda) = f(p,q) + \lambda g(p,q).$$

Pour tout $\lambda \in \mathbb{R}$, la fonction $\mathcal{L}(\cdot,\lambda)$ est convexe, il suffit d'identifier les points critiques de cette fonction pour la minimiser sur l'ouvert U. On a $\nabla_{(p,q)}\mathcal{L}((p,q),\lambda) = 0$ ssi $p = q = \exp(-\lambda)/(1+\exp(-\lambda))$. On conclut que la fonction duale en $\lambda \in \mathbb{R}$ vaut

$$\psi(\lambda) = \mathcal{L}\left(\left(\frac{\exp(-\lambda)}{1+\exp(-\lambda)}, \frac{\exp(-\lambda)}{1+\exp(-\lambda)}\right), \lambda\right) = \theta\left(\frac{\exp(-\lambda)}{1+\exp(-\lambda)}\right)$$

où $\theta : u \in]0,1[\to (2-s)\log(1-u) + s\log(u)$. Par ailleurs, θ est maximale en $u^* = s/2$ et donc ψ est maximale en $\lambda^* = \log((1-u^*)/u^*)$.

On cherche maintenant les points $(p,q) \in U$ tels que $((p,q),\lambda^*)$ vérifie les conditions KKT. On sait déjà que $\nabla_{(p,q)}\mathcal{L}((p,q),\lambda^*) = 0$ ssi $p = q = \exp(-\lambda^*)/(1+\exp(-\lambda^*)) = u^*$. Il reste donc à véfifier que la solution $(p^*,q^*) = (u^*,u^*)$ est bien faisable pour le problème primal : c'est bien le cas vue que $p^* + q^* = 2u^* = s$ et comme il n'y a pas de contrainte d'inégalité, il n'y a pas de 'complementary condition' à vérifier. On a donc bien que $((u^*,u^*),\lambda^*)$ vérifie KKT, on en déduit donc que $(s/2,s/2)$ est l'unique point solution de (2) (et qu'il y a aussi dualité forte).

Autre manière de résoudre ce problème : On peut résoudre ce problème directement en utilisant un argument de symétrie si on sait prouver l'existence d'une unique solution. Ce qui n'est pas une chose facile à faire vue que f n'est pas coercive sur son domaine (en fait on peut prolonger f par continuité sur $[0,1]^2$ en posant $f(p,q) = 0$ si $(p,q) \in \partial U$). On pourrait donc très bien n'avoir aucune solution au problème

(2) car U est ouvert. Cependant, supposons pour le moment qu'on sait démontrer l'existence d'une solution. Dans ce cas, on sait par stricte convexité qu'elle serait unique. Maintenant, par symétrie du problème, on voit que si (p, q) est solution alors (q, p) l'est aussi et donc par unicité, on aurait $p = q$ et comme $p + q = s$, on en déduirait que forcément $p = q = s/2$. Mais ce raisonement ne tient que si on a réussi à prouver l'existence d'une solution. Pour démontrer une telle existence (et sans passer par un argument de dualité, càd par une preuve directe), on 'doit' se ramener à une application du Théorème de Weierstrass. Pour cela, on peut prolonger f par continuité sur $[0, 1]^2$ en posant f égale à 0 sur le bord de son domaine de définition ; on continue à noter la fonction f prolongée par continuité. Par continuité de f et compacité de $[0, 1]^2$, il existe une solution au problème de minimisation de f sur $\{(p, q) \in [0, 1]^2 : p + q = s\}$. Il reste à démontrer la stricte convexité de f sur $[0, 1]^2$. On sait pour le moment que f est fortement convexe sur $]0, 1[^2$ et comme $p(1 - p) \le (1/4), \forall p \in [0, 1]$, on a que $\nabla f(p, q) \succeq (1/4)I_2$. On a donc pour tout $(p_0, q_0), (p_1, q_1) \in]0, 1[^2$ et $0 \le \alpha \le 1$ que

$$f(\alpha(p_0, q_0) + (1 - \alpha)(p_1, q_1)) \le \alpha f(p_0, q_0) + (1 - \alpha)f(p_1, q_1) - \frac{\alpha(1 - \alpha)}{8} \|(p_0, q_0) - (p_1, q_1)\|_2^2 .$$

En particulier, si on fait tendre (p_0, q_0) vers le bord de $[0, 1]^2$ ou si on fait tendre (p_0, q_0) et (p_1, q_1) vers le bord de $[0, 1]^2$ alors cette égalité reste vraie et donc le prolongement de f sur son bord reste fortement convexe et donc strictement convexe. On en déduit que f est fortement convexe et continue sur $[0, 1]^2$ donc elle admet un unique minimum sur $[0, 1]^2$ et en utilisant le raisonnement du dessus on voit que ce minimum est forcément $(s/2, s/2)$. Finalement, comme $(s/2, s/2) \in]0, 1[^2$, c'est aussi l'unique minimum de f sur $\{(p, q) \in]0, 1[^2 : p + q = s\}$.

Autre manière de résoudre ce problème : Une autre manière de résoudre le problème est de se ramener à un problème d'optimisation en une variable. Vu que $q = p - s$, on peut résoudre en p le problème

$$\max\left(\mathrm{Ent}(p) + \mathrm{Ent}(s - p) : 0 < p < 1, 0 < s - p < 1\right)$$

et en déduire une valeur optimale de q (vu que $q = p - s$) pour résoudre le problème (2).

****

Exercice 0.4
On considère le problème

$$\max\left((1 - A)\log(1 - p) + A\log(p) + (1 - B)\log(1 - q) + B\log(q) : 0 < q \le p < 1\right) \tag{3}$$

où $A, B \in]0, 1[$**.**

a) Appliquer le théorème de KKT pour résoudre (3). (On pourra distinguer les 3 cas $A < B$, $A > B$ et $A = B$ avant de de considérer les cas de la complementary condition).

b) Montrer que (3) est un problème d'optimisation convexe différentiable

c) Trouver un point-selle de la fonction de Lagrange quand $A > B$ et $A \le B$.

****

Correction de l'exercice 0.4 **a)** On pose $U =]0, 1[^2$. On voit que la fonction objective $f : (p, q) \in U \to (1 - A)\log(1 - p) + A\log(p) + (1 - B)\log(1 - q) + B\log(q)$ est fortement concave et que les ensembles

$\{(p,q) \in U : f(p,q) \geq \alpha\}$ sont soit vides soit des compacts de $\mathbb{R}^2$ car $-f$ est coercive sur U : $\lim(-f) = +\infty$ quand $p \to 0$ ou $p \to 1$ ou $q \to 0$ ou $q \to 1$. Alors comme l'ensemble $\{(p,q) \in \mathbb{R}^2 : p \geq q\}$ est un fermé de $\mathbb{R}^2$, on conclut que (3) admet une unique solution.

Il n'y a qu'une seule contrainte d'inégalité qui est affine. Donc la contrainte est qualifiée.

Soit (p,q) une solution au problème. D'après KKT, il existe $\mu \geq 0$ tel que $-\nabla f(p,q) + \mu \nabla h(p,q) = 0$ et $\mu h(p,q) = 0$ où $h : (p,q) \to q - p$. On a alors $\mu \geq 0$, $\mu(p-q) = 0$, $p \geq q$ et

$$\frac{p-A}{p(1-p)} = \mu = \frac{B-q}{q(1-q)}. \tag{4}$$

Si $A > B$: Si $p = q$ alors d'après (4), on a $p = q = (A+B)/2$ et $\mu = (p-A)/(p(1-p)) = [(A+B)/2 - A]/(p(1-p)) < 0$, ce qui est impossible donc $p \neq q$ et donc $\mu = 0$ par la complementayr condition et donc dans (4), on obtient que $p = A$ et $q = B$. Alors $(p = A, q = B)$ est l'unique solution au problème (3).

Si $A < B$: Si $p \neq q$ alors d'après la complementary condition $\mu = 0$ et donc $p = A$ et $q = B$. En particulier, $q > p$, ce qui n'est pas possible. On a donc $p = q$ et donc dans (4), $p = q = (A+B)/2$. Alors $(p = (A+B)/2, q = (A+B)/2)$ est l'unique solution au problème (3).

Si $A = B$: Si $p \neq q$ alors $\mu = 0$ et donc $p = A$ et $q = B$ càd $p = q$ ce qui est une contradiction donc $p = q$ et dans (4), on obtient que $p = q = (A+B)/2 = A = B$. Alors $(p = A, q = B)$ est l'unique solution au problème (3).

b)On calcul la Hessienne de f : pour tout $(p,q) \in U$,

$$\nabla f(p,q) = \begin{pmatrix} \frac{A-p}{p(1-p)} \\ \frac{B-q}{q(1-q)} \end{pmatrix} \text{ et } \nabla^2 f(p,q) = \begin{pmatrix} \frac{-p^2 - A(1-2p)}{p^2(1-p)^2} & 0 \\ 0 & \frac{-q^2 - B(1-2q)}{q^2(1-q)^2} \end{pmatrix} \prec 0.$$

Donc la fonction objective est concave et il n'y a qu'une seule contrainte qui est une contrainte d'égalité affine donc le problème (3) est un problème d'optimisation convexe (c'est le maximum d'une fonction concave sous contrainte affine ; ce qui est équivalenent au minimum d'une fonction convexe sous contrainte affine).

c)On peut voir que quand $A > B$ alors $((A,B),0)$ est un point-selle de la fonction de Lagrange en vérifiant les conditions KKT alors il y a dualité forte et (A,B) est l'unique solution du primal (càd de (3)) (et 0 est solution du problème duale). Quand $A \leq B$ alors $((A+B)/2, (A+B)/2), (B-A)/[(A+B)(1 - (A+B)/2)]$ est un point-selle de la fonction de Lagrange donc $p = q = (A+B)/2$ est l'unique solution du problème (3). On peut aussi trouver qu'il y a dualité forte vue qu'on a montré que la contrainte est qualifiée et qu'il existe une solution au problème. Donc la contrainte est qualifiée en une solution du problème et donc il y a dualité forte.