Correction partielle Feuille 8-9

Exercice 3. Voir quelles sont les contraintes actives en $(0,1)$ et $(1,0)$. Calculez les dérivées des contraintes en ce points. Les dérivées des contraintes peuvent êtres représentées par leurs gradients (vecteurs contenant les dérivées par rapport à chaque variable). Voir si les gradients forment une famille libre ou pas. (pour un de deux points ça va être vrai, pour l'autre non)

Le deuxième point concerne un problème d'optimisation sur l'ensemble K. On voit immédiatement qu'on a deux solutions $(1,0)$ et $(0,1)$. Le but est de voir concrètement que la conclusion du théorème des multiplicateurs de Lagrange est n'est pas vérifiée si les contraintes ne sont pas qualifiées. Il suffit de calculer le gradient du J et de voir si il est représentable comme une combinaison linéaire des gradients des contraintes actives en $(1,0)$ et $(0,1)$.

Exercice 4. 1. L'existence est immédiate car la fonctionnelle est continue sur un ensemble compacte. en plus la fonction $x \mapsto x \ln x$ est convexe, ce qui montre que $s(p)$ est concave strictement. En plus K est convexe non-vide. La fonctionnelle admet, donc un maximiseur unique p^*.

2. Si $p_i^* > 0$ pour tout i alors les contraintes d'inégalité ne sont pas actives. On a tout simplement deux contraintes d'égalité avec gradients $(1,1,1,...,1)$ et $(\varepsilon_1,...,\varepsilon_N)$ qui sont indépendants par hypothèse. En écrivant la relation des multiplicateurs de Lagrange pour p^* on obtient la conclusion. Notez que soit vous utilisez le résultat pour les conditions d'égalité sous l'hypothèse que votre minimiseur p^* est à l'intérieur du domaine $\mathbb{R}_+^N$, soit vous utilisez la condition d'optimalité pour les contraintes d'inégalité en écrivant chaque contrainte d'égalité affine comme deux contraintes d'inégalité.

Notez que la relation $\nabla s(p) + \lambda(1,1...) + \mu(\varepsilon_1,...,\varepsilon_N) = 0$ vous donne un système d'équations vous permettant de trouver chaque p_i^* en fonction de μ et λ.

3. On a $p_i^* = \frac{1}{Z}\exp(-\beta\varepsilon_i)$. Il faut montrer que β et Z sont définis de manière unique à partir de l'énergie totale E. Une manière de récupérer cette dépendance est d'utiliser les contraintes qui sont vérifiées a l'optimum.

Si $\sum_{i=1}^N p_i^* = 1$ on obtient $Z = \sum_{i=1}^N \exp(-\beta\varepsilon_i)$. Donc Z dépend de manière régulière de β. La deuxième contrainte s'écrit

$$\sum_{i=1}^N \frac{\exp(-\beta\varepsilon_i)}{\sum_{j=1}^N \exp(-\beta\varepsilon_j)} = E.$$

Pour finir, il suffit d'étudier la fonction

$$g(\beta) = \sum_{i=1}^N \frac{\exp(-\beta\varepsilon_i)}{\sum_{j=1}^N \exp(-\beta\varepsilon_j)}.$$

Ceci est une fonction régulière, décroissante, avec domaine $\mathbb{R}$ et image $]\varepsilon_1, \varepsilon_N[$. Pour voir ceci il suffit de calculer $g'(\beta)$, voir que $g'(\beta) \leq 0$ (par exemple en utilisant une inégalité type Cauchy-Schwarz) et de calculer les limites quand $\beta \to \infty$ et $\beta \to -\infty$.

Si g est régulière et strictement décroissante, alors $\beta = g^{-1}(E)$ dépend de manière régulière de E.

Le but de cette première partie de l'exercice est de montrer que si on se donne l'énergie totale E, on peut retrouver β, Z qui définissent p_i^* et qui sont un point selle du Lagrangien, i.e.

$$\nabla\mathcal{L}(p,\lambda,\mu) = \nabla(-s(p)) + \lambda(1,1...,1) + \mu(\varepsilon_1,...,\varepsilon_N) = 0.$$

4. On voudrait appliquer Kuhn-Tucker pour p^* qui est un point selle d'un Lagrangien correspondant au une fonctionnelle convexe $(-s)$ et à des contraintes convexes. Il y a plusieurs points à régler, d'abord :

- Kuhn-Tucker s'applique pour des contraintes d'inégalité. Ici il n'y a pas de souci, car les contraintes d'égalité affines peuvent êtres écrites comme deux contraintes d'inégalité convexes :

$$\sum_{i=1}^{N} p_i - 1 \leq 0$$
$$1 - \sum_{i=1}^{N} p_i \leq 0$$
$$\sum_{i=1}^{N} \varepsilon_i p_i - E \leq 0$$
$$E - \sum_{i=1}^{N} \varepsilon_i p_i \leq 0$$

- les contraintes d'inégalités sont affines, donc qualifiées
- à ce stade on ne peut pas appliquer Kuhn-Tucker, car s n'est pas défini dans un voisinage de la contrainte $p \geq 0$. Il faut trouver une manière d'étendre le Lagrangien au delà du $p \geq 0$.

Pour définir un Lagrangien régulier sur $\mathbb{R}^N \times \mathbb{R}^{4+N}$ on ne peut pas étendre $x \ln x$ qui vient en 0 avec une dérivée infinie. L'astuce est de s'arrêter avant 0. On sait que chaque p_i^* défini en 2 par Z et β trouvés en 3. est strictement positif. Prenons $\varepsilon > 0$ tel que $\varepsilon < p_i^*$ pour $i = 1, ..., N$. On considère une fonction f_ε telle que $f_\varepsilon(x) = x \ln x$ pour $x \geq \varepsilon$, f_ε est affine sur $]-\infty, \varepsilon]$ et f_ε est dérivable en ε.

Considérez maintenant $s_\varepsilon(p) = -k_B \sum_{i=1}^{N} f_\varepsilon(p_i)$. Alors le Lagrangien

$$\mathcal{L}_\varepsilon(p, \lambda, \mu) = -s_\varepsilon(p) + \mathbf{\Lambda} \cdot (\text{contraintes})$$

est bien défini sur tout l'espace défini pour une fonctionnelle $-s_\varepsilon$ convexe pour des contraintes d'inégalité affines, donc convexes. En plus, il est possible de montrer que p^* est un point selle du $\mathcal{L}_\varepsilon$. On sait qu'il existent $\lambda, \mu \in \mathbb{R}$ tels que

$$\nabla(-s_\varepsilon(p^*)) + \lambda(1, 1..., 1) + \mu(\varepsilon_1, .., \varepsilon_N) = 0.$$

Pourquoi ceci suffit pour avoir le point selle :

- les contraintes $p_i^* \geq 0$ ne sont pas actives, donc elles n'interviennent pas
- il faudrait pouvoir donner un signe au λ et μ (dans l'énoncé de Kuhn-Tucker ils sont positifs). Il est possible de donner un signe en remarquant qu'on peut choisir d'associer λ soit au vecteur $(1, 1, ..., 1)$ soit au vecteur $(-1, -1, ..., -1)$. Notez qu'on a deux contraintes d'inégalité affines opposées, avec des gradients opposés. On va choisir donc a mettre le λ comme multiplicateur pour la contrainte qui lui permet d'être positif. La même chose pour μ.

On peut donc appliquer Kuhn-Tucker et obtenir que p^* est un point selle pour $\mathcal{L}_\varepsilon$. Alors on en déduit que p^* est un minimum global de $-s_\varepsilon(p)$ sur K. Par construction on sait que $-s_\varepsilon(q) \leq -s(q)$ pour tout $q \in \mathbb{R}^N$. Donc si $q \in K$ on déduit que

$$-s(q) \geq -s_\varepsilon(q) \geq -s_\varepsilon(p^*) = -s(p^*).$$

Finalement on obtient que p^* est bien un maximiseur de s sur K.

Les points 5. et 6. sont plutôt calculatoires. Pour 5 notez que

$$S(E) = k_B \ln(Z + \beta E)$$

(en introduisant tout simplement l'expression de p_i^* dans la fonctionnelle).

En derivant $S(E)$ vous obtenez une expression contenant $Z'(E)$ et $\beta'(E)$. En utilisant que $Z = \sum_{i=1}^N \exp(-\beta\varepsilon_i)$ vous pouvez trouver une formule de $Z'(E)$ en fonction de β', E et $Z(E)$ qui vous permet de trouver le résultat.

6. Exercice élémentaire (mais calculatoire) avec des sup. Il faut voir comment splitter un supremum en deux en modifiant petit à petit ce qui est à l'intérieur.