Université de Picardie Jules Verne. Année 2024-2025.

M1 : Optimisation

Correction du TD 8

On considère une fonctionnelle α-elliptique J définie sur $I\!R^n$ et on pose :

$$U = \{v \in I\!R^n;\ \ Cv \leq d\}, \quad C \in M_{m,n}(I\!R),\ d \in I\!R^m.$$

On suppose que U est non vide.
On considère le problème (P) (problème primal) trouver $u \in U$ tel que

$$J(u) = \inf_{v \in U} J(v).$$

1. On pose $H_i(x) = (Cx - d)_i$ pour $i = 1, \cdots, m$. On cherche donc à minimiser J sous la contrainte

$$U := \{x \in I\!R^n |\, H_i(x) \leq 0,\ \ \forall\, i = 1, \cdots, m\}.$$

Les fonctions H_i sont affines donc continues et U est fermé comme intersection de fermés ($H_i^{-1}(]-\infty, 0])$ est fermé comme image réciproque d'un fermé par une application continue).
De plus, C est convexe puisque pour tout $u, v \in U$ et pour tout $t \in [0, 1]$, on a

$$C(tu + (1-t)v) = tCu + (1-t)Cv \leq td + (1-t)d = d.$$

La fonctionnelle est elliptique, donc continue, strictement convexe et coercice (résultat établi dans le TD 2). On en déduit que le problème primal (P) admet une unique solution.
Soit L le lagrangien associé au problème primal.

Le Lagrangien L associé au problème primal, défini sur $I\!R^n \times I\!R^m_+$, est donné par

$$L(v, \mu) = J(v) + \sum_{i=1}^{m} \mu_i H_i(v) = J(v) + ({}^tC\mu, v)_n - (\mu, d)_m.$$

D'après un théorème vu en cours, comme J et H_i $(i = 1, \cdots, m)$ sont convexes et dérivables (les fonctions H_i sont affines, donc convexes), les contraintes étant qualifiées (car *affines*), à une solution du problème (P), on peut associer un vecteur $\lambda \in I\!R^m_+$ tel que le couple (u, λ) soit point-selle du Lagrangien L.
Le second problème est encore un problème avec contraintes, mais l'avantage est que la contrainte peut être prise en compte plus facilement. Pour résoudre le second problème, on applique la méthode du gradient projeté à la fonction $\mu \mapsto L(u_k, \mu)$. Il est alors aisé de déterminer l'opérateur de projection sur $I\!R^m_+$ (voir TD 7).
2. Le couple (u, λ) est point-selle du lagrangien, donc

$$L(u, \lambda) = \inf_{v \in I\!R^n} L(v, \lambda) = \sup_{\mu \in I\!R^m_+} L(u, \mu).$$

De la première égalité ($L(u,\lambda) = \inf_{v\in I\!R^n} L(v,\lambda)$), on déduit que

$$\nabla_v L(u,\lambda) = 0$$

donc

$$\nabla J(u) + {}^t C\lambda = 0.$$

D'autre part, comme $L(u,\lambda) = \sup_{\mu\in I\!R^m_+} L(u,\mu)$, on a aussi

$$-L(u,\lambda) = \inf_{\mu\in I\!R^m_+} -L(u,\mu).$$

et la contrainte étant *convexe*, une condition nécessaire d'extrémun (inéquation d'Euler) est que

$$(\nabla_\mu(-L(v,\lambda)), \mu - \lambda)_m \geq 0 \quad \forall\, \mu \in I\!R^m_+.$$

Or $\nabla_\mu L(v,\lambda) = Cv - d$, donc la condition précédente s'exprime sous la forme

$$(\phi(u), \mu - \lambda)_m \leq 0 \quad \forall\, \mu \in I\!R^m_+,$$

où $\phi(u) := Cu - d$. L'inégalité précédente peut s'écrire pour tout nombre $\rho > 0$

$$(\lambda - (\lambda + \rho\phi(u)), \mu - \lambda)_m \geq 0, \quad \forall\, \mu \in I\!R^m_+,$$

ce qui montre que λ peut s'interpréter comme la projection sur $I\!R^m_+$ de l'élément $\lambda + \rho\phi(u)$. On en déduit que

$$\left\{\begin{array}{l} \nabla J(u) + {}^t C\lambda = 0, \\ \lambda = P_+(\lambda + \rho\phi(u)). \end{array}\right. \tag{1}$$

3. La méthode d'Uzawa consiste à résoudre successivement la suite de problèmes : étant donné $\lambda^k \in I\!R^m_+$ ($k \in I\!N$), trouver $u^k \in I\!R^n$ tel

$$L(u^k, \lambda^k) = \inf_{v\in I\!R^n} L(v, \lambda^k)$$

puis trouver $\lambda^{k+1} \in I\!R^m_+$ tel que

$$L(u^k, \lambda^{k+1}) = \sup_{\mu\in I\!R^m_+} L(u^k, \mu).$$

Le premier problème de minimisation admet bien une unique solution puisque la fonction ψ définie par $v \mapsto J(v) + ({}^tC\mu, v)_n - (\mu, d)_m$ est continue, strictement convexe comme somme d'une fonction strictement convexe et d'une fonction convexe et coercive. On obtient la coercivité de ψ à partir de l'inégalité

$$\psi(v) \geq J(0) + (\nabla J(0), v) + \frac{\alpha}{2}\|v\|^2 - \|{}^tC\mu\|\|v\| - (\mu, d)_m.$$

Le second problème d'optimisation admet également au moins une solution car d'une part, la fonction J est convexe, dérivable et d'autre part, les contraintes sont convexes, dérivables et qualifiées.

4. Par définition de la méthode d'Uzawa, on a également

$$\begin{cases} \nabla J(u^k) +^t C\lambda^k = 0, \\ \lambda^{k+1} = P_+(\lambda^k + \rho\phi(u^k)). \end{cases} \tag{2}$$

5. En soustrayant membre à membre les égalités $\lambda = P_+(\lambda + \rho\phi(u))$ et $\lambda^{k+1} = P_+(\lambda^k + \rho\phi(u^k))$ et comme l'opérateur de projection est 1-lipschitzien, on déduit que

$$\|\lambda^{k+1} - \lambda\|_m \leq \|\lambda^k - \lambda + \rho C(u_k - u)\|_m.$$

Élevant au carré les deux membres de l'inégalité précédente, et développant le membre de droite, on obtient

$$\|\lambda^{k+1} - \lambda\|_m^2 \leq \|\lambda^k - \lambda\|_m^2 + 2\rho(\lambda^k - \lambda, C(u_k - u)) + \rho^2\|C\|^2\|u_k - u\|_n^2.$$

Or, d'après (**??**)-(**??**) et comme J est α-elliptique, on a aussi

$$(\lambda_k - \lambda, C(u_k - u)) = (\nabla J(u) - \nabla J(u_k), u_k - u) \leq -\alpha\|u_k - u\|_n^2.$$

On en déduit que

$$\|\lambda^{k+1} - \lambda\|_m^2 \leq \|\lambda^k - \lambda\|^2 - \rho\{2\alpha - \rho\|C\|^2\}\|u_k - u\|_n^2.$$

6. Compte tenu de l'hypothèse $0 < \rho < \frac{2\alpha}{\|C\|^2}$, on déduit que

$$\|\lambda^{k+1} - \lambda\|_m^2 \leq \|\lambda^k - \lambda\|_m^2, \quad \in I\!\!N.$$

Ainsi, la suite $(\|\lambda^k - \lambda\|_m)$ est décroissante et minorée par 0, donc elle converge. Comme

$$\rho\{2\alpha - \rho\|C\|^2\}\|u_k - u\|_n^2 \leq \|\lambda^k - \lambda\|_m^2 - \|\lambda^{k+1} - \lambda\|_m^2,$$

on déduit que la suite (u^k) converge vers u.
7. La suite (λ^k) est bornée, donc par compacité, on peut en extraire une sous-suite qui converge vers un élément $\lambda' \in I\!\!R_+^m$. Par passage à la limite dans la première équation de (**??**), on obtient que cet élément λ' satisfait

$$\nabla J(u) +^t C\lambda' = 0.$$

Comme le rang de C est égal à m, l'image de C est égale à $I\!\!R^m$ ce qui équivaut à dire que le noyau de ${}^tC$ est réduit à $\{0\}$. En effet, soit u un élément du noyau de ${}^tC$. Compte tenu de la surjectivité de C, il existe $z \in I\!\!R^n$ tel que $Cz = u$ et ${}^tCCz = 0$. Mais alors

$$({}^tCCz, z) = (Cz, Cz) = 0,$$

donc $u = 0$.
Comme ${}^tC(\lambda - \lambda') = 0$, on a $\lambda = \lambda'$. La suite bornée (λ^k) admet une unique valeur d'adhérence, donc elle converge vers cette valeur d'adhérence (argument utilisé à plusieurs reprises cette année).