Ecole de recherche ENS Lyon – janvier 2011
TD optimisation et convexité
Claude Lemaréchal, Jérôme Malick

TD – dualité lagrangienne

Exercice 1 – Dualiser d'autres contraintes. Avec les notation du cours, on considère le problème

$$\begin{cases} \max & \varphi(u) \\ & u \in U \\ & c(u) \in B \end{cases}$$

où B est une partie de $\mathbb{R}^n$. On suppose qu'on dispose d'un oracle qui résout

$$\theta(\lambda) := \max_{u \in U} \varphi(u) - \lambda^\top c(u).$$

a) Introduisant une variable supplémentaire, écrire le dual du problème

b) Appliquer le résultat avec $B = \{0\}$, $B = \mathbb{R}^n_+$ et B la boule euclidienne de rayon ε.

Exercice 2 – Entropie et dualité. Nous souhaitons estimer un vecteur inconnu $\bar{x} \in \mathbb{R}^n$ à coefficients positifs, dont nous ne connaissons que certaines de ses « réalisations », c'est-à-dire nous ne connaissons que ${a_i}^\top \bar{x} = b_i$, pour des $a_i \in \mathbb{R}^n$ et $b_i \in \mathbb{R}$ connus ($i = 1, \dots, m$), avec $m < n$. On note A la matrice dont les lignes sont les ${a_i}^\top$, et $b \in \mathbb{R}^m$ le vecteur des b_i.

Parmi tous les vecteurs vérifiant ces conditions, on décide de préférer un vecteur « entropique » défini comme une solution du problème (où log est le logarithme népérien)

$$(\mathrm{P}) \begin{cases} \min & \sum_{k=1}^n x_k \log(x_k) \\ & Ax = b, \\ & x \geqslant 0. \end{cases}$$

On suppose qu'il existe un vecteur positif strictement réalisable (autrement dit, il existe un vecteur x tel que $x_k > 0,\ (k = 1, \dots, n)$ vérifiant $Ax = b$).

a) Considérons la fonction $\varphi_\alpha \colon \mathbb{R}_+ \to \mathbb{R}$ définie pour une constante $\alpha \in \mathbb{R}$ donnée par

$$\varphi_\alpha(t) = \begin{cases} t \log(t) + \alpha\, t & \text{si } t > 0 \\ 0 & \text{si } t = 0. \end{cases}$$

Montrer que φ_α est convexe et coercive sur $\mathbb{R}^*_+$.

b) Noter que φ_α est continue sur $\mathbb{R}_+$ tout entier, en déduire que φ_α est convexe sur $\mathbb{R}_+$.

c) Donner le t qui atteint le minimum de φ_α sur $\mathbb{R}_+$, ainsi que la valeur de ce minimum.

Revenons à présent au problème (P), que l'on va résoudre par dualité. On suggère de dualiser uniquement la contrainte couplante $Ax = b$.

d) Mettre le problème sous la forme du cours. Donner l'expression du lagrangien, et donner la définition de la fonction duale et du problème dual. On notera λ la variable duale, θ la fonction duale et (D) le problème dual.

e) Observer que la maximisation du lagrangien (à λ fixé) se découple en n problèmes. Donner l'unique $x_\lambda \in \mathbb{R}^n$ qui maximise le lagrangien (à λ fixé).

f) Montrer que θ est différentiable, et donner l'expression de $\theta(\lambda)$ et de $\nabla\theta(\lambda)$ en fonction de x_λ.

g) Supposons qu'il existe une solution duale qu'on note λ^*. Comment la calculer ?

h) Montrer que x_{λ^*} est réalisable dans le primal.

i) En déduire que x_{λ^*} est une solution primale.

j) Expliquer finalement comment répondre à la question initiale de cet exercice.

Exercice 3 – Affectation linéaire. Reprenons le problème d'affectation évoqué lors du premier cours :

$$(\mathrm{P})\left\{\begin{array}{l}\min \quad \sum_{i,j=1}^{n} a_{ij}u_{ij} \\ \quad \sum_{i=1}^{n} u_{ij}=1, \quad \text{pour tout } j=1,\ldots,n, \\ \quad \sum_{j=1}^{n} u_{ij}=1, \quad \text{pour tout } i=1,\ldots,n, \\ \quad u_{ij}\in\{0,1\}, \quad \text{for all } i,j=1,\ldots,n.\end{array}\right.$$

a) On veut dualiser les n premières contraintes dans (P) : écrire le problème (P) sous la forme du cours (c'est-à-dire exhiber, φ, c et $\mathcal{U}$).

b) Écrire par ailleurs (P) sous forme matricielle, et donner une interprétation « matricielle » du problème.

c) Écrire le lagrangien $L(\cdot;\cdot)$ associé à ce problème. Montrer qu'il existe une matrice U_λ qui maximise $L(\cdot;\lambda)$ à λ fixé. Définir θ la fonction duale associée.

d) Montrer qu'on peut calculer U_λ explicitement. Indice : remarquer que la maximisation de $L(\cdot,\lambda)$ se découple par rapport à chaque ligne i.

Exercice 4 – Méthode des régions de confiance. Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ de classe C^2 ; on veut utiliser une méthode par région de confiance pour résoudre

$$(\mathrm{P})\left\{\begin{array}{l}\min \quad f(x) \\ \quad x\in\mathbb{R}^n.\end{array}\right.$$

On suppose qu'à l'itéré courant x on dispose d'une estimation $M\in\mathcal{S}_n^+$ du hessien $\nabla^2 f(x)$; l'itéré suivant $x+d$ est construit avec la solution du problème suivant

$$(\mathrm{C})\left\{\begin{array}{l}\min \quad \frac{1}{2}\, d^\top M d+\nabla f(x)^\top d \\ \quad \frac{1}{2}\|d\|^2\leqslant\delta, \\ \quad d\in\mathbb{R}^n,\end{array}\right.$$

où $\delta>0$ est un paramètre choisi par ailleurs.

a) Montrer qu'il existe une solution à (C).

b) Écrire (C) avec un max (pour être mis sous la forme du cours), puis écrire le lagrangien associé à la dualisation de la contrainte $\frac{1}{2}\|d\|^2-\delta\leqslant 0$, en introduisant une variable duale réelle $\mu\geqslant 0$.

c) Fixons $\mu>0$; montrer que le lagrangien admet alors un unique maximum (selon d).

d) Dans ce cas, calculer le vecteur $d(\mu)$ qui réalise le maximum.

e) Montrer que si x n'est pas tel que $\nabla f(x)=0$, alors $d(\mu)$ est une direction de descente pour f.

f) Appelons (D) le problème dual associé au lagrangien précédent. Montrer que (D) admet une solution $\mu^\star$ et qu'il n'y a pas de saut dual.

On suppose pour les deux dernières questions que $\mu^\star>0$.

g) Montrer que la fonction duale est dérivable sur $\mathbb{R}_+^*$. Calculer sa valeur $\theta(\mu)$ et sa dérivée $\theta'(\mu)$.

h) Que donne un simulateur de la fonction duale ? Comment peut-on résoudre ce problème dual (D) ? Est-ce un problème "facile" ? Comment résoudre le problème (C) ? Pourquoi est-ce important de résoudre rapidement (C) ?