Ecole de recherche ENS Lyon – janvier 2011
Claude Lemaréchal, Jérôme Malick

Optimisation & Convexité - Examen

– Tous les documents sont autorisés.
– La qualité de la présentation et de l'argumentation sera un élément important de la notation.

On souhaite minimiser le temps de tranfert d'informations dans un réseau de télécom. Après modélisation (voir quelques détails à la fin de la feuille d'énoncé), on obtient le problème :

$$\begin{cases} \min \quad \displaystyle\sum_{j=1}^{n} f_{c_j}\Big(\sum_{k=1}^{K} x_j^k\Big) \\ \quad x^k \in P_k, \quad \text{pour } k = 1, \dots, K \end{cases}$$

dont les données sont les suivantes :

- Variables : $x^1, \dots, x^K$ sont K vecteurs de $\mathbb{R}^n$ (et on note donc $x_j^k \in \mathbb{R}$ la j-ème composante du vecteur $x^k \in \mathbb{R}^n$) ;
- Contraintes : chaque x^k vérifie un certain nombre de contraintes affines qui définissent un polytope (convexe) P_k de $\mathbb{R}^n$ qui s'avère être inclus dans $\mathbb{R}^n_+$ (tous les points ont toutes leurs composantes positives). On suppose aussi que les P_k sont compacts.
- Objectif : à chaque composante j est associée un coût défini (pour $c_j > 0$ fixé) par la fonction $f_{c_j} : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$

$$f_{c_j}(t) = \begin{cases} \dfrac{t}{c_j - t} & \text{si } 0 \leqslant t < c_j \\ +\infty & \text{sinon.} \end{cases}$$

D'après un exercice fait en TD, on a l'expression explicite de la fonction conjuguées de f_{c_j}

$$f_{c_j}^*(\lambda_j) = \begin{cases} (\sqrt{c_j \lambda_j} - 1)^2 & \text{si } \lambda_j \geqslant 1/c_j \\ 0 & \text{sinon.} \end{cases}$$

Ce problème se prête à une résolution par dualité, c'est ce que nous allons étudier dans cet exercice. Considérons la nouvelle variable $y = \sum_{k=1}^{K} x^k \in \mathbb{R}^n$ pour formuler le problème comme

$$\begin{cases} \min \quad \displaystyle\sum_{j=1}^{n} f_{c_j}(y_j) \\ \quad \sum_{k=1}^{K} x^k - y = 0 \\ \quad x^k \in P_k, \quad \text{pour } k = 1, \dots, K. \end{cases}$$

a) On dualise la contrainte couplant les x^k et y : écrire le problème sous la forme du cours, introduire la variable duale $\lambda \in \mathbb{R}^n$ et le lagrangien associé.

b) Montrer que la fonction duale s'écrit

$$\theta(\lambda) = \sum_{j=1}^{n} f_{c_j}^*(\lambda_j) + \sum_{k=1}^{K} \sigma_{P_k}(-\lambda)$$

où $f_{c_j}^*$ est la fonction conjuguée de f_{c_j} et σ_{P_k} est la fonction support de l'ensemble P_k. Écrire le problème dual.

c) En utilisant que les P_k sont inclus dans $\mathbb{R}^n_+$, montrer que : si une composante j de λ est telle que $\lambda_j \leqslant 1/c_j$ alors $\sigma_{P_k}(-\lambda) \geqslant \sigma_{P_k}(-\tilde{\lambda})$ et $\theta(\lambda) \geqslant \theta(\tilde{\lambda})$, où $\tilde{\lambda}$ est le vecteur de $\mathbb{R}^n$ tel que

$$\tilde{\lambda}_j = 1/c_j \qquad \text{et} \qquad \tilde{\lambda}_i = \lambda_i \text{ pour tout } i \neq j.$$

En déduire que le problème dual est inchangé (= même valeur optimale et même solution) si on ajoute la contrainte $\lambda \in \Lambda$ avec

$$\Lambda = \{\lambda \in \mathbb{R}^n : \ \lambda_j \geqslant 1/c_j \text{ pour tout } j = 1, \ldots, n\ \}.$$

d) Montrer que le problème dual peut ainsi se formuler

$$\text{(D)} \quad \min_{\lambda \in \Lambda} \quad \Phi(\lambda) + \Psi(\lambda)$$

avec deux fonctions Φ et Ψ convexes et Φ de classe C^∞ (sur l'intérieur de Λ).

e) Montrer que $\nabla^2 \Phi(\lambda)$ est définie-positive (si $\lambda_j > 1/c_j$ pour tout j).

f) Écrire précisement les sorties d'un simulateur de Φ. Même question avec Ψ (utiliser que les P_k sont compacts pour calculer un sous-gradient).

g) Expliquer brièvement ce qu'est la méthode des faisceaux classique pour résoudre le problème (D). (Se placer à l'itération T et noter $\lambda^1, \ldots, \lambda^t, \ldots, \lambda^T$ les itérés précédents.) Écrire explicitement le problème quadratique qui calcule λ^{T+1} (avec objectif et contraintes concrètes).

h) On suppose qu'il existe une solution à (D). Montrer qu'il n'y a pas de saut dual.

On dispose donc d'une approche par dualité pour résoudre le problème initial. On va maintenant améliorer cette approche en construisant une méthode (de type faisceaux) adaptée à la structure particulière du problème dual (donc plus efficace en pratique).

i) Écrire le modèle des plans sécants de σ_{P_k} à l'itération T. En déduire un modèle de Ψ.

j) Quel est le modèle quadratique de Φ le plus précis ?

k) Proposer un modèle pour la fonction duale $\Phi + \Psi$ (qui permet construire une méthode de faisceaux adaptée). Écrire le problème quadratique qui calcule l'itéré suivant λ^{T+1}.

Remarque sur l'exercice : Cet exercice porte sur un problème classique d'optimisation dans les réseaux de télécom, appelé "multicommodity flow". Voici une intuition d'où vient ce problème :

- Il y a K paquets à faire passer dans un réseau, représenté par un graphe à n arcs.
- Pour un arc j, on note sa capacité $c_j > 0$ et sa fonction de congestion f_{c_j} qui est convexe, croissante et explose en c_j (la fonction de congestion choisie dans cet exercice s'appelle la fonction de Kleinrock).
- Les contraintes P^k codent le flot transportant le paquet k d'un point à un autre dans le réseau. Chaque P_k est donc défini par $x^k \geqslant 0$ et des contraintes affines du type $Ax^k = a^k$, avec A la matrice d'incidence du graphe et a^k représentant le flot à transporter (a^k n'a que 2 coordonnées non nulles qui ont des valeurs opposées). Simuler Ψ consiste ainsi à résoudre K problèmes de plus court chemin.
- On veut minimiser le temps de trajet total des K paquets à travers ce réseau.