

Cours OPT 201

Optimisation Différentiable – Théorie et Algorithmes

Exercices de la séance 6
(dualité)

Dualisation de problèmes d'optimisation

1. Optimisation linéaire et dualité
2. Dualisation d'un problème quadratique
3. Dualisation d'un problème linéaire avec contrainte conique
4. Dualisation d'une contrainte scalaire dans un problème à deux variables

Utilisation de la dualité

5. Identité du minimax de von Neumann
6. Méthodes de décomposition par les prix et les ressources
7. Approximation de Tchebychev d'un système linéaire surdéterminé

Mise en œuvre des méthodes de dualité

8. Projection par relaxation lagrangienne

1 Optimisation linéaire et dualité

Considérez le problème d'optimisation linéaire suivant (sous sa forme standard) :

$$(P_L) \quad \begin{cases} \inf_{x \in \mathbb{R}^n} c^\top x \\ Ax = b \\ x \geq 0, \end{cases} \quad (6.1)$$

avec $c \in \mathbb{R}^n$, A une matrice $m \times n$ et $b \in \mathbb{R}^m$.

1. *Problème dual.*

- (a) Montrez que le problème dual de (6.1) pour le lagrangien classique est le problème d'optimisation linéaire suivant :

$$(D_L) \quad \begin{cases} \sup_{(y,s) \in \mathbb{R}^m \times \mathbb{R}^n} b^\top y \\ A^\top y + s = c \\ s \geq 0. \end{cases} \quad (6.2)$$

- (b) Montrez que si l'on dualise seulement les contraintes d'égalité dans (6.1), le problème dual résultant est encore (6.2).
 (c) Montrez que si $b \in \mathcal{R}(A)$ et que l'on dualise seulement la contrainte d'inégalité dans (6.1), le problème dual résultant est encore (6.2).

2. *Problème bidual.* Montrez que le dual de (6.2) est (6.1), si la dualisation se fait comme précédemment.
 3. *Dualité faible.* Montrez que si x_0 est admissible pour (6.1) et (y_0, s_0) est admissible pour (6.2), alors

$$c^\top x_0 - b^\top y_0 = x_0^\top s_0 \geq 0. \quad (6.3)$$

En déduire que

$$b^\top y_0 \leq \sup_{\substack{(y,s) \\ A^\top y + s = c \\ s \geq 0}} b^\top y = \text{val}(D) \leq \text{val}(P) = \inf_{\substack{x \\ Ax = b \\ x \geq 0}} c^\top x \leq c^\top x_0. \quad (6.4)$$

4. *Dualité forte*⁴. Montrez que les propriétés suivantes sont équivalentes :

- (i) (P_L) et (D_L) sont réalisables,
- (ii) (P_L) a une solution,
- (iii) (D_L) a une solution.

Montrez que lorsque ces propriétés ont lieu il n'y a pas de saut de dualité, c'est-à-dire $\text{val}(D_L) = \text{val}(P_L)$.

Remarque. L'implication (i) \Rightarrow (ii) est l'argument le plus souvent utilisé pour montrer qu'un problème d'optimisation linéaire a une solution.

⁴ Cet exercice requiert l'utilisation du théorème affirmant que

$$\text{sol}(P_L) \neq \emptyset \iff \text{val}(P_L) \in \mathbb{R}. \quad (6.5)$$

2 Dualisation d'un problème quadratique

On considère le problème quadratique :

$$(P) \quad \begin{cases} \inf \frac{1}{2}x^T Qx + c^T x \\ Ax = b \\ x \geq 0, \end{cases}$$

où Q est une matrice $n \times n$ symétrique définie positive, $c \in \mathbb{R}^n$, A est une matrice $m \times n$ et $b \in \mathbb{R}^m$. Écrire un problème dual.

3 Dualisation d'un problème linéaire avec contrainte conique

Soient \mathbb{E} et \mathbb{F} deux espaces euclidiens, dont on note $\langle \cdot, \cdot \rangle$ les produits scalaires. Soit K un cône non vide de \mathbb{E} (i.e., $tx \in K$ si $t > 0$ et $x \in K$), $c \in \mathbb{E}$, $b \in \mathbb{F}$ et A une application linéaire de \mathbb{E} dans \mathbb{F} . On considère le problème sur \mathbb{E} suivant :

$$\begin{cases} \inf \langle c, x \rangle \\ Ax = b \\ x \in K. \end{cases} \quad (6.6)$$

1. *Dualisation de la contrainte d'égalité.* En dualisant la contrainte d'égalité avec le lagrangien classique $\ell(x, y) = \langle c, x \rangle - \langle y, Ax - b \rangle$, montrez que l'on obtient comme problème dual

$$\begin{cases} \sup \langle b, y \rangle \\ A^*y + s = c \\ s \in K^+, \end{cases} \quad (6.7)$$

où $K^+ := \{s \in \mathbb{R}^n : \langle s, x \rangle \geq 0, \forall x \in K\}$ est le cône dual de K .

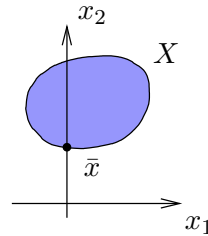
2. *Dualisation de la contrainte d'appartenance au cône.*

- (a) Montrez que, si K soit un cône convexe fermé non vide, on a $K^{++} = K$.
- (b) Supposons que K soit un cône convexe fermé non vide. Montrez que l'on retrouve le dual (6.7) si l'on dualise la contrainte linéaire *et* la contrainte d'appartenance au cône.

4 Dualisation d'une contrainte scalaire dans un problème à deux variables

Soit X une partie quelconque de \mathbb{R}^2 . On considère le problème :

$$(P) \quad \begin{cases} \inf x_2 \\ x \in X \\ x_1 = 0. \end{cases}$$



1. *Dualisation lagrangienne.* En dualisant la contrainte « $x_1 = 0$ » par le lagrangien classique, montrez que le problème dual (D) consiste à trouver la droite de \mathbb{R}^2 qui est en-dessous de X et qui rencontre l'axe des x_2 le plus haut possible.

2. *Existence de point-selle*⁵. Montrez que si X est un convexe contenant un point d'abscisse > 0 et un point d'abscisse < 0 et si le problème primal a une solution \bar{x} , alors le problème dual a une solution $\bar{\lambda}$ et $(\bar{x}, \bar{\lambda})$ est un point-selle du lagrangien.
3. *Construction de cas particuliers*. En choisissant des ensembles X particuliers, montrez que l'on peut rencontrer les situations suivantes.

- (a) Il y a un saut de dualité et des solutions du problème de Lagrange (où $\bar{\lambda}$ est une solution duale)

$$\inf_{x \in X} (\bar{\lambda}x_1 + x_2) \quad (6.8)$$

ne sont pas solutions de (P) .

- (b) Le lagrangien a un point-selle, mais certaines solutions du problème de Lagrange (6.8) ne sont pas solutions de (P) (il y a des *solutions importunes*).
- (c) Tout $\lambda \in \mathbb{R}$ est solution duale.
- (d) $\text{val}(P) = \text{val}(D) \in \mathbb{R}$, $\text{sol}(D) = \emptyset$.

4. *Dualisation lagrangienne augmentée*.

- (a) Dualisez la contrainte « $x_1 = 0$ » par le lagrangien *augmenté* et donnez une interprétation géométrique du problème dual.
- (b) Donnez un exemple d'ensemble X pour lequel il n'y a pas de saut de dualité avec cette dualisation par le lagrangien augmenté, alors qu'il y en a un avec la dualisation par le lagrangien classique.

5 Identité du minimax de von Neumann

Soit A une matrice réelle de type $m \times n$. On note

$$\Delta_p := \left\{ x \in \mathbb{R}^p : x \geq 0, \sum_{i=1}^p x_i = 1 \right\}$$

le simplexe unité de \mathbb{R}^p . On cherche à montrer que

$$\max_{y \in \Delta_m} \min_{x \in \Delta_n} y^T A x = \min_{x \in \Delta_n} \max_{y \in \Delta_m} y^T A x. \quad (6.9)$$

On note α la valeur à gauche dans (6.9), β la valeur à droite et e un vecteur dont toutes les composantes valent 1 (il peut varier en dimension, suivant le contexte).

1. Montrez que les problèmes

$$\sup_{y \in \Delta_m} y^T A x \quad \text{et} \quad \inf_{x \in \Delta_n} \sup_{y \in \Delta_m} y^T A x$$

ont une solution (le premier pour tout $x \in \mathbb{R}^n$ fixé).

On pourrait faire de même pour les problèmes dans le membre de gauche de (6.9). Cela justifie donc l'utilisation des opérateurs « min » et « max » dans (6.9) (au lieu de « inf » et « sup »).

⁵ La démonstration de ce point requiert la notion de normale à un convexe.

2. Montrez que

$$\min_{x \in \Delta_n} v^\top x = \min_{1 \leq i \leq n} v_i.$$

On admettra que l'on montrerait de la même manière que $\max\{v^\top x : x \in \Delta_n\} = \max\{v_i : 1 \leq i \leq n\}$.

En ajoutant si besoin à A la matrice tE ($E \in \mathbb{R}^{m \times n}$ est la matrice dont tous les éléments valent 1), on ajoute t aux deux membres de (6.9), si bien que l'on peut supposer que $\alpha > 0$ et $\beta > 0$ (on prend t assez grand), ce que l'on fait ci-dessous.

3. (a) Soit t un réel strictement positif. Montrez les équivalences suivantes :

$$\begin{aligned} t \leq \alpha &\iff \exists y \in \mathbb{R}^m : y \geq 0, e^\top y = 1, A^\top y \geq te \\ &\iff \inf_{\substack{y \geq 0 \\ A^\top y \geq e}} e^\top y \leq 1/t. \end{aligned}$$

(b) En déduire que

$$\inf_{\substack{y \geq 0 \\ A^\top y \geq e}} e^\top y = \frac{1}{\alpha}.$$

(c) Montrez que le problème dans le membre de gauche ci-dessus a une solution.

On admettra que l'on montrerait de la même manière que

$$\frac{1}{\beta} = \max_{\substack{x \geq 0 \\ Ax \leq e}} e^\top x.$$

4. Conclure.

Note. L'identité (6.9) est ce que l'on appelle l'*identité du minimax de von Neumann*. Elle fut obtenue en 1928 et marqua le point de départ de la dualité min-max et de l'étude mathématique de la théorie des jeux.

6 Méthodes de décomposition par les prix et les ressources

On considère le problème

$$\begin{cases} \inf f(x) \\ c(x) \leq 0 \\ x \in X, \end{cases} \quad (6.10)$$

où X est un ensemble, $f : X \rightarrow \mathbb{R}$ et $c : X \rightarrow \mathbb{R}^m$.

On suppose que le problème (6.10) a une **structure décomposable**, c'est-à-dire :

1) X s'écrit comme un produit cartésien de N ensembles quelconques X^j :

$$X = \prod_{j=1}^N X^j, \quad x = (x^1, \dots, x^N) \in X, \quad \text{avec } x^j \in X^j \text{ pour tout } j, \quad (6.11)$$

2) les fonctions f et c se décomposent en une somme de fonctions élémentaires f^j et c^j définies sur X^j à valeurs dans \mathbb{R} et \mathbb{R}^m respectivement :

$$f(x) = \sum_{j=1}^N f^j(x^j) \quad \text{et} \quad c(x) = \sum_{j=1}^N c^j(x^j). \quad (6.12)$$

Les contraintes de (6.10) sont dites **couplantes**, car ce ne sont pas des contraintes sur chaque $c^j(x^j)$, mais sur leur somme. Si l'on avait eu des contraintes du type $c^j(x^j) \leq 0$ pour tout $j = 1, \dots, N$, (6.10) aurait pu se décomposer en N problèmes indépendants :

$$\begin{cases} \inf f^j(x^j) \\ c^j(x^j) \leq 0 \\ x^j \in X^j, \end{cases}$$

pour $j = 1, \dots, N$. Le couplage des contraintes c^j dans (6.10) rend le problème plus difficile à résoudre.

Exemple. Une entreprise a sa production répartie sur N usines. L'usine $j \in \{1, \dots, N\}$ peut produire une quantité x^j en utilisant des ressources en quantité $c^j(x^j) \in \mathbb{R}^m$ (la composante i de c^j donne la quantité de i -ième ressource utilisée). Le coût de cette production propre à l'usine est $f^j(x^j)$ (il ne tient pas compte du coût des ressources). Il s'agit de minimiser le coût total $\sum_j f^j(x^j)$ de production tout en satisfaisant des contraintes globales $\sum_j c^j(x^j) \leq 0$ sur l'ensemble des ressources utilisées par toutes les usines.

Voici deux méthodes de résolution de (6.10)–(6.12) qui tiennent compte de la structure de ce problème et font appel à la dualité. Ces approches seront d'autant plus intéressantes (par rapport à l'approche standard n'utilisant pas la structure du problème) que N est grand et les X^j sont « petits » (par exemple des espaces vectoriels de faible dimension).

1. Décomposition par les prix (ou relaxation lagrangienne)

- (a) Écrire l'algorithme d'Uzawa pour le problème (6.10)–(6.12) et montrez son intérêt.
- (b) Dans le cas de l'exemple cité ci-dessus, donnez une interprétation du procédé de résolution et des multiplicateurs de Lagrange en termes de prix des ressources.

2. Décomposition par les ressources.

- (a) Montrez que le problème (6.10)–(6.12) peut s'écrire

$$\inf_{\substack{p^j \in \mathbb{R}^m \\ \sum_{j=1}^N p^j \geq 0}} \sum_{j=1}^N v^j(p^j), \quad (6.13)$$

où

$$v^j(p^j) := \inf_{\substack{x^j \in X^j \\ c^j(x^j) + p^j \leq 0}} f^j(x^j). \quad (6.14)$$

Donnez l'intérêt de cette formulation.

- (b) On suppose dorénavant que les fonctions f^j et c^j sont convexes et que les problèmes (6.14) en p^j ont une solution primale-duale (x^j, μ^j) , le multiplicateur μ^j étant associé à la contrainte « $c^j(x^j) + p^j \leq 0$ ». Montrez que
- i. les fonctions v^j sont convexes;
 - ii. pour tout $q^j \in \mathbb{R}^m$, on a $v^j(q^j) \geq v^j(p^j) + (\mu^j)^\top (q^j - p^j)$;
 - iii. (μ^1, \dots, μ^N) est un sous-gradient du critère de (6.13).
- (c) Dans le cas de l'exemple cité ci-dessus, donnez une interprétation du procédé de résolution de (6.10)–(6.12) par (6.13)–(6.14) et précisez l'utilité des multiplicateurs μ^j .

7 Approximation de Tchebychev d'un système linéaire surdéterminé

Soient A une matrice de type $m \times n$ et $b \in \mathbb{R}^m$. On considère le problème d'approximation au sens de Tchebychev ou en norme ℓ_∞ suivant :

$$v_{\text{tch}} := \min_{x \in \mathbb{R}^n} \|Ax - b\|_\infty. \quad (6.15)$$

Le problème (6.15) n'a pas nécessairement une unique solution, mais il en a au moins une, d'où l'utilisation du « min ».

1. Montrez que le problème (6.15) a une solution.

On ne connaît pas d'expression analytique de v_{tch} , ni a fortiori l'expression analytique des solutions de (6.15), alors que les solutions du problème de moindres-carrés sont connues. En particulier, on connaît la forme de ses solutions x_{mc} et son résidu optimal

$$r_{\text{mc}} := Ax_{\text{mc}} - b.$$

Il est donc naturel de chercher à savoir si ce résidu permet d'estimer v_{tch} et plus précisément d'en donner une borne inférieure (positive bien sûr).

2. Montrez que

$$\frac{1}{\sqrt{m}} \|r_{\text{mc}}\|_\infty \leq v_{\text{tch}}. \quad (6.16)$$

On montre maintenant que la dualité permet de resserrer cette estimation de v_{tch} .

3. Montrez dans quel sens on peut considérer que le problème

$$\begin{aligned} \max_{y \in \mathbb{R}^m} \quad & b^\top y \\ \text{s.t.} \quad & \|y\|_1 \leq 1 \\ & A^\top y = 0 \end{aligned} \quad (6.17)$$

est dual du problème (6.15).

Montrez que (6.17) a une solution (si bien que le « max » y est utilisé à bon escient).

4. On suppose ici que $r_{\text{mc}} \neq 0$ (dans le cas contraire, $v_{\text{tch}} = 0$ et il n'y a donc pas lieu de trouver un minorant strictement positif de v_{tch}). En choisissant bien un point admissible du problème dual (6.17), montrez que l'on a

$$\frac{\|r_{\text{mc}}\|_2^2}{\|r_{\text{mc}}\|_1} \leq v_{\text{tch}}.$$

Montrez que ce minorant est meilleur que celui donné en (6.16).

8 Projection par relaxation lagrangienne

On considère le problème dans \mathbb{R}^2

$$\begin{cases} \inf x_1^2 + x_2^2 \\ x_1 + x_2 = 1 \\ x_2 \geq 0. \end{cases} \quad (6.18)$$

1. Écrire un problème dual de (6.18).
2. Expliciter l'algorithme d'Uzawa sur cet exemple, en partant de $\lambda = 0$ et en prenant un pas dans un compact de $]0, 1[$.
3. Même question pour l'algorithme d'Arrow-Hurwicz.