Chapitre 4

RELAXATION LAGRANGIENNE

4.1 Introduction

La relaxation lagrangienne est une manipulation classique en optimisation sous contraintes. Bien qu'intimement liée à la convexité, elle est couramment utilisée dans le cas non convexe. En particulier, elle permet d'obtenir des bornes de la valeur optimale de certains problèmes d'optimisation combinatoire durs. L'idée consiste à relaxer une partie des contraintes (en principe, celles qui rendent le problème compliqué) qui sont introduites dans la fonction objectif sous la forme d'une pénalité qui combine linéairement les contraintes relaxées. Les coefficients de cette combinaison linéaire sont appelées les variables duales associées à la relaxation lagrangienne.

Considérons le problème d'optimisation dans $I\!R^n$ suivant :

$$(\mathbf{P}) \qquad \begin{array}{ll} \text{Minimiser} & f(x) \\ & g(x) \leq 0 \\ & x \in S \end{array}$$

où S est un sous-ensemble de $I\!R^n$, $f : S \mapsto I\!R$ et $g : S \mapsto I\!R^p$.

(P) sera appelé le problème *primal*. Ici, les contraintes $g(x) \leq 0$ sont considérées comme les contraintes complicantes (on dit aussi couplantes) dans le sens où on suppose que l'on dispose d'un algorithme 'efficace' pour minimiser une fonction de $I\!R^n$ sur l'ensemble S. Observez qu'aucune hypothèse particulière n'a été formulée sur le problème primal.

4.2 Lagrangien et fonction duale

On définit le *Lagrangien* du problème (P) comme la fonction $L : S \times (I\!R^p)^+ \mapsto I\!R$ suivante :

$$L(x, u) = f(x) + \langle u, g(x) \rangle$$

où $u \in I\!R^p, u \geq 0$ est appelé le vecteur des multiplicateurs (de Lagrange dans le cas de contraintes égalités ou de Kuhn-Tucker dans le cas de contraintes inégalités, cf. chap. I-4) ou vecteur des *variables duales*. On appelle alors *fonction duale* associée au problème (P) la fonction $h : (I\!R^p)^+ \mapsto I\!R$ telle que :

$$h(u) = \inf_{x \in S} \; L(x, u)$$

définie sur $U = \{u \in (I\!R^p)^+ \mid \inf_{x \in S} L(x, u) > -\infty\}$.

Théorème 4.1 *La fonction h est concave sur tout sous-ensemble convexe de U.*

Démonstration Pour chaque $x \in S$, $f(x) + \langle u, g(x)\rangle$ est une fonction affine en u. Or, l'enveloppe inférieure d'une collection de fonctions affines est concave sur tout sous-ensemble convexe fermé de son domaine. □

En plus de cette propriété remarquable, la fonction duale fournit une borne inférieure pour le problème (P) :

Théorème 4.2

$$h(u) \leq f(x),\ \forall u \in U \text{ et } \forall x \in S \text{ tel que } g(x) \leq 0$$

Démonstration Par définition, $h(u) \leq f(x) + \langle u, g(x)\rangle, \forall x \in S$ et $\forall u \in U$. Donc, $h(u) \leq f(x)$ si $\langle u, g(x)\rangle \leq 0$, c.a.d. si $g(x) \leq 0$. □

Il est naturel de rechercher la meilleure borne inférieure de (P), c.a.d. la plus grande afin d'approcher, voire d'atteindre, la valeur optimal de(P). En effet, on peut écrire :

$$\sup_{u \geq 0} h(u) \leq \inf\{f(x) \mid x \in S, g(x) \leq 0\}$$

Corollaire 4.1 *En notant* $\sup h$ *et Si* $\sup h = +\infty$*, le problème primal n'a pas de solution réalisable.*

Si $\inf f = -\infty$ *?, le problème dual n'a pas de solution réalisable.*

Le problème dual associé à (P) est donc :

$$(\mathbf{D}) \quad \begin{array}{ll} \text{Maximiser} & h(u) \\ & u \in U \end{array}$$

On appellera solution primale réalisable ou P-réalisable tout $x \in S$ tel que $g(x) \leq 0$. De même, tout $u \in U$ est une solution duale réalisable ou D-réalisable.

4.3 Point-selle du Lagrangien

Définition 4.1 *On appelle point-selle de la fonction Lagrangien L toute paire (x^*, u^*) telle que :*

$$\forall u \geq 0,\ L(x^*, u) \leq L(x^*, u^*) \leq L(x, u^*),\ \forall x \in S$$

Le point-selle signifie donc que x^* minimise $L(x, u)$ par rapport à x sur S et que u^* maximise $L(x, u)$ par rapport à u sur $u \geq 0$ (voir représentation sur la figure 4.1).

Théorème 4.3 *(x^*, u^*) est un point-selle de $L(x, u)$ si et seulement si :*

i) x^* *minimise* $L(x, u^*)$ *sur* S

ii) $g(x^*) \leq 0$

iii) $\langle u^*, g(x^*)\rangle = 0$

FIGURE 4.1 – Point-selle du Lagrangien

Démonstration i) correspond naturellement à l'inégalité de droite de la définition. L'inégalité de gauche s'écrit :
$\forall u \geq 0, f(x^*) + \langle u, g(x^*)\rangle \leq f(x^*) + \langle u^*, g(x^*)\rangle \Leftrightarrow \langle (u - u^*), g(x^*)\rangle \leq 0$
Quand u_i tend vers $+\infty$, on doit donc avoir $g_i(x^*) \leq 0$.
Par ailleurs, si $u = 0$, $-\langle u^*, g(x^*)\rangle \leq 0$. Comme $u^* \geq 0$ et $g(x^*) \leq 0$, on obtient :
$\langle u^*, g(x^*)\rangle = 0$. □

Théorème 4.4 *Si (x^*, u^*) est un point-selle de L, alors x^* est solution de (P).*

Démonstration D'après le théorème précédent, si (x^*, u^*) est un point-selle de L, x^* minimise $L(x, u^*)$ sur S. Donc, $f(x^*) \leq f(x) + \langle u^*, g(x)\rangle, \forall x \in S$. D'où, $f(x^*) \leq f(x), \forall x \in S$ et tels que $g(x) \leq 0$. C'est-à-dire, x^* est solution de (P). □

L'implication réciproque n'est en général pas vraie. Quand $\sup h < \inf f$, on dit qu'il y a un *saut de dualité*.

Corollaire 4.2 *S'il existe x^* P-réalisable et u^* D-réalisable tels que $h(u^*) = f(x^*)$, alors (x^*, u^*) est un point-selle de L. Par conséquent, x^* est solution optimale de (P) et u^* est solution optimale de (D).*

Il est difficile de donner une condition nécessaire et suffisante d'existence d'un point-selle. Le théorème d'existence le plus précis est dû à Karlin :

Théorème 4.5 *Supposons S convexe et les fonctions $f, g_1, \ldots, g_p$ convexes sur S. Supposons de plus qu'il existe $x_0 \in S$ tel que $g(x_0) < 0$. Si $x^* \in S$ est une solution optimale de (P), alors il existe un vecteur de multiplicateurs $u^* \geq 0$ tel que (x^*, u^*) est un point-selle du Lagrangien $L(x, u)$.*

Démonstration cf. Mangasarian, p. 79. □

L'idée est donc bien de substituer la résolution de (P) jugée difficile par la résolution du problème dual. On a supposé en effet que le calcul de la fonction h en un point u de U est relativement simple. On va voir dans la section suivante que le prix à payer pour cette relaxation est le fait que la fonction duale n'est en général pas différentiable.

4.4 Différentiabilité de la fonction duale

Pour simplifier l'étude, on suppose S compact et les fonctions $f, g_1, \ldots, g_p$ continues sur S. Cela implique $U = (I\!R^p)^+$. La fonction h est alors sous-différentiable sur U (on emploiera par abus de langage les mots sous-différentiel et sous-gradient même s'il s'agit d'une fonction concave, les supports affines correspondants étant alors des majorants de la fonction).

Soit $X(u) = \{x \in S \mid x \text{ minimise } L(x,u) \text{ sur } S\}$. D'après les résultats du chapitre II-3 (Théorème de Danskin), on a :

$$\partial h(u) = \mathrm{conv}\{g(x) \mid x \in X(u)\}$$

Donc, h sera différentiable en $u \in U$ si $g(x)$ est constant sur $X(u)$, ou de manière plus pratique, si le sous-problème $\inf_{x \in S}\ f(x) + \langle u, g(x)\rangle$ possède un unique minimum. Cela sera le cas en particulier si f est strictement convexe, $g_1, \ldots, g_p$ convexes sur S convexe borné.

Exemples :

a) Programmation Linéaire.
Soit le programme linéaire :

$$\begin{aligned} \text{Max} \quad & z = c^T x \\ & Ax \leq b \\ & x \geq 0 \end{aligned}$$

Relaxons les contraintes $Ax \leq b$ pour obtenir le Lagrangien :
$L(x,u) = c^T x + u^T(b - Ax)$ pour tout $u \geq 0$
La fonction duale s'écrit :

$$\begin{aligned} h(u) &= \inf\{(c - A^T u)^T x \mid x \geq 0\} + b^T u \\ &= \begin{cases} b^T u & \text{si } c - A^T u \geq 0, u \geq 0 \\ -\infty & \text{sinon} \end{cases} \end{aligned}$$

b)

$$\begin{aligned} \text{Min} \quad & (x_1 - 1)^2 + (x_2 - 2)^2 \\ & x_1 + x_2 \leq 2 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Relaxons la contrainte $x_1 + x_2 \leq 2$. La fonction duale d'une seule variable u est :

$$\begin{aligned} h(u) &= \inf\{(x_1 - 1)^2 + ux_1 \mid x_1 \geq 0\} + \inf\{(x_2 - 2)^2 + ux_2 \mid x_2 \geq 0\} - 2u \\ &= \begin{cases} -1/2u^2 + u & \text{si } 0 \leq u \leq 2 \\ -1/4u^2 + 1 & \text{si } 2 < u \leq 4 \\ 5 - 2u & \text{si } u > 4 \end{cases} \end{aligned}$$

Remarques :

FIGURE 4.2 – Fonction duale différentiable

- Le calcul a été séparé en deux calculs indépendants sur x_1 et x_2 (on dit qu'on a décomposé le problème).
- La fonction h est concave, donc continue aux points de jonction des différents morceaux et également différentiable sur $u \geq 0$ (cf. figure 4.2).

c)

$$\begin{array}{ll} \text{Min} & -x^2 \\ & 1-2x=0 \\ & 0 \leq x \leq 1 \end{array}$$

La solution est bien sûr $x = 0.5$ avec $f(x) = -0.25$.
En relaxant la contrainte $1 - 2x = 0$, on obtient la fonction duale :
$h(u) = \begin{cases} u & \text{si } u \leq -1/2 \\ -1-u & \text{si } u > -1/2 \end{cases}$
Ici, u n'est pas astreint à être ≥ 0 car la contrainte relaxée est une égalité. On remarque que h n'est pas différentiable en $u = -1/2$ qui est d'ailleurs le maximum de h et que $\max h = -1/2 < \min f$. On a ici un saut de dualité car le problème n'est pas convexe (Fig. 4.3).

4.5 Dualité et perturbations des contraintes

On associe à (P) le problème perturbé suivant :

Figure 4.3 – Saut de dualité

$$(Py)\quad \begin{array}{ll} \text{Minimiser} & f(x) \\ \text{sous} & g(x) \leq y \\ & x \in S \end{array}$$

où $y \in I\!R^p$ est un vecteur de perturbations du second membre. Le problème (P) correspond bien sûr à $y = 0$.

Soit $w(y)$ la fonction perturbation :

$$w(y) = \min\{f(x) \mid g(x) \leq y, x \in S\}$$

définie sur

$$Y = \{y \in I\!R^p \mid \exists x \in S \text{ tel que } g(x) \leq y\}$$

Proposition 4.1 – *La fonction w est monotone non croissante sur Y*
- *Si f, g sont des fonctions convexes et S est convexe fermé, alors Y est un ensemble convexe et w est convexe sur Y.*

Démonstration : exercice

La figure 4.4 montre l'épigraphe de la fonction w dans $I\!R^{p+1}$.

Théorème 4.6 *La fonction duale h associée à (P) satisfait :*

$$h(u) = -w^*(-u)$$

Démonstration Montrons que, pour un $u \geq 0$ tel que $h(u)$ soit fini, on a :

$$y_0 + \langle u, y\rangle \geq h(u), \forall \begin{pmatrix} y \\ y_0 \end{pmatrix} \in \text{epi}(w)$$

Observez que, si $\bar{x} \in S$ est tel que $h(u) = f(\bar{x}) + \langle u, g(\bar{x})\rangle$, la relation ci-dessus signifie que $H = \{\begin{pmatrix} y \\ y_0 \end{pmatrix} \in I\!R^{p+1} \mid y_0 + \langle u, y\rangle = h(u)\}$ est un hyperplan support de l'épigraphe

FIGURE 4.4 – Fonction w(y)

de la fonction w en $\begin{pmatrix} g(\bar{x}) \\ f(\bar{x}) \end{pmatrix}$. En effet, $h(u) = L(\bar{x}, u) \leq f(x) + \langle u, g(x) \rangle, \forall x \in S$, ce qui implique que $h(u) \leq y_0 + \langle u, y \rangle$ pour tout $y_0 \geq f(\bar{x})$ et tout $y \geq g(\bar{x})$. On vérifie alors que $f(\bar{x}) = w(y)$.
En se reportant à la définition de la fonction conjuguée (chap.3), on retrouve la relation cherchée entre h et w^*. $\Box$

4.6 Relaxation Lagrangienne et optimisation combinatoire

De nombreux problèmes d'optimisation combinatoire peuvent se modéliser sous la forme :

$$
\textbf{(IP)} \quad \begin{array}{ll} \text{Minimiser} & c^T x \\ & Ax \geq b \\ & Ex \geq d, x \in Z^n \end{array}
$$

Ici encore, on suppose que les contraintes $Ax \geq b$ sont complicantes et qu'on dispose d'un outil pour minimiser une fonction linéaire sur $S = \{x \in Z^n \,|\, Ex \geq d\}$. Soit $v_{RL} = \sup_{u \geq 0} \inf_{x \in S} c^T x + u^T(b - Ax)$, la valeur maximale du dual associé à la relaxation lagrangienne des contraintes complicantes dans (IP). Le résultat suivant est dû à Geoffrion (1974) :

Théorème 4.7 *Soit v_{RC} la valeur optimale de la relaxation de (IP) obtenue en remplaçant S par conv$S$, l'enveloppe convexe de S. Alors*

$$v_{RL} = v_{RC}$$

Démonstration On remarque d'abord qu'on peut remplacer S par conv$S$ dans la définition de la fonction duale h. En effet, le Lagrangien étant linéaire en x, le minimum est atteint en un sommet de S (supposé borné pour simplifier l'exposé). Soient $x^k, k \in K$ un sommet de S. On a alors :

$$h(u) = \min_{k \in K} [c^T x^k + u^T(b - Ax^k)]$$

La fonction duale est donc concave linéaire par morceaux, chaque support affine étant associé à un sommet de S.

Le problème dual peut donc s'écrire comme un programme linéaire :

$$\begin{array}{ll} \text{Maximiser} & w \\ & w - u^T(b - Ax^k) \leq c^T x^k, k \in K \\ & u \geq 0 \end{array}$$

dont le dual est, en appelant λ_k la variable duale associée à la contrainte k :

$$\begin{array}{ll} \text{Minimiser} & c^T(\sum_k \lambda_k x^k) \\ & A(\sum_k \lambda_k x^k) \geq b \\ & \sum_k \lambda_k = 1 \\ & \lambda_k \geq 0, \forall k \end{array}$$

et on retrouve bien l'enveloppe convexe de S donnée par :

$$\text{conv}(S) = \{x = \sum_k \lambda_k x^k, \sum_k \lambda_k = 1, \lambda_k \geq 0\}$$

□

Observation : Si v_{LP} est la valeur optimale de la relaxation continue de (IP), on a la hiérarchie suivante entre les bornes inférieures :

$$v_{LP} \leq v_{RL} \leq v^*$$

On aura de plus $v_{LP} = v_{RL}$ si $\text{conv}(S) = \{x \in I\!R^n \mid Ex \geq d, x \geq 0\}$ (ce qui revient à dire que le sous-problème a la propriété d'intégralité de ses solutions de base).

Références :

L.S. Lasdon, Optimization for Large Systems, Mac Millan, 1970
C. Lemaréchal, Lagrangian Relaxation, INRIA Report, 2000
O. Mangasarian, Nonlinear Programming, Prentice-Hall, 1969