

Relaxation lagrangienne

Considérons le problème de programmation mathématique suivant:

$$(P) \quad \begin{array}{ll} \text{Min} & f(x) \\ \text{Sujet à} & x \in X_1 \subset R^n \\ & \underline{x \in X_2 \subset R^n} \\ & x \in X \subset R^n \end{array}$$

$\begin{array}{ll} \text{Min} & c^T x \\ \text{Sujet à} & Ax \geq b \\ & \underline{Dx \geq d} \\ & x \geq 0, \text{ entier} \end{array}$

où $f : R^n \rightarrow R^1$.

Supposons que X_1 et X_2 sont engendrés respectivement comme suit:

$$X_1 = \{x \in R^n : f_i(x) \leq 0, i = 1, \dots, m\}$$
$$X_2 = \{x \in R^n : g_i(x) \leq 0, i = 1, \dots, q\}.$$

Si les contraintes $x \in X_1$ et $x \in X$ sont des contraintes **faciles** à traiter, alors il est intéressant de trouver un moyen d'en tirer avantage en éliminant les contraintes $x \in X_2$.

Par contre il faut trouver un moyen de simuler leur présence dans le problème résultant.

Considérons le problème de programmation mathématique suivant:

$$(P) \quad \text{Min } f(x)$$
$$\text{Sujet à } \begin{aligned} x &\in X_1 \subset R^n \\ x &\in X_2 \subset R^n \\ x &\in X \subset R^n \end{aligned}$$

où $f : R^n \rightarrow R^1$.

Supposons que X_1 et X_2 sont engendrés respectivement comme suit:

$$X_1 = \{x \in R^n : f_i(x) \leq 0, i = 1, \dots, m\}$$

$$X_2 = \{x \in R^n : g_i(x) \leq 0, i = 1, \dots, q\}. \quad \boxed{\lambda_i}$$

Le principe de la relaxation lagrangienne consiste à remplacer la résolution de (P) par celle d'une suite de problèmes de la forme suivante (**dénotés relaxations lagrangiennes**):

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$
$$\text{Sujet à } \begin{aligned} f_i(x) &\leq 0 & i = 1, \dots, m \\ x &\in X. \end{aligned}$$

obtenus en utilisant diverses valeurs pour les multiplicateurs

$$\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0.$$

Le principe de la relaxation lagrangienne consiste à remplacer la résolution de (P) par celle d'une suite de problèmes de la forme suivante (**dénotés relaxations lagrangiennes**):

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$
$$\text{Sujet à } \quad f_i(x) \leq 0 \quad i = 1, \dots, m$$
$$x \in X.$$

obtenus en utilisant diverses valeurs pour les multiplicateurs

$$\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0.$$

Ainsi, les contraintes considérées comme **contrariantes**

$g_1(x), \dots, g_q(x)$ sont introduites dans la fonction économique avec des multiplicateurs $\lambda_1, \dots, \lambda_q \geq 0$, afin de simuler leur violation, en quelque sorte.

Le principe de la relaxation lagrangienne consiste à remplacer la résolution de (P) par celle d'une suite de problèmes de la forme suivante (**dénotés relaxations lagrangiennes**):

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$

$$\text{Sujet à } \begin{array}{l} f_i(x) \leq 0 \quad i = 1, \dots, m \\ x \in X. \end{array}$$

obtenus en utilisant diverses valeurs pour les multiplicateurs

$$\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0.$$

$$(P) \quad \begin{array}{l} \text{Min } f(x) \\ \text{Sujet à } \begin{array}{l} x \in X_1 \subset R^n \\ x \in X_2 \subset R^n \\ x \in X \subset R^n \end{array} \end{array}$$

Se référant à la théorie de la dualité lagrangienne, nous notons que $G(\lambda)$ est la fonction économique du dual lagrangien (D) du problème (P) par rapport aux contraintes $g_1(x), \dots, g_q(x)$:

$$(D) \quad \text{Max}_{\lambda \geq 0} \{ G(\lambda) \} = \text{Max}_{\lambda \geq 0} \left\{ \text{Min}_{\substack{x \in X_1 \\ x \in X}} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\} \right\}$$

Le principe de la relaxation lagrangienne consiste à remplacer la résolution de (P) par celle d'une suite de problèmes de la forme suivante (dénotés relaxations lagrangiennes) :

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$

$$\text{Sujet à } \begin{array}{l} f_i(x) \leq 0 \quad i = 1, \dots, m \\ x \in X. \end{array}$$

obtenus en utilisant diverses valeurs pour les multiplicateurs

$$\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0.$$

(P)	$\text{Min } f(x)$
Sujet à	$x \in X_1 \subset R^n$ $x \in X_2 \subset R^n$ $x \in X \subset R^n$

$$(D) \quad \text{Max}_{\lambda \geq 0} \{ G(\lambda) \} = \text{Max}_{\lambda \geq 0} \left\{ \text{Min}_{\substack{x \in X_1 \\ x \in X}} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\} \right\}$$

Ainsi, par le théorème de dualité faible, la valeur optimale $G(\lambda)$ du problème (P_λ) est une borne inférieure sur la valeur optimale de (P) .

C'est dans ce sens que (P_λ) est une relaxation de (P) .

Le principe de la relaxation lagrangienne consiste à remplacer la résolution de (P) par celle d'une suite de problèmes de la forme suivante (dénotés relaxations lagrangiennes) :

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$

$$\text{Sujet à } \quad f_i(x) \leq 0 \quad i = 1, \dots, m$$

$$x \in X.$$

obtenus en utilisant diverses valeurs pour les multiplicateurs

$$\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0.$$

$$(P) \quad \text{Min } f(x)$$

$$\text{Sujet à } \quad x \in X_1 \subset R^n$$

$$x \in X_2 \subset R^n$$

$$x \in X \subset R^n$$

De plus, se référant aux conditions d'optimalité du théorème du lagrangien, si nous pouvons identifier un vecteur de multiplicateurs $\lambda \geq 0$ tel que la solution optimale x^λ de (P_λ) satisfait les conditions:

$$g_i(x^\lambda) \leq 0 \quad i = 1, \dots, q \quad (\text{réalisabilité})$$

$$\lambda_i g_i(x^\lambda) = 0 \quad i = 1, \dots, q \quad (\text{complémentarité})$$

alors x^λ est une solution optimale de (P) .

À chaque itération de la **méthode de relaxation lagrangienne**, un problème

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$

$$\text{Sujet à } \begin{array}{l} f_i(x) \leq 0 \quad i = 1, \dots, m \\ x \in X. \end{array}$$

est résolu pour un vecteur de multiplicateurs $\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0$ fixé.

Le problème à résoudre à la prochaine itération est obtenu en modifiant les valeurs des multiplicateurs $\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0$.

Comment modifier le vecteur des multiplicateurs pour que la procédure résultante ait du sens et converge vers quelque chose de pertinent.

La réponse repose sur le fait que le vecteur $\lambda \geq 0$ est une solution réalisable du dual (D)

$$(D) \quad \text{Max}_{\lambda \geq 0} \{G(\lambda)\} = \text{Max}_{\lambda \geq 0} \left\{ \text{Min}_{\substack{x \in X_1 \\ x \in X}} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\} \right\}$$

À chaque itération de la **méthode de relaxation lagrangienne**, un problème

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$

$$\text{Sujet à } \begin{array}{l} f_i(x) \leq 0 \quad i = 1, \dots, m \\ x \in X. \end{array}$$

est résolu pour un vecteur de multiplicateurs $\lambda = [\lambda_1, \dots, \lambda_q]^T \geq 0$ fixé.

La réponse repose sur le fait que le vecteur $\lambda \geq 0$ est une solution réalisable du dual (D)

$$(D) \quad \text{Max}_{\lambda \geq 0} \{G(\lambda)\} = \text{Max}_{\lambda \geq 0} \left\{ \text{Min}_{\substack{x \in X_1 \\ x \in X}} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\} \right\}$$

$$(P) \quad \begin{array}{l} \text{Min } f(x) \\ \text{Sujet à } \begin{array}{l} x \in X_1 \subset R^n \\ x \in X_2 \subset R^n \\ x \in X \subset R^n \end{array} \end{array}$$

Ainsi $\lambda \geq 0$ est modifié à chaque itération de sorte à optimiser le problème dual (D) et par le fait même à générer la meilleure borne inférieure sur la valeur optimale du problème (P) .

Comme nous l'avons observé dans la section sur la dualité lagrangienne, la fonction $G(\lambda)$

$$(P_\lambda) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$

Sujet à $f_i(x) \leq 0 \quad i = 1, \dots, m$
 $x \in X.$

est concave en λ .

Si la fonction $G(\lambda)$ était continuellement différentiable, il suffirait de résoudre le dual (D) avec une méthode itérative où à chaque itération la solution actuelle est modifiée en se déplaçant dans une direction s'alignant sur le gradient de la fonction évalué à la solution actuelle.

Du fait que $G(\lambda)$ est concave, il s'ensuit que cette fonction atteint un maximum global en un point où le gradient s'annule.

$$\begin{aligned}
 (P_\lambda) \quad G(\lambda) &= \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\} \\
 \text{Sujet à} \quad & f_i(x) \leq 0 \quad i = 1, \dots, m \\
 & x \in X.
 \end{aligned}$$

Mais la fonction $G(\lambda)$ n'est pas continuellement différentiable en général. Pour ce genre de problème, des méthodes itératives ont été développées où le gradient est remplacé par le sous-gradient de la fonction.

Heureusement il est facile d'identifier un sous-gradient pour la fonction $G(\lambda)$ à une solution actuelle $\tilde{\lambda}$ du dual.

Si $x^{\tilde{\lambda}}$ dénote une solution optimale de $(P_{\tilde{\lambda}})$, alors le vecteur

$$\left[g_1(x^{\tilde{\lambda}}), \dots, g_q(x^{\tilde{\lambda}}) \right]^T$$

est un sous-gradient de G au point $\tilde{\lambda}$.

$$(P_{\lambda}) \quad G(\lambda) = \text{Min} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) \right\}$$

Sujet à $f_i(x) \leq 0 \quad i = 1, \dots, m$
 $x \in X.$

Si $x^{\tilde{\lambda}}$ dénote une solution optimale de $(P_{\tilde{\lambda}})$, alors le vecteur

$$\left[g_1(x^{\tilde{\lambda}}), \dots, g_q(x^{\tilde{\lambda}}) \right]^T$$

est un sous-gradient de G au point $\tilde{\lambda}$.

$$G(\lambda) \leq G(\tilde{\lambda}) + \left[g_1(x^{\tilde{\lambda}}), \dots, g_q(x^{\tilde{\lambda}}) \right]^T (\lambda - \tilde{\lambda})$$

En effet, pour tout $\lambda \geq 0$

$$\begin{aligned} G(\tilde{\lambda}) + \left[g_1(x^{\tilde{\lambda}}), \dots, g_q(x^{\tilde{\lambda}}) \right]^T (\lambda - \tilde{\lambda}) & \\ &= f(x^{\tilde{\lambda}}) + \sum_{i=1}^q \tilde{\lambda}_i g_i(x^{\tilde{\lambda}}) + \sum_{i=1}^q (\lambda_i - \tilde{\lambda}_i) g_i(x^{\tilde{\lambda}}) \\ &= f(x^{\tilde{\lambda}}) + \sum_{i=1}^q \lambda_i g_i(x^{\tilde{\lambda}}) \\ &\geq \text{Inf}_{x \in X} \left\{ f(x) + \sum_{i=1}^q \lambda_i g_i(x) : f_i(x) \leq 0, i = 1, \dots, m \right\} \\ &= G(\lambda) \end{aligned}$$

Notes importantes:

- i)* Le fait de résoudre le problème dual (D) ne nous assure en rien de trouver une solution optimale de (P).
- ii)* À cause du saut de dualité qui peut exister entre les valeurs optimales de (P) et de (D) lorsque les hypothèses de convexité et de stabilité ne sont pas satisfaites, il se peut que la borne inférieure obtenue avec la relaxation lagrangienne soit strictement inférieure à la valeur optimale de (P).
- iii)* Tout de même à chaque itération de la relaxation lagrangienne, la valeur de $G(\lambda)$ constitue une borne inférieure sur la valeur optimale de (P) qui est d'autant meilleure que λ se rapproche d'une solution optimale de (D).