

Master d'Informatique - Spécialité Androide
Module MAOA

Recherche Opérationnelle
et Optimisation Combinatoire

(Hors-Programme) Décomposition et relaxation
Lagrangienne

Pierre Fouilhoux

Sorbonne Université

2019-2020

Reformulation

Pour un même problème, il existe plusieurs formulations PLNE utilisant des variables différentes.

Lorsque l'on change l'espace des variables d'un problème, on parle souvent de **reformulation** (c'est une façon de dire qu'il existe plusieurs formulations).

Qu'est une bonne (re)formulation ? On la juge :

- en fonction de *la valeur de relaxation* obtenue : est-ce une amélioration à la formulation précédente ? Plus elle est proche de l'optimum entier, plus la formulation a des chances de bien se comporter dans un algorithme de branchement.
- en fonction des possibilités efficaces de branchement : éviter qu'il existe trop de solutions *symétriques* c'est-à-dire qui ont un même coût et/ou des structures proches : par exemple deux ensembles de sommets d'un graphe indentiques à un sommet près. Plus il y a de symétries, plus l'algorithme de branchement va devoir explorer de solutions...
- en fonction de la structure de la formulation : un programme peut-être "presque" découpable en sous-programmes indépendants. Des **techniques de décomposition** permettent d'utiliser cet structure pour construire des principes algorithmiques particuliers (décomposition de Dantzig-Wolfe, de Benders, par relaxation lagrangienne).

1. Décompositions

2. Relaxations

3. Relaxation Lagrangienne

4. Décomposition par relaxation lagrangienne

Décomposition

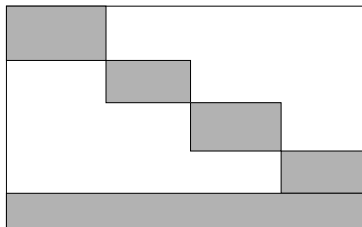
Un PLNE (P) est défini par les vecteurs et matrices suivantes :

- $M \in \mathbf{R}^{n \times m}$ matrice des coefficients
- $v \in \mathbf{R}^m$ vecteur de droite (Right Hand Side (rhs))
- $c \in \mathbf{R}^n$ vecteur coût

Et les variables entières (ou mixtes) $x \in \mathbf{R}^n$.

$$(P) \begin{cases} \text{Min } c^T x \\ Mx \geq v \\ x \in \mathbf{N}^n \end{cases}$$

Dans les problèmes pratiques, la matrice M des coefficients a bien souvent cette structure bloc-diagonale où les zones grises représentent les coefficients non-nuls (parfois après des permutations simultanées des lignes et des colonnes) :



Décomposition

Seul le bloc du bas concernent toutes les variables du problèmes : on appelle les contraintes de ce bloc **les contraintes couplantes**.

Les autres blocs diagonaux sont eux plus petits et ne concernent qu'un lot limité de variables où ces variables apparaissent exclusivement.

On appelle **résolution frontale** le fait de résoudre un PLNE sans utiliser cette composition de la matrice.

On appelle **décomposition** d'un PLNE est le fait d'utiliser la composition du PLNE en plusieurs blocs pour résoudre le PLNE en divisant la résolution en plusieurs programmes et sous-programmes.

Vocabulaire :

On appelle **problème maître** le PLNE obtenu à partir du PLNE d'origine qui va donner la solution du problème d'origine : il conserve en général les contraintes couplantes et joue un rôle de pilote dans la méthode de décomposition.

Et **sous-problèmes auxiliaires (ou satellites)** des sous-programmes qui vont aider à résoudre la totalité du problème : le plus souvent ces sous-problèmes seront associés aux blocs isolés de la matrice M des coefficients.

Décomposition

L'ensemble du programme maître et des programmes satellites peut alors s'intégrer dans un processus algorithmique qui converge vers une borne inférieure ou vers la solution optimale du problème.

Suivant la manière selon laquelle on décompose le problème et la technique mathématique utilisée, on obtient des décompositions diverses pour lesquelles plusieurs techniques algorithmiques ont été mises au point. On peut citer principalement :

- **la décomposition par relaxation lagrangienne**, dite aussi "décomposition par les prix" : on utilise la relaxation lagrangienne des contraintes couplantes en les glissant dans la fonction objective. On obtient ainsi un problème maître qui pilote la résolution en fixant un prix (le coût lagrangien) associés aux inégalités couplantes. Ce problème maître est alors un ensemble de PLNE indépendants : chacun d'entre eux sont alors vu comme des sous-problèmes.
Attention : cette méthode converge vers une borne inférieure du problème et non vers une solution.
- **la génération de colonnes et branchement** en reformulant un problème avec un nombre exponentiel de variables, on peut obtenir un programme ayant une matrice décomposable. Le principe algorithmique de la génération de colonnes peut-être alors vu comme un algorithme de décomposition où le problème-maître pilote les sous-problèmes par les coûts duaux de ses inégalités.

Décomposition

- **la décomposition de Dantzig-Wolfe** : il existe une technique quasi-automatique pour reformuler un PLNE en un programme décomposable possédant un nombre exponentiel de variables. Une telle décomposition, dite de Dantzig-Wolfe, se résoud alors par génération de colonnes et branchement (on confond d'ailleurs souvent les deux termes).
- **la décomposition de Benders** : il existe une technique quasi-automatique pour reformuler un PLNE en un programme décomposable possédant un nombre exponentiel d'inégalités. Une telle décomposition, dite de Benders, se résoud alors par une méthode de coupes (potentiellement non polynomiale) et branchement.

La mise en oeuvre d'une décomposition demande à chaque fois une étude dédiée pour la mettre en oeuvre. Ces techniques de décomposition ont mené à de grandes réussites sur des problèmes industriels de très grandes tailles.

1. Décompositions

2. Relaxations

3. Relaxation Lagrangienne

4. Décomposition par relaxation lagrangienne

Les bonnes performances de la programmation mathématique reposent essentiellement sur sa capacité à fournir des bornes intéressantes (borne inf en maximisation et borne sup en minimisation).

Ces bornes peuvent être obtenues par **relaxation** du PLNE, c'est-à-dire :

- soit le fait d'élargir l'ensemble des solutions de manière à optimiser sur un espace plus large (et plus facile à optimiser) : cela comporte la relaxation de contraintes (et en particulier la relaxation continue).
- soit le fait de remplacer la fonction objective qui est une borne supérieure (en maximisation) (par exemple la relaxation lagrangienne).

Relaxations continues

Pour un programme mathématique, les relaxations continues possibles sont :

- la **(simple) relaxation continue** qui est, pour la PLNE (par l'algorithme du simplexe, des points intérieurs ou du volume) et la programmation quadratique (avec matrice définie positive) très efficace. Néanmoins, nous verrons que la relaxation continue est rarement la meilleure !
- En ce qui concerne plus spécifiquement la PLNE, on peut améliorer la valeur de relaxation linéaire en ajoutant des contraintes, le **renforcement** : ce que l'on a vu dans les chapitres précédents ! C'est très très efficace !
- On peut également ajouter des variables, la **reformulation**. Nous verrons cela dans le dernier chapitre
- Une autre piste est la **relaxation Lagrangienne** que l'on voit ici en "hors-programme".

1. Décompositions
2. Relaxations
3. Relaxation Lagrangienne
4. Décomposition par relaxation lagrangienne

Relaxation lagrangienne

Un cadre particulièrement intéressant pour la relaxation (mais aussi pour la décomposition voir plus loin) est celui donné par la **dualité lagrangienne**.

Principe :

- choisir une inégalité dans le PL et l'enlever du PL.
- ajouter cette inégalité dans la fonction objective avec un multiplicateur visant à pénaliser la fonction objective si la solution trouvée ne vérifie pas cette inégalité.

Cette relaxation fournit fréquemment de très bonne valeur de relaxation, bien souvent meilleure que la relaxation continue.

On considère ici un PLNE écrit de la façon suivante :

$$\begin{array}{ll} \text{Min } f(x) \\ g_i(x) \leq 0 & i \in I \\ x \in S \subset \mathbf{R}^n. \end{array}$$

On associe à chaque contrainte $i \in I$ une valeur réelle $\lambda_i \geq 0$ que l'on appelle **multiplicateur de Lagrange**. On considère également le vecteur réel $\lambda = (\lambda_i, i \in I)$. On pose alors la *fonction de Lagrange*

$$L(x, \lambda) = f(x) + \sum_{i \in I} \lambda_i g_i(x)$$

qui est une fonction qui a pour paramètres les deux vecteurs x et λ et prend ses valeurs dans \mathbf{R} . On suppose ici que le PLNE est tel que le problème $\text{Min}_{x \in S} \{L(x, \lambda)\}$ a une solution optimale pour tout $\lambda \geq 0$.

Prenons par exemple le PLNE suivant où $S = \{0, 1\}^3$

$$\begin{aligned} \text{Min } f(x) &= 10 - 3x_1 - 2x_2 - x_3 \\ 2x_1 + 3x_2 + 4x_3 &\leq 4 \\ x &= (x_1, x_2, x_3) \in \{0, 1\}^3. \end{aligned}$$

L'ensemble des points de S est

$$y^1 = (0, 0, 0), y^2 = (1, 0, 0), y^3 = (0, 1, 0), y^4 = (1, 1, 0), y^5 = (0, 0, 1), \\ y^6 = (1, 0, 1), y^7 = (0, 1, 1) \text{ et } y^8 = (1, 1, 1).$$

Parmi elles, on trouve une solution optimale au PLNE pour $y^2 = (1, 0, 0)$ qui vaut $f(y^2) = 7$.

En posant $g(x) = 2x_1 + 3x_2 + 4x_3 - 4$, on définit la fonction de Lagrange pour chaque point y^i de S comme étant $\forall \lambda \in \mathbf{R}_+$,

$$L(y^i, \lambda) = f(y^i) + \lambda g(y^i) = 10 - 3y_1^i - 2y_2^i - y_3^i + (2y_1^i + 3y_2^i + 4y_3^i - 4)\lambda$$

c'est-à-dire

$$\begin{array}{ll} L(y^1, \lambda) = 10 - 4\lambda & \forall \lambda \in \mathbf{R}_+, & L(y^5, \lambda) = 9 & \forall \lambda \in \mathbf{R}_+, \\ L(y^2, \lambda) = 7 - 2\lambda & \forall \lambda \in \mathbf{R}_+, & L(y^6, \lambda) = 6 + 2\lambda & \forall \lambda \in \mathbf{R}_+, \\ L(y^3, \lambda) = 8 - \lambda & \forall \lambda \in \mathbf{R}_+, & L(y^7, \lambda) = 7 + 3\lambda & \forall \lambda \in \mathbf{R}_+, \\ L(y^4, \lambda) = 5 + \lambda & \forall \lambda \in \mathbf{R}_+, & L(y^8, \lambda) = 4 + 5\lambda & \forall \lambda \in \mathbf{R}_+. \end{array}$$

Pour des vecteurs \bar{x} et $\bar{\lambda} \geq 0$ donnés, on dit que le couple $(\bar{x}, \bar{\lambda})$ est un *point-selle* de la fonction de Lagrange si

$$\begin{aligned} L(\bar{x}, \lambda) &\leq L(\bar{x}, \bar{\lambda}) && \forall \lambda \geq 0 \\ L(\bar{x}, \bar{\lambda}) &\leq L(x, \bar{\lambda}) && \forall x \in S. \end{aligned}$$

En d'autre terme, un point-selle $(\bar{x}, \bar{\lambda})$ est tel que la fonction L pour x fixé à \bar{x} atteind un maximum en $\bar{\lambda}$ et que la fonction L pour λ fixé à $\bar{\lambda}$ atteind un minimum en \bar{x} .

On a alors le résultat suivant pour caractériser les points-selles :

Theorem

Etant donnés des vecteurs \bar{x} et $\bar{\lambda} \geq 0$, le couple $(\bar{x}, \bar{\lambda})$ est un point-selle de la fonction de Lagrange si et seulement si

i) $L(\bar{x}, \bar{\lambda}) = \text{Min}_{x \in S} \{L(x, \bar{\lambda})\}$

ii) $g_i(\bar{x}) \leq 0, \forall i \in I,$

iii) $\bar{\lambda}_i g_i(\bar{x}) = 0, \forall i \in I.$

Le théorème suivant permet de faire le lien entre point-selle et optimum du PLNE.

Theorem

Etant donné un point-selle $(\bar{x}, \bar{\lambda})$ pour L défini par les vecteurs \bar{x} et $\bar{\lambda} \geq 0$, alors \bar{x} est un optimum global pour le PMD.

Ces théorèmes permettent ainsi d'étudier la fonction de Lagrange pour résoudre le PLNE. Cette relaxation ramène parfois à des problèmes d'optimisation continue **en ôtant des contraintes difficiles**.

Malheureusement un PLNE n'a pas toujours de point-selle !

Dualité lagrangienne

On considère la fonction, dite **duale lagrangienne**, suivante :

$$w(\lambda) = \inf_{x \in S} \{L(x, \lambda)\} \quad \forall \lambda \geq 0.$$

Dans le cas qui nous intéresse ici, c'est-à-dire des ensembles S discrets finis, cette fonction peut s'écrire plus simplement

$$w(\lambda) = \min_{x \in S} \{L(x, \lambda)\} \quad \forall \lambda \geq 0.$$

On définit alors le *problème dual (lagrangien)* du PM :

$$\text{Max}_{\lambda \in \mathbf{R}_+^m} w(\lambda) \quad (= \text{Max}_{\lambda} \{ \text{Min}_{x \in S} L(x, \lambda) \})$$

En fait, ce problème dual correspond au fait de rechercher un point-selle pour L . On peut néanmoins noter que ce problème est défini même s'il n'existe pas de point-selle (comme c'est le cas pour un PLNE).

On a alors les théorèmes suivants de la dualité.

Theorem

(Théorème de la dualité faible)

Si l'on note $f(x^)$ l'optimum global de f et $w(\lambda^*)$ l'optimum global du problème dual, alors*

Pour tout $\lambda \geq 0$, $w(\lambda) \leq w(\lambda^) \leq f(x^*)$.*

Theorem

(Concavité de la fonction duale)

La fonction duale $\lambda \mapsto w(\lambda)$ est une fonction concave de λ .

Ainsi la fonction w est concave !!!

Ce qui permet d'utiliser des outils efficaces de l'optimisation continue. Il est important de noter qu'aucune hypothèse n'a été faite sur la nature des fonctions f et g_i qui peuvent être quelconque, ni sur la convexité de l'ensemble S .

Dans le cas des ensembles S discrets, la fonction f n'est pas différentiable en tout point mais la concavité de w permet cependant de montrer que tout optimum local de w est global : ce qui fait que le problème dual est en général plus facile à résoudre que le problème dual.

Enfin, ce dernier théorème indique que, si le PLNE admet un point-selle, optimiser w revient à optimiser f .

Theorem

(Théorème de la dualité)

Si le problème PM admet un point-selle (x^, λ^*) alors $w(\lambda^*) = f(x^*)$.*

Inversement, s'il existe x^ et $\lambda^* \geq 0$ tels que $w(\lambda^*) = f(x^*)$, alors le PM admet un point-selle et (x^*, λ^*) est un tel point-selle.*

En fait, cette théorie de la dualité permet de ramener, lorsqu'il y a des points-selles, le PLNE à la minimisation d'une fonction concave. De plus, on sait facilement utiliser des méthodes de sous-gradient pour cette fonction. C'est le cas pour la dualité lagrangienne appliquée à un PL est en fait équivalente à la dualité "classique" de la PL. **mais pas pour la PLNE.**

Saut de dualité

Malheureusement, un PLNE ne possède pas en général de point-selle, ce dernier théorème n'est cependant pas applicable pour S discret, l'écart $f(x^*) - w(\lambda^*)$ est souvent non nul.

On appelle *saut de dualité* le fait que la valeur $f(x^*) - w(\lambda^*)$ soit non nul. C'est le cas lorsqu'un PM ne possède de point-selle pour sa fonction duale associée. C'est le cas par exemple de la plupart des PMD.

Dans l'exemple du PLNE à 3 variables.

La fonction duale du PMD est $w(\lambda) = \text{Min}_{i \in \{1, \dots, 8\}} \{L(y^i, \lambda)\} \quad \forall \lambda \geq 0$ qui est facile à calculer pour un λ donné. On obtient par exemple $w(0) = 4$, $w(0,5) = 5,5$ et $w(1) = 5$.

On veut en fait résoudre le programme dual qui est alors $\text{Max}_{\lambda \geq 0} w(\lambda)$. Or la fonction w est concave ! On sait donc par ces trois valeurs (en 0, 0,5 et 1) que son optimum sera atteint sur l'intervalle $[0, 1]$. On peut ainsi, par exemple par recherche dichotomique, on obtient alors un maximum pour que l'on peut visualiser comme étant le point le plus haut de l'enveloppe des droites $L(y^i, \lambda)$, c'est-à-dire, après calcul simple, $\lambda^* = \frac{2}{3}$ et $w(\lambda^*) = \frac{17}{3}$.

(On peut noter que comme la relaxation linéaire de ce problème est égale à la relaxation lagrangienne, on obtient également un optimum fractionnaire à $x_R = (1, \frac{2}{3}, 0)$ qui donne $f(x_R) = \frac{17}{3}$.)

Ainsi, pour cet exemple, on obtient un saut de dualité de $f(y^2) - w(\lambda^*) = \frac{4}{3}$.

La relaxation lagrangienne

La théorie de la dualité lagrangienne trouve son application la plus fréquente dans le cadre de la *relaxation lagrangienne* c'est-à-dire le fait d'appliquer la fonction de Lagrange pour un sous-ensemble bien choisi de contraintes.

Considérons le PM suivant :

$$\begin{aligned} \text{Min } & f(x) \\ & g_i(x) \leq 0 \quad i \in I \\ & x \in S \subset \mathbf{R}^n. \end{aligned}$$

où S est un ensemble non triviale pouvant être lui-même défini par des contraintes. On choisira d'appliquer la relaxation lagrangienne aux contraintes g_i dans le cas où :

- le problème de minimisation de la fonction lagrangienne obtenue sur S doit pouvoir être fait de manière efficace et exacte.
- l'ensemble des contraintes g_i relaxée n'est pas trop important et bien choisie pour limiter le nombre de multiplicateur à manipuler.

L'objectif de cette relaxation est de donner de bonne borne d'évaluation du problème. Cependant il n'est pas rare que l'on obtienne également des solutions primales intéressantes.

Contraintes couplantes

Un exemple très fréquent en pratique se situe dans le cadre des PLNE où certaines contraintes sont jugées “difficiles” à traiter. On peut alors écrire le PLNE de la façon suivante :

$$\begin{array}{ll} \text{Min} & c^T x \\ (P) & Ax \leq b \\ & Cx \leq d \\ & x \in \mathbf{N}^n. \end{array}$$

où les contraintes $Ax \leq b$ sont des contraintes que l'on voudrait ôter du programme : c'est principalement le cas si elles sont **couplantes** c'est-à-dire si elles concernent des variables de différentes sections du PLNE (voir section Décomposition).

Qualité de la relaxation Lagrangienne

Pour mesurer la qualité de la relaxation Lagrangienne, on va tenter de comparer la valeur d'évaluation obtenue par la relaxation linéaire du PLNE et celle obtenue par la relaxation lagrangienne L des contraintes $Ax \leq B$, c'est-à-dire les valeurs de

$$(P_R) \quad \begin{array}{l} \text{Min } c^T x \\ Ax \leq b \\ Cx \leq d \\ x \in \mathbf{R}^n \end{array} \quad \text{et} \quad (D_L) \quad \begin{array}{l} \text{Max}_\lambda \text{ Min}_{x \in S} \{L(x, \lambda)\} \\ \text{avec } S = \{x \in \mathbf{N}^n \mid Cx \leq d\}. \end{array}$$

En fait, on peut prouver le résultat suivant

Lemma

La valeur optimale du dual lagrangien (D_L) est égale à la valeur optimale du problème suivant :

$$(P') \quad \begin{array}{l} \text{Min } c^T x \\ Ax \leq b \\ x \in \text{conv}(S). \end{array}$$

Ainsi dans le cas où l'enveloppe convexe de S est telle que $\text{conv}(S) \subsetneq \{x \in \mathbf{R}^n \mid Cx \leq d\}$, ce qui est le cas plus fréquent en PLNE, on voit que la relaxation lagrangienne donne une meilleure évaluation que la relaxation linéaire.

Application au problème du voyageur de commerce

On illustre le principe de la relaxation lagrangienne à l'exemple célèbre du voyageur de commerce.

Soit $G = (V, E)$ un graphe pour lequel on recherche un circuit hamiltonien de plus petite longueur où $c(e)$, $e \in E$, est la longueur associée à une arête.

On choisit un sommet particulier que l'on nomme sommet 1 et on nomme $G \setminus 1$ le graphe obtenu à partir de G en supprimant dans G le sommet 1 et ses arêtes incidentes. On appelle un *1-arbre* de G un arbre-couvrant de $G \setminus 1$ auquel on ajoute deux arêtes quelconque incidente à 1. On notera par la suite \mathcal{T} l'ensemble des 1-arbre de G .

La suite de cet exercice repose sur la remarque suivante : "un cycle hamiltonien est un 1-arbre pour lequel chaque sommet a exactement un degré 2".

Pour $T \in \mathcal{T}$ un 1-arbre de G , on pose le *vecteur d'incidence* χ^T tel que pour toute arête $e \in E$, $\chi^T(e) = 1$ si $e \in T$ et 0 sinon. On note également $\chi(\mathcal{T}) = \{\chi^T \mid T \in \mathcal{T}\}$, c'est-à-dire l'ensemble des vecteurs d'incidence des 1-arbres de G .

On pose alors les variables binaires suivantes : $x(e) = 1$ si on choisit l'arête $e \in e$ dans la solution et 0 sinon et on formule le problème selon le PMD suivant :

$$\begin{aligned} \text{Min } & \sum_{e \in E} c(e)x(e) \\ & \sum_{e \in \delta(u)} x(e) = 2 \quad \forall u \in V \\ & x \in \chi(\mathcal{T}). \end{aligned}$$

On remarque que le vecteur x solution est nécessairement un 1-arbre tel que le degré de chacun de ses sommets est exactement 2.

Etant donné un poids w positif associé aux arêtes de G , on sait résoudre efficacement le problème suivant : optimiser une fonction linéaire $\sum_{e \in E} \bar{c}(e)x(e)$ sur l'ensemble $\chi(\mathcal{T})$. Ceci se fait en recherchant un arbre couvrant de poids minimum au sens du poids w dans $G \setminus 1$ et en lui ajoutant les deux arêtes incidentes à 1 de plus faibles valeurs selon w .

Pour ramener le problème du voyageur de commerce à la résolution du problème précédent, on va réaliser une relaxation lagrangienne des contraintes $\sum_{e \in \delta(u)} x(e) = 2 \quad \forall u \in V$. Pour chaque sommet $u \in V$, on pose les contraintes $g_u(x) = \sum_{e \in \delta(u)} x(e) - 2$ pour tout $x \in \chi(\mathcal{T})$.

La fonction de Lagrange est alors :

$$L(x, \lambda) = c^T x + \sum_{u \in V} \lambda_u g_u(x) \quad \forall \lambda \in \mathbf{R}_+^n.$$

On remarque que cette fonction est la somme, pour chaque arête $e = uv \in E$, de la variable $x(e)$ multipliée par la somme de $c(e)$ et de $\lambda_u + \lambda_v$. Ainsi

$$L(x, \lambda) = \sum_{uv \in E} (c(uv) + \lambda_u + \lambda_v)x(uv) - 2 \sum_{u \in V} \lambda_u.$$

On remarque alors que le dernier terme de L est la constante $K = -2 \sum_{u \in V} \lambda_u$ lorsque λ est fixé. De plus, en posant $\bar{c}(uv) = c(uv) + \lambda_u + \lambda_v$ pour toute $uv \in E$, on a $L(x, \lambda) = \sum_{uv \in E} \bar{c}(uv)x(uv) - K$ qui est alors une fonction linéaire. On sait donc calculer pour tout vecteur $\lambda \geq 0$, la valeur de la fonction duale

$$w(\lambda) = \text{Min}_{x \in \mathcal{X}(\mathcal{T})} L(x, \lambda).$$

On peut facilement appliquer un algorithme de sous-gradient pour calculer le maximum de la fonction w sur λ (en fait, il existe même des formules automatique pour calculer ici le gradient). Les expérimentations de cette technique permettent de résoudre le problème de voyageur de commerce pour des tailles intéressantes : en effet, on dispose d'algorithme permettant de découvrir une solution x correspondant aux bonnes valeurs de λ . En utilisant simplement un algorithme pour déterminer une bonne valeur de λ , on donne également une très bonne évaluation du coût d'une tournée hamiltonienne.

1. Décompositions

2. Relaxations

3. Relaxation Lagrangienne

4. Décomposition par relaxation lagrangienne

4.1 Le Unit Commitment Problem

4.2 Relaxation Lagrangienne pour l'UCP

1. Décompositions

2. Relaxations

3. Relaxation Lagrangienne

4. Décomposition par relaxation lagrangienne

4.1 Le Unit Commitment Problem

4.2 Relaxation Lagrangienne pour l'UCP

Le Unit Commitment Problem (UCP)

On considère le problème de planification de production d'énergie appelé *Unit Commitment Problem* (UCP). On considère ici une version simplifiée pour l'exemple.

Une période de planification $\{1, \dots, T\}$ consiste en T *pas de temps* d'une demi-heure pour lesquels on connaît à l'avance la prévision de production électrique consommée. On appelle D_t la demande en kWatt\h à produire sur le pas de temps $t \in \{1, \dots, T\}$. Le problème consiste à planifier quelles unités de production doivent être en marche à chaque pas de temps de la période.

On considère n unités de production d'énergie électrique (hydro-électricité, centrales thermiques,...).

Chaque unité a un état "up" (en-marche) et un état "down" (non-en-marche).

L'objectif du problème est de décider si une unité est up ou down à chaque pas de temps de la période.

Le Unit Commitment Problem (UCP)

Chaque centrale $i \in \{1, \dots, n\}$ est caractérisée par les données suivantes :

- la production P^i d'électricité (en kWatt.h) qui est produite sur un pas de temps par l'unité si elle est up (si elle est down, elle ne produit rien sur tout le pas de temps)
- le nombre minimal de pas de temps I^i de repos : c'est-à-dire que si l'unité est up au pas de temps t et que l'on désire qu'elle soit down en $t + 1$ alors elle doit rester down de $t + 1$ à $t + I^i$ avant de pouvoir être éventuellement mis up en $t + I^i + 1$
- le nombre minimal de pas de temps L^i d'activité : c'est-à-dire que si l'unité est down au pas de temps t et que l'on désire qu'elle soit up en $t + 1$ alors elle doit rester up de $t + 1$ à $t + L^i$ avant de pouvoir être éventuellement mis down en $t + L^i + 1$
- un coût de production c^i fixe par pas de temps si elle est sur un pas de temps.

Le problème UCP consiste à décider l'état up/down de chacune de unités, en minimisant le coût total de production, en respectant les contraintes de repos/activité et en satisfaisant le fait que le total, pour chaque pas de temps, de la production soit supérieure ou égale à la demande.

Le Unit Commitment Problem (UCP)

Comme même avec un seul pas de temps ($T = 1$), ce problème est déjà aussi-difficile que le problème du sac-à-dos entier, ce problème est NP-difficile.

On peut modéliser ce problème en utilisant les variables suivantes :

- $x_t^i \in \{0, 1\}$ indique si l'unité i est up au temps t .
- $u_t^i \in \{0, 1\}$ indique si l'unité i démarre au temps t .

$$\begin{aligned}
 \text{Min } & \sum_{i=1}^n \sum_{t=1}^T c^i x_t^i \\
 & u_t^i \geq x_t^i - x_{t-1}^i & \forall i \in \{1, \dots, n\}, \forall t \in \{2, \dots, T\}, \\
 & \sum_{t'=t-L^i+1}^t u_{t'}^i \leq x_t^i & \forall i \in \{1, \dots, n\}, \forall t \in \{L^i, \dots, T\}, \\
 & \sum_{t'=t-l^i+1}^t u_{t'}^i \leq 1 - x_{t-l^i}^i & \forall i \in \{1, \dots, n\}, \forall t \in \{l^i, \dots, T\}, \\
 & \sum_{i=1}^n P^i x_t^i \geq D_t & \forall t \in \{1, \dots, T\}, \\
 & x_t^i, u_t^i \in \{0, 1\} & \forall i \in \{1, \dots, n\}, \forall t \in \{1, \dots, T\}.
 \end{aligned}$$

Le Unit Commitment Problem (UCP)

La première inégalité si une unité i passe de down à up, alors u_t^i est égale à 1.
Les deux suivantes obligent à respecter les contraintes de repos et d'activité.
Enfin la dernière inégalité exige qu'à chaque pas de temps, la production soit d'au moins D_t .

On peut noter que cette formulation correspond exactement au cas de structure bloc-diagonale! En effet, les 3 premières inégalités peuvent être groupées par unité et seule la contrainte de production est couplante!

Il apparaît très naturel de décomposer par unités de production!

1. Décompositions

2. Relaxations

3. Relaxation Lagrangienne

4. Décomposition par relaxation lagrangienne

4.1 Le Unit Commitment Problem

4.2 Relaxation Lagrangienne pour l'UCP

Relaxation Lagrangienne pour l'UCP

Relaxons les contraintes de production par relaxation lagrangienne.

On utilise des multiplicateurs de lagrange $\lambda_t \geq 0$. Le dual lagrangien est alors

$$\max_{\lambda \in \mathbb{R}_+^T} \left(\min \left(\sum_{i=1}^n \sum_{t=1}^T c^i x_t^i - \sum_{t=1}^T \lambda_t \left(\sum_{i=1}^n P^i x_t^i - D_t \right) \right) \right)$$

$$\begin{aligned} u_t^i &\geq x_t^i - x_{t-1}^i && \forall i \in \{1, \dots, n\}, \forall t \in \{2, \dots, T\}, \\ \sum_{t'=t-L^i+1}^t u_{t'}^i &\leq x_t^i && \forall i \in \{1, \dots, n\}, \forall t \in \{L^i, \dots, T\}, \\ \sum_{t'=t-l^i+1}^t u_{t'}^i &\leq 1 - x_{t-l^i}^i && \forall i \in \{1, \dots, n\}, \forall t \in \{l^i, \dots, T\}, \\ x_t^i, u_t^i &\in \{0, 1\} && \forall i \in \{1, \dots, n\}, \forall t \in \{1, \dots, T\}. \end{aligned}$$

La particularité ici est que le problème est directement décomposée : en effet, il s'agit de n sous-problèmes d'optimisation de la production : un par unité !

Relaxation Lagrangienne pour l'UCP

Pour un vecteur λ donné, les coût sont eux-aussi décomposables :

$$\min \sum_{i=1}^n \sum_{t=1}^T (c^i - P^i \lambda_t) x_t^i - \sum_{t=1}^T \lambda_t D_t$$

Pour une unité i donnée, le sous-problème est alors

$$\begin{aligned} & \sum_{t=1}^T (c^i - P^i \lambda_t) x_t^i \\ & u_t^i \geq x_t^i - x_{t-1}^i \quad \forall t \in \{2, \dots, T\}, \\ & \sum_{t'=t-L^i+1}^t u_{t'}^i \leq x_t^i \quad \forall t \in \{L^i, \dots, T\}, \\ & \sum_{t'=t-J^i+1}^t u_{t'}^i \leq 1 - x_{t-J^i}^i \quad \forall t \in \{J^i, \dots, T\}, \\ & x_t^i, u_t^i \in \{0, 1\} \quad \forall t \in \{1, \dots, T\}. \end{aligned}$$

Or ce sous-problème est très simple à résoudre : c'est un problème polynomial qui se résoud par programmation dynamique (ou même par ce PLNE car le problème est en fait entièrement caractérisé par ces 3 inégalités : c'est donc un PL dont tous les points extrêmes sont entiers!).

Relaxation Lagrangienne pour l'UCP

Dans ce cas, la relaxation lagrangienne donne directement une décomposition naturelle du problème : on peut nommer cette décomposition, **décomposition par les prix** car les multiplicateurs de Lagrange agissent comme un régulateur économique qui pilote la production de chaque unité en fonction d'un prix à payer pour produire ou pour ne pas produire : il indique à chaque sous-problème des coûts (positifs ou négatifs) pour la production de chacune des unités.

On a vu dans le chapitre sur les relaxations que résoudre un dual lagrangien consiste à maximiser une fonction concave sur \mathbf{R}_+^T : on peut utiliser des techniques d'optimisation continue pour cela.

Ce processus semble assez simple à mettre en place. Toutefois :

- la relaxation lagrangienne est une relaxation : elle ne mène pas à une solution entière optimale.
- mais dans ce cas de l'UCP, elle consiste à piloter des solutions entières : parmi-elles, il y a de très bonne solution proche de l'optimale.
- la convergence des méthodes d'optimisation continue pour le dual lagrangien est parfois difficile à obtenir : des versions plus réalistes du problème de l'UCP ont d'ailleurs conduit à inventer des méthodes avancées de convergence (comme la méthode des faisceaux (bundle method) de C. LeMaréchal.)