La programmation mathématique en nombres entiers

Décompositions

Stanislas Francfort

Orange Labs

5 Juin 2017

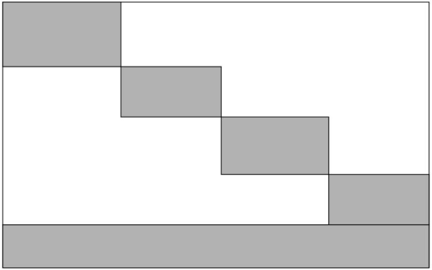
Décomposition

- On appelle *décomposition* d'un PLNE $Ax \leq b$ le fait de le diviser en plusieurs sous-problèmes
- grâce à la structure de la matrice des contraintes A

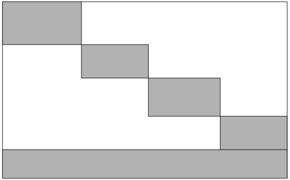
- On note A la *matrice de contraintes*
- Une colonne par variable du PLNE
- Une ligne par contrainte du PLNE

Décomposition

- Pafois la structure de A est très particulière
- éventuellement après rangement (parfois peu évident)



Décomposition



- Le bloc du bas concerne toutes les variables du problèmes
- on appelle ces contraintes des *contraintes couplantes*
- Les autres blocs diagonaux sont eux plus petits et sont indépendants les uns des autres

Décomposition

- Chaque bloc diagonal : une sous-partie des variables et des contraintes
- Ils sont appelés problèmes *satellites*

- Ces sous-problèmes n'ont pas de sens par rapport à l'optimisation globale du programme
- **Mais** : on peut par contre guider la résolution de chaque sous-problème en les reliant entre eux
- on réalise ce lien en modifiant chaque sous-problème a partir de données obtenues a partir d'un programme *maître* qui contient les contraintes couplantes.
- L'ensemble du programme maître et des programmes satellites peut alors s'intégrer dans un processus algorithmique qui converge vers la solution optimale du problème.

Décomposition

- Plusieurs manière de décomposer le problème
- Plusieurs techniques mathématiques utilisées pour recombiner
  - *décomposition Lagrangienne* : relaxation Lagrangienne des contraintes couplantes pour obtenir des sous-problèmes indépendants. Utilisation des multiplicateurs de Lagrange.
  - *décomposition de Dantzig-Wolfe* : cette décomposition s'appuie sur un découplage des solutions en utilisant les points extrêmes du polyèdre.
  - *décomposition de Benders* : un problème maître associé aux contraintes couplantes et des problèmes-satellites associé à chaque bloc de la matrice. On rassemble les différents résultats dans un processus itératif utilisant les coûts duaux du problème maître dans les problèmes satellites

Décomposition

- On va chercher à résoudre un PLNE composé de deux ensembles de contraintes

$$
\begin{aligned}
&Max\ c^T x \\
&\qquad Bx \leq b \text{ (m contraintes)} \\
&\qquad Ax \leq a \text{ (p contraintes)} \\
&\qquad x \in \mathbb{N}^n
\end{aligned}
$$

Décomposition Lagrangienne

- On utilise cette décomposition lorsque la conjonction de deux blocs de contraintes couplantes ($Ax \leq a$ et $Bx \leq b$) rend le problème difficile
- On pose :

$$
\begin{aligned}
Max\ c^T x & \\
Bx &\leq b \\
x &= y \\
Ay &\leq a
\end{aligned}
$$

- on relâche la contrainte $x = y$ par la méthode du multiplicateur de Lagrange

Décomposition de Dantzig-Wolfe

- On pose $X = \{x \in \mathbb{N}^n \mid Bx \leq b\}$ (qui est discret)
- et on note $(x_i)_{i \in I}$ l'ensemble de tous ses point
- Alors (PLNE) est équivalent à (PLNE-DW) :

$$\begin{aligned} Max \ & \sum_{i \in I} (c_i x_i) \lambda_i \\ & \sum_{i \in I} (A^i) x_i \lambda_i \leq a \\ & \lambda_i \in \{0, 1\}, \ \forall i \in I \end{aligned}$$

Décomposition de Dantzig-Wolfe

- (PLNE-DW) est équivalent à résoudre le PLNE initial
- Mais leurs relaxation continues sont différentes
- la relaxation continue de PLNE-DW ne relâche que les contraintes d'intégrité correspondants à $Ax \leq b$
- meilleure relaxation continue de (PLNE-DW) que de (PLNE)

Décomposition de Dantzig-Wolfe

Theorem

Soit $X = \{x \in \mathbb{R}^n \mid Bx \leq b\}$, un polyèdre borné non vide. $x \in X$ si et seulement si x peut s'écrire comme combinaison convexe des points extrêmes de P
c'est à dire :
$X = \{x \in \mathbb{R}^n \mid x = \sum_{j \in J} \lambda_j x_j \mid \sum_{j \in J} \lambda_j = 1, \lambda_j \geq 0, \forall j \in J\}$
où J est l'ensemble des points extrêmes de X

- Décomposition de Dantzig-Wolfe : revient à exprimer un polyèdre convexe par l'intermédiaire de ses points extrémaux

Décomposition de Dantzig-Wolfe

- remplaçons l'expression de $x \in X$ dans (PLNE-DW) :

$$
\begin{aligned}
Max \ & \sum_{j \in J} \lambda_j c_j x_j \\
& \sum_{j \in J} \lambda_j A^j x_j \leq a \\
& \sum_{j \in J} \lambda_j = 1 \\
& \lambda_j \in [0, 1], \ \forall j \in J
\end{aligned}
$$

Décomposition de Dantzig-Wolfe

- En posant $\hat{c}_j = c_j x_j$ et $\hat{A}^j = A^j x_j$, on obtient :

$$
\begin{aligned}
Max \ & \sum_{j \in J} \lambda_j \hat{c}_j \\
& \sum_{j \in J} \lambda_j \hat{A}^j \leq a \\
& \sum_{j \in J} \lambda_j = 1 \\
& \lambda_j \in [0,1], \ \forall j \in J
\end{aligned}
$$

- $m + 1$ contraintes (au lieu de $m + p$)
- Mais : une variable par point extrême de X...

Décomposition de Dantzig-Wolfe

- Pour résoudre le problème précédent :
- relaxation continue (en λ)
- génération dynamique des variables pour obtenir une base optimale (génération de colonnes)
- On cherche la variable hors base de coût réduit maximal (Rappel : coût réduit = réduction de l'optimum si on abaisse la contrainte saturée correspondante)
- en résolvant le problème suivant : $Max\{\hat{c}_j - u^t \hat{A}^j \mid j \in J\}$
- (u est le multiplicateur du dual correspondant aux contraintes couplantes $\sum_{j \in J} \hat{A}^j \lambda^j \leq a$)
- si ce maximum est ≤ 0, alors on stoppe, car on a atteint l'optimum

Décomposition de Dantzig-Wolfe

- Si B est bloc-diagonale (B_k les sous-matrices)
- Alors on peut exprimer les solutions de (PLNE-DW) en fonction des points extrêmes des X^k
- avec $X^k = \{y \in \mathbb{N}^{n_k} \mid B_k y \leq b_k\}$, pour $k = 1, \ldots, p$

Décomposition de Dantzig-Wolfe

- On obtient alors le *problème-père* (ou *problème-maitre*) $(\hat{Q})$:

$$
\begin{aligned}
Max \ & \sum_{k=1}^{p} \sum_{j \in J_k} \lambda_j^k \hat{c_j^k} \\
& \sum_{k=1}^{p} \sum_{j \in J_k} \lambda_j^k \hat{A_k^j} \leq a \\
& \sum_{j \in J_k} \lambda_j^k = 1, \ \forall k = 1, \ldots, p \\
& \lambda_j^k \in \{0, 1\}, \ \forall j \in J_k, \ \forall k = 1, \ldots, p
\end{aligned}
$$

Décomposition de Dantzig-Wolfe

- Le nombre de variables de $(\hat{Q})$ est le nombre de points extrêmes des X^k
- On résoud $(\hat{Q})$ par génération de colonnes
- Chaque sous-problèmes fournit des points extrêmes du polyèdre X^k associé
- Approche intéressante lorsque le nombre de blocs indépendants est grand

Décomposition de Benders

- Utilisé principalement dans des problèmes à variables mixtes
- Principe : si au cours de la résoution, une solution partielle ne peut pas être complétée en solution réalisable
  - Alors toutes les solutions engendrées sont ignorées
- On partitionne les variables en deux vecteurs x et y
- On fixe le vecteur y
- On regarde uniquement le sous-problème basé sur x
- Ceci peut nous révéler des y non réalisables (avec l'aide du dual)

Décomposition de Benders

$$
\begin{aligned}
Max\ & cx + sy \\
& Ax + By \leq b \\
& y \in Y \\
& x \in \mathbb{R}_+^{n_1}
\end{aligned}
$$

Avec Y un espace quelconque de dimension n_2 ($n = n_1 + n_2$). Par exemple, $Y = \mathbb{N}^{n_2}$

Décomposition de Benders

En fixant le vecteur y, on obtient :

$$\begin{aligned} &Max\ cx \\ &\quad Ax \leq b - By \\ &\quad x \in \mathbb{R}_+^{n_1} \end{aligned}$$

Et son dual (D_y) :

$$\begin{aligned} &Min\ (b - By)^T u \\ &\quad uA \geq c \\ &\quad u \in \mathbb{R}_+^{m} \end{aligned}$$

Décomposition de Benders

Le problème d'optimisation s'écrit :

$$
\begin{aligned}
& Max\ \{sy + cx \mid Ax + By \leq b, y \in Y, x \in \mathbb{R}_+^{n_1}\} \\
= & Max\ \{sy + Max\ \{cx \mid Ax \leq b - By, x \in \mathbb{R}_+^{n_1}\} \mid y \in Y\} \\
= & Max\ \{sy + Min\ \{(b - By)^T u \mid uA \geq c, u \in \mathbb{R}_+^{m}\} \mid y \in Y\}
\end{aligned}
$$

Décomposition de Benders

En linéarisant la dernière ligne, on obtient le *problème maître de Benders* :

$$\begin{aligned} Max\ & w \\ & w \leq sy + (b - By)^T u_y \\ & y \in Y \\ & w \in \mathbb{R} \end{aligned}$$

- Avec u_y une solution optimale du dual (D_y), pour tout y
- Le principe consiste alors à alimenter le problème maître avec des solutions u_y qui sont optimales pour (D_y)

Décomposition de Benders

- l'ajout de nouvelles solutions u_y se traduit par des contraintes
- ces coupes restreignent petit à petit le domaine d'optimisation du problème maître
- La difficulté est d'énoncer les $w \leq sy + (b - By)^T u_y$, où u_y est défini implicitement comme solution optimale du dual
- Pour éviter d'avoir à lister toutes les contraintes, on part d'un problème maître non contraint
- puis on ajoute petit à petit les contraintes par un algorithme de génération de coupes