

Du sous-problème de séparation vers le sous-problème d'intersection dans la décomposition de Benders

Daniel Porumbel

<http://cedric.cnam.fr/~porumbed/>

Contexte :

On veut minimiser un PL avec un
nombre exorbitant de contraintes

⇒ Plans Coupants

- 1 La séparation et le concept d'intersection
- 2 La décomposition Benders et son sous-problème d'intersection
- 3 Résultats et conclusions

Sous-problème de séparation

$$\min \{ \mathbf{d}^\top \mathbf{y} : \mathbf{u}^\top \mathbf{y} \geq b, \forall (\mathbf{u}, b) \in \mathcal{C} \} = \min \{ \mathbf{d}^\top \mathbf{y} : \mathbf{y} \in \mathcal{Z} \},$$

- \mathcal{C} : un ensemble avec un **nombre énorme de contraintes**
- \mathcal{Z} : Zone (polytope) réalisable

Algo Plans Coupants classique

- On optimise à chaque itération un PL avec un sous-ensemble des contraintes \mathcal{C}
- Soit \mathbf{r} la solution optimale à l'itération courante. On cherche une contrainte violée via le sous-problème de séparation :

$$\min_{(\mathbf{u}, b) \in \mathcal{C}} \mathbf{u}^\top \mathbf{r} - b$$

Sous-problème de séparation

$$\min \{ \mathbf{d}^\top \mathbf{y} : \mathbf{u}^\top \mathbf{y} \geq b, \forall (\mathbf{u}, b) \in \mathcal{C} \} = \min \{ \mathbf{d}^\top \mathbf{y} : \mathbf{y} \in \mathcal{Z} \},$$

- \mathcal{C} : un ensemble avec un **nombre énorme de contraintes**
- \mathcal{Z} : Zone (polytope) réalisable

Algo Plans Coupants classique

- On optimise à chaque itération un PL avec un sous-ensemble des contraintes \mathcal{C}
- Soit \mathbf{r} la solution optimale à l'itération courante. On cherche une contrainte violée via le sous-problème de séparation :

$$\max_{(\mathbf{u}, b) \in \mathcal{C}} b - \mathbf{u}^\top \mathbf{r}$$

Pour quoi s'embêter avec un sous-problème d'intersection ?

- Soit $\mathbf{r} = [1 \ 1 \ 1]^T$ la solution optimale à l'itération courante, Il faut choisir une contrainte :

(1) $4y_1 + 6y_2 + 8y_3 \geq 20$

(2) $y_1 + 2y_2 + y_3 \geq 5$

La séparation choisit (1)

! Sous-problème d'intersection : normaliser et séparer !

- Cela revient à maximiser :

$$\max_{(u,b) \in \mathcal{C}} \frac{b}{\mathbf{u}^T \mathbf{r}}$$

- L'intersection choisit (2) :

(1) $\frac{20}{4+6+8} = 1 + \frac{1}{8}$

<

(2) $\frac{5}{1+2+1} = 1 + \frac{1}{4}$

(2) est plus violée que (1)
par $\mathbf{r} = [1 \ 1 \ 1]^T$

Pour quoi s'embêter avec un sous-problème d'intersection ?

- Soit $\mathbf{r} = [1 \ 1 \ 1]^T$ la solution optimale à l'itération courante, Il faut choisir une contrainte :

$$(1) \quad 4y_1 + 6y_2 + 8y_3 \geq 20$$

$$(2) \quad y_1 + 2y_2 + y_3 \geq 5$$

La séparation choisit (1)

! Sous-problème d'intersection : normaliser et séparer !

- Cela revient à maximiser :

$$\max_{(\mathbf{u}, b) \in \mathcal{C}} \frac{b}{\mathbf{u}^T \mathbf{r}}$$

- L'intersection choisit (2) :

$$(1) \quad \frac{20}{4+6+8} = 1 + \frac{1}{9}$$

<

$$(2) \quad \frac{5}{1+2+1} = 1 + \frac{1}{4}$$

(2) est plus violée que (1)
par $\mathbf{r} = [1 \ 1 \ 1]^T$

Pour quoi s'embêter avec un sous-problème d'intersection ?

- Soit $\mathbf{r} = [1 \ 1 \ 1]^T$ la solution optimale à l'itération courante, Il faut choisir une contrainte :

$$(1) \quad 4y_1 + 6y_2 + 8y_3 \geq 20$$

$$(2) \quad y_1 + 2y_2 + y_3 \geq 5$$

La séparation choisit (1)

! Sous-problème d'intersection : normaliser et séparer !

- Cela revient à maximiser :

$$\max_{(\mathbf{u}, b) \in \mathcal{C}} \frac{b}{\mathbf{u}^T \mathbf{r}}$$

- L'intersection choisit (2) :

$$(1) \quad \frac{20}{4+6+8} = 1 + \frac{1}{9}$$

<

$$(2) \quad \frac{5}{1+2+1} = 1 + \frac{1}{4}$$

(2) est plus violée que (1)
par $\mathbf{r} = [1 \ 1 \ 1]^T$

Deuxième avantage : l'intersection trouve des solutions réalisables et bornes sup

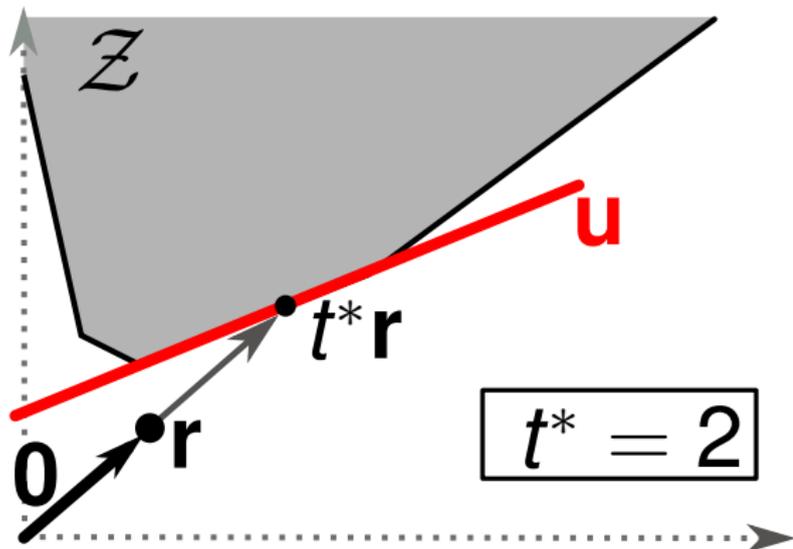
$$t^* = \max_{(\mathbf{u}, b) \in \mathcal{C}} \frac{b}{\mathbf{u}^\top \mathbf{r}}$$

\Rightarrow

$$\mathbf{u}^\top (t^* \mathbf{r}) \geq b \quad \forall (\mathbf{u}, b) \in \mathcal{C}$$

Donc $t^* \mathbf{r}$ est une solution réalisable qui donne une borne sup à chaque itération !

- Impossible avec un Plans coupants standard (via la séparation) avant la fin de la convergence



Deuxième avantage : l'intersection trouve des solutions réalisables et bornes sup

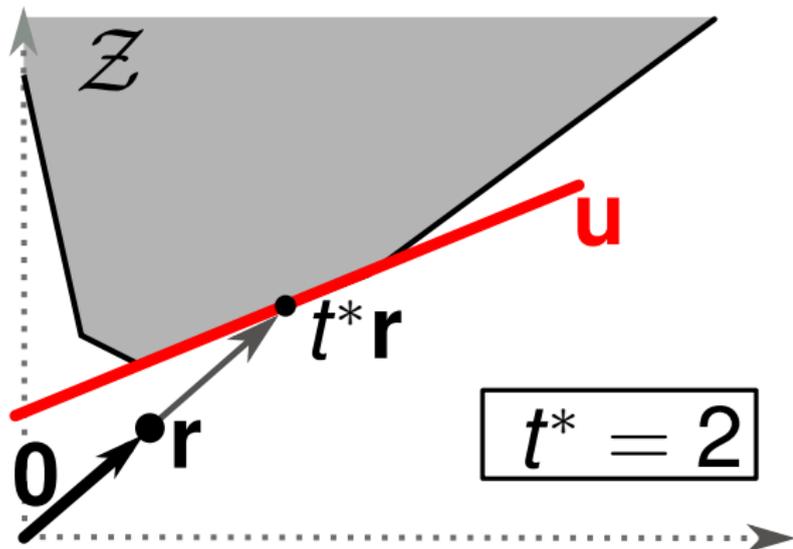
$$t^* = \max_{(\mathbf{u}, b) \in \mathcal{C}} \frac{b}{\mathbf{u}^\top \mathbf{r}}$$

\Rightarrow

$$\mathbf{u}^\top (t^* \mathbf{r}) \geq b \quad \forall (\mathbf{u}, b) \in \mathcal{C}$$

Donc $t^* \mathbf{r}$ est une solution réalisable qui donne une borne sup à chaque itération !

- Impossible avec un Plans coupants standard (via la séparation) avant la fin de la convergence



Résoudre le sous-problème d'intersection

1 L'intersection généralise la séparation

2 L'intersection demande de maximiser une fraction

$$\max_{(\mathbf{u}, b) \in \mathcal{C}} \frac{b}{\mathbf{u}^T \mathbf{r}}$$

et la séparation maximise un objectif linéaire

$$\max_{(\mathbf{u}, b) \in \mathcal{C}} b - \mathbf{u}^T \mathbf{r}$$



Dans l'absolu, le sous-problème d'intersection est plus difficile.

! Dans certains cas, le sous-problème d'intersection peut être aussi tractable que celui de séparation :

A Si on travaille avec le dual d'un PL de génération de colonnes et on peut utiliser la programmation dynamique.

- On associe un état à chaque sous-problème (cf Bellman)
- il faut trouver l'état qui maximise l'objectif. Pas important si l'objectif est une fonction linéaire, une fraction, etc

B Si on est dans une décomposition de Benders, il faut maximiser une **programme linéaire-fractionnaire** qui peut être **converti en PL** à l'aide de la transformation Charness-Cooper.

Étapes et Plan

- 1 La séparation et le concept d'intersection
- 2 La décomposition Benders et son sous-problème d'intersection
- ~~3 Exemple de problème résolu par Benders à l'aide de Plans Coupants basés sur l'intersection~~
- ~~4 Plans Coupants avancés pour la séparation et l'intersection~~
- ~~5 Implementation en C++~~
- 6 Résultats et conclusions

Étapes et Plan

- 1 La séparation et le concept d'intersection
- 2 La décomposition Benders et son sous-problème d'intersection
- 3
- 4
- 5
- 6 Résultats et conclusions

Du modèle de départ aux reformulations

coût des flux, on va utiliser $\mathbf{c} = \mathbf{0}$

$$\min \mathbf{d}^\top \mathbf{y} + \mathbf{c}^\top \mathbf{x}$$

$$\mathbf{D}\mathbf{y} \geq \mathbf{e}$$

Contraintes de design

$$\mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{x} \geq \mathbf{b}$$

$$\mathbf{y} \in \mathbb{Z}_+^n, \mathbf{x} \geq \mathbf{0}$$

Le flux \mathbf{x} doit pouvoir passer

\mathbf{y} est un nombre d'unités à faire fonctionner.

- câbles à monter
- entrepôts à ouvrir

\mathbf{x} est un coût des flux qui passent, coût d'affectations, etc

Du modèle de départ aux reformulations

coût des flux, on va utiliser $\mathbf{c} = \mathbf{0}$

$$\min \mathbf{d}^\top \mathbf{y} + \mathbf{c}^\top \mathbf{x}$$

$$\mathbf{D}\mathbf{y} \geq \mathbf{e}$$

Contraintes de design

$$\mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{x} \geq \mathbf{b}$$

$$\mathbf{y} \in \mathbb{Z}_+^n, \mathbf{x} \geq \mathbf{0}$$

Le flux \mathbf{x} doit pouvoir passer

Reformulation 1 :

$$\min \mathbf{d}^\top \mathbf{y} + \hat{z}$$

$$\mathbf{D}\mathbf{y} \geq \mathbf{e}$$

$$\hat{z} = \min \{ \mathbf{c}^\top \mathbf{x} : \mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}$$

$$\mathbf{y} \in \mathbb{Z}_+^n$$

Du modèle de départ aux reformulations

$$\begin{aligned} \min \quad & \mathbf{d}^\top \mathbf{y} + \mathbf{c}^\top \mathbf{x} \\ & \mathbf{D}\mathbf{y} \geq \mathbf{e} \\ & \mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{y} \in \mathbb{Z}_+^n, \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Reformulation 1 :

$$\begin{aligned} \min \quad & \mathbf{d}^\top \mathbf{y} + \hat{z} \\ & \mathbf{D}\mathbf{y} \geq \mathbf{e} \\ & \hat{z} = \min \{ \mathbf{c}^\top \mathbf{x} : \mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0} \} \\ & \mathbf{y} \in \mathbb{Z}_+^n \end{aligned}$$

On va dualiser ce PL

Du modèle de départ aux reformulations

coût des flux, on va utiliser $\mathbf{c} = \mathbf{0}$
 $\Rightarrow \hat{z} = 0$

$$\begin{aligned} \min \mathbf{d}^\top \mathbf{y} + \mathbf{c}^\top \mathbf{x} \\ \mathbf{Dy} \geq \mathbf{e} \\ \mathbf{By} + \mathbf{Ax} \geq \mathbf{b} \\ \mathbf{y} \in \mathbb{Z}_+^n, \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Reformulation 2 :

$$\begin{aligned} \min \mathbf{d}^\top \mathbf{y} + \hat{z} \\ \mathbf{Dy} \geq \mathbf{e} \\ \hat{z} = \max\{(\mathbf{b} - \mathbf{By})^\top \mathbf{u} : \mathbf{u} \in \mathcal{P}\}, \\ \mathbf{y} \in \mathbb{Z}_+^n \end{aligned}$$

$$\mathcal{P} = \{\mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{c}\}$$

Du modèle de départ aux reformulations

coût des flux, on va utiliser $\mathbf{c} = \mathbf{0}$
 $\Rightarrow \hat{z} = 0$

$$\begin{aligned} \min \mathbf{d}^\top \mathbf{y} + \mathbf{c}^\top \mathbf{x} \\ \mathbf{Dy} \geq \mathbf{e} \\ \mathbf{By} + \mathbf{Ax} \geq \mathbf{b} \\ \mathbf{y} \in \mathbb{Z}_+^n, \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Reformulation 3 :

$$\begin{aligned} \min \mathbf{d}^\top \mathbf{y} + \hat{z} \\ \mathbf{Dy} \geq \mathbf{e} \\ \hat{z} = \max\{(\mathbf{b} - \mathbf{By})^\top \mathbf{u} : \mathbf{u} \in \mathcal{P}\} \\ 0 \geq (\mathbf{b} - \mathbf{By})^\top \mathbf{u} \quad \forall \mathbf{u} \in \mathcal{P}, \\ \mathbf{y} \in \mathbb{Z}_+^n \\ \mathcal{P} = \{\mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{0}\} \end{aligned}$$

Du modèle de départ aux reformulations

$$\begin{aligned} \min \quad & \mathbf{d}^\top \mathbf{y} + \mathbf{c}^\top \mathbf{x} \\ \mathbf{D}\mathbf{y} \geq & \mathbf{e} \\ \mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{x} \geq & \mathbf{b} \\ \mathbf{y} \in \mathbb{Z}_+^n, \quad & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Reformulation 4 :

$$\begin{aligned} \min \quad & \mathbf{d}^\top \mathbf{y} \\ \mathbf{D}\mathbf{y} \geq & \mathbf{e} \\ \mathbf{B}\mathbf{y}^\top \mathbf{u} \geq & \mathbf{b}^\top \mathbf{u} \quad \forall \mathbf{u} \in \mathcal{P}, \\ \mathbf{y} \in \mathbb{Z}_+^n & \end{aligned}$$

$$\mathcal{P} = \{ \mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{0} \}$$

Les coupes Benders sont définies par les rayons \mathbf{u} du polytope Benders \mathcal{P}

Du modèle de départ aux reformulations

$$\min \mathbf{d}^\top \mathbf{y} + \mathbf{c}^\top \mathbf{x}$$

$$\mathbf{D}\mathbf{y} \geq \mathbf{e}$$

$$\mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{x} \geq \mathbf{b}$$

$$\mathbf{y} \in \mathbb{Z}_+^n$$

Reformulation 4 :

$$\min \mathbf{d}^\top \mathbf{y}$$

$$\mathbf{D}\mathbf{y} \geq \mathbf{e}$$

$$\mathbf{B}\mathbf{y}^\top \mathbf{u} \geq \mathbf{b}^\top \mathbf{u} \quad \forall \mathbf{u} \in \mathcal{P},$$

$$\mathbf{y} \in \mathbb{Z}_+^n$$

$$\mathcal{P} = \{\mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{0}\}$$

Sous-probl. séparation sur $\mathbf{y} = \mathbf{r}$

$$\max_{\mathbf{u} \in \mathcal{P}} \mathbf{b}^\top \mathbf{u} - \mathbf{B}\mathbf{r}^\top \mathbf{u}$$

Sous-probl. d'intersection sur \mathbf{r}

$$\max_{\mathbf{u} \in \mathcal{P}} \frac{\mathbf{b}^\top \mathbf{u}}{\mathbf{B}\mathbf{r}^\top \mathbf{u}}$$

Résolution sous-problème d'intersection

Il faut résoudre

$$t^* = \max_{\mathbf{u} \in \mathcal{P}} \frac{\mathbf{b}^\top \mathbf{u}}{\mathbf{B} \mathbf{r}^\top \mathbf{u}}, \text{ avec } \mathcal{P} = \{\mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{0}\}$$

Résolution sous-problème d'intersection

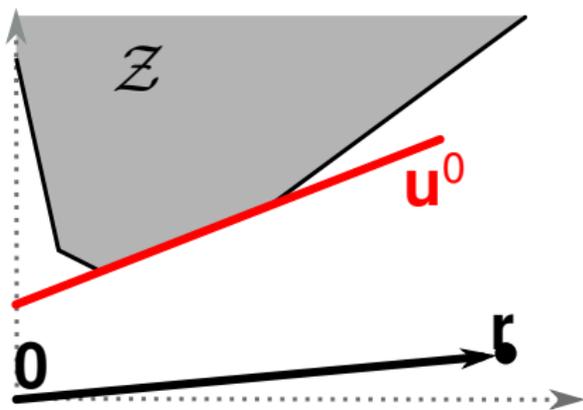
Il faut résoudre

$$t^* = \max_{\mathbf{u} \in \mathcal{P}} \frac{\mathbf{b}^\top \mathbf{u}}{\mathbf{B}\mathbf{r}^\top \mathbf{u}}, \text{ avec } \mathcal{P} = \{\mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{0}\}$$

1 $\exists \mathbf{u}^0 \in \mathcal{P}$ tel que $\mathbf{B}\mathbf{r}^\top \mathbf{u}^0 = 0$
et $\mathbf{b}^\top \mathbf{u}^0 > 0 \implies t^* = \infty$

- t^* non réalisable $\forall t \geq 0$: le rayon $\mathbf{0} \rightarrow \mathbf{r}$ ne **touche pas** la zone réalisable \mathcal{Z}

On détecte ce cas à l'aide de :
 $\max \{\mathbf{b}^\top \mathbf{u} : \mathbf{B}\mathbf{r}^\top \mathbf{u} = 0, \mathbf{u} \in \mathcal{P}\}$

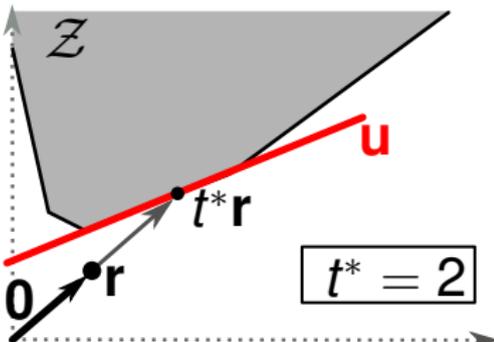


Résolution sous-problème d'intersection

Il faut résoudre

$$t^* = \max_{\mathbf{u} \in \mathcal{P}} \frac{\mathbf{b}^\top \mathbf{u}}{\mathbf{B}\mathbf{r}^\top \mathbf{u}}, \text{ avec } \mathcal{P} = \{\mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{0}\}$$

2 Le cas classique résout :

$$t^* = \max \left\{ \frac{\mathbf{b}^\top \mathbf{u}}{\mathbf{B}\mathbf{r}^\top \mathbf{u}} : \mathbf{B}\mathbf{r}^\top \mathbf{u} > 0, \mathbf{u} \in \mathcal{P} \right\}$$


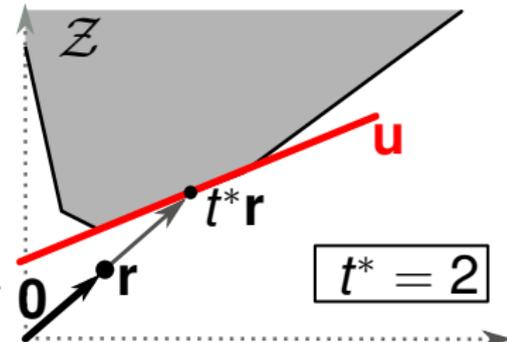
$t^* = 2$

Résolution sous-problème d'intersection

Il faut résoudre

$$t^* = \max_{\mathbf{u} \in \mathcal{P}} \frac{\mathbf{b}^\top \mathbf{u}}{\mathbf{B}\mathbf{r}^\top \mathbf{u}}, \text{ avec } \mathcal{P} = \{\mathbf{u} \geq \mathbf{0} : \mathbf{A}^\top \mathbf{u} \leq \mathbf{0}\}$$

2 Le cas classique résout :

$$t^* = \max \left\{ \frac{\mathbf{b}^\top \mathbf{u}}{\mathbf{B}\mathbf{r}^\top \mathbf{u}} : \mathbf{B}\mathbf{r}^\top \mathbf{u} > 0, \mathbf{u} \in \mathcal{P} \right\}$$


Charnes Cooper :

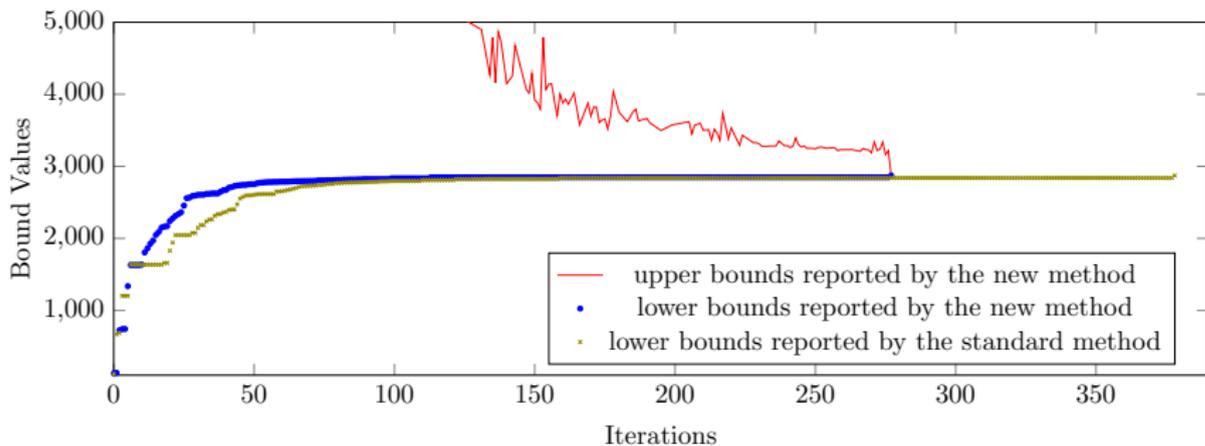
$$\mathbf{v} = \mathbf{u} \frac{1}{\mathbf{B}\mathbf{r}^\top \mathbf{u}}$$

\Rightarrow

$$\begin{aligned} t^* &= \max \mathbf{b}^\top \mathbf{v} \\ \mathbf{A}^\top \mathbf{v} &\leq \mathbf{0} \\ (\mathbf{B}\mathbf{r})^\top \mathbf{v} &= 1 \\ \mathbf{v} &\geq \mathbf{0} \end{aligned}$$

- 1 La séparation et le concept d'intersection
- 2 La décomposition de Benders et son sous-problème d'intersection
- 3 Résultats et conclusions**

Quelques résultats 1



taille vecteur \mathbf{y} : $|E| = 600$

Quelques résultats 2

The speed-up obtained by the new method becomes increasingly higher as the graph size grows, reaching speed-ups close to 10, when $|E| = 5000$. More detailed results on these new runs are publicly available on-line at cedric.cnam.fr/~porumbed/benders/, along with the code source.

| | $ E $ 1500 | $ V $ 130 | $ E $ 2000 | $ V $ 150 | $ E $ 2500 | $ V $ 175 | $ E $ 3000 | $ V $ 200 | $ E $ 3500 | $ V $ 225 | $ E $ 5000 | $ V $ 275 |
|--------------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|--------------|
| CPU Time speed-up | 2.27 | | 3.27 | | 3.78 | | 4.49 | | 5.65 | | 9.21 | |
| CPU Time new meth. | 536 | | 1533 | | 1886 | | 2912 | | 4164 | | 8636 | |
| CPU Time old meth. | 1220 | | 5013 | | 7124 | | 13063 | | 23521 | | 79564 | |
| Iters new meth. | 162 | | 184 | | 221 | | 257 | | 280 | | 340 | |
| Iters old meth. | 532 | | 804 | | 961 | | 1179 | | 1518 | | 1983 | |

Table 3: Speed-up of the new method compared to the standard method on larger **random-10** graphs (all with $id = 1$). The figures for the first two graphs are taken from Table 2 and they represent averages over 20 runs. For the remaining larger graphs, we only used 5 runs because the running times are very large.

On a aussi fait des calculs avec un Branch-and-bound à base d'une relaxation linéaire des variables y : gap/écart à la racine de 0.5% à 10%

Conclusions

Il est possible de remplacer le sous-problème de séparation avec celui d'intersection avec des avantages :

- 1 contraintes plus fortes
- 2 une borne sup disponible à chaque itération (impossible avec la séparation standard)
 - Très utile si on résout une relaxation linéaire des \mathbf{y} et on a des bornes **inf-sup** pour exécuter un Branch-and-bound

Il y a une étape difficile : au début les rayons ne touchent pas la zone réalisable \mathcal{Z} , on peut toujours se servir de la séparation !

Dans un contexte Benders :

- la séparation nécessite de résoudre un PL sur \mathcal{P}
- l'intersection nécessite plusieurs PLs sur le même \mathcal{P}