Quelques Propriétés Particulières de la Relaxation Lagrangienne.

Monique Guignard
the Wharton School
University of Pennsylvania

Journée Francilienne de Recherche Opérationnelle, 11 juin 2004

Un programme en nombres entiers

(P): $\mathbf{Min}\ \ fx$

$$\text{s.c. } Ax \leq b$$

$$\boxed{Cx \leq d}$$

$$x \in X$$

avec des contraintes $Cx \leq d$ qui compliquent le problème:

Min fx

s.t. Ax £ **b**

x Î **X**

On dualise les contraintes $Cx \leq d$:

Min fx

s.t. Ax £ b

Cx £ d

x Î X

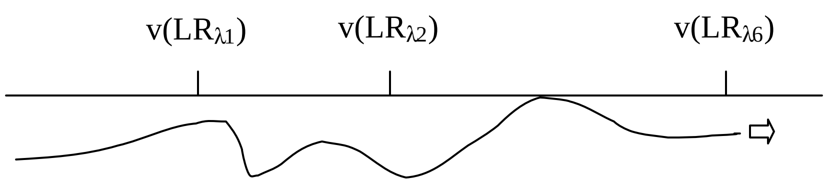
i.e., on les intègre à la fonction économique:

(LRl) **Min fx - l(d-Cx)**

s.c. Ax £ b

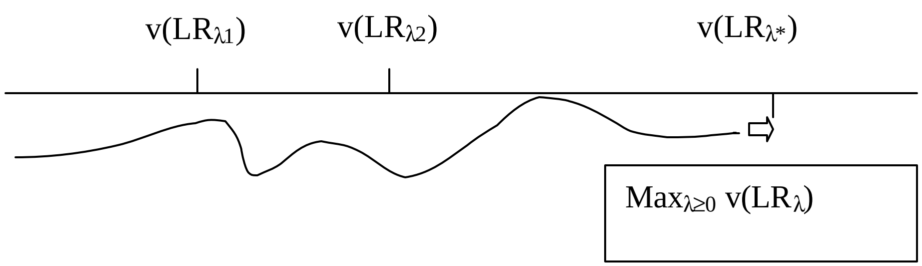
x Î X

ce qui fournit une borne inférieure pour v(P):



Le problème dual lagrangien consiste à chercher la meilleure borne inférieure. Il s'écrit

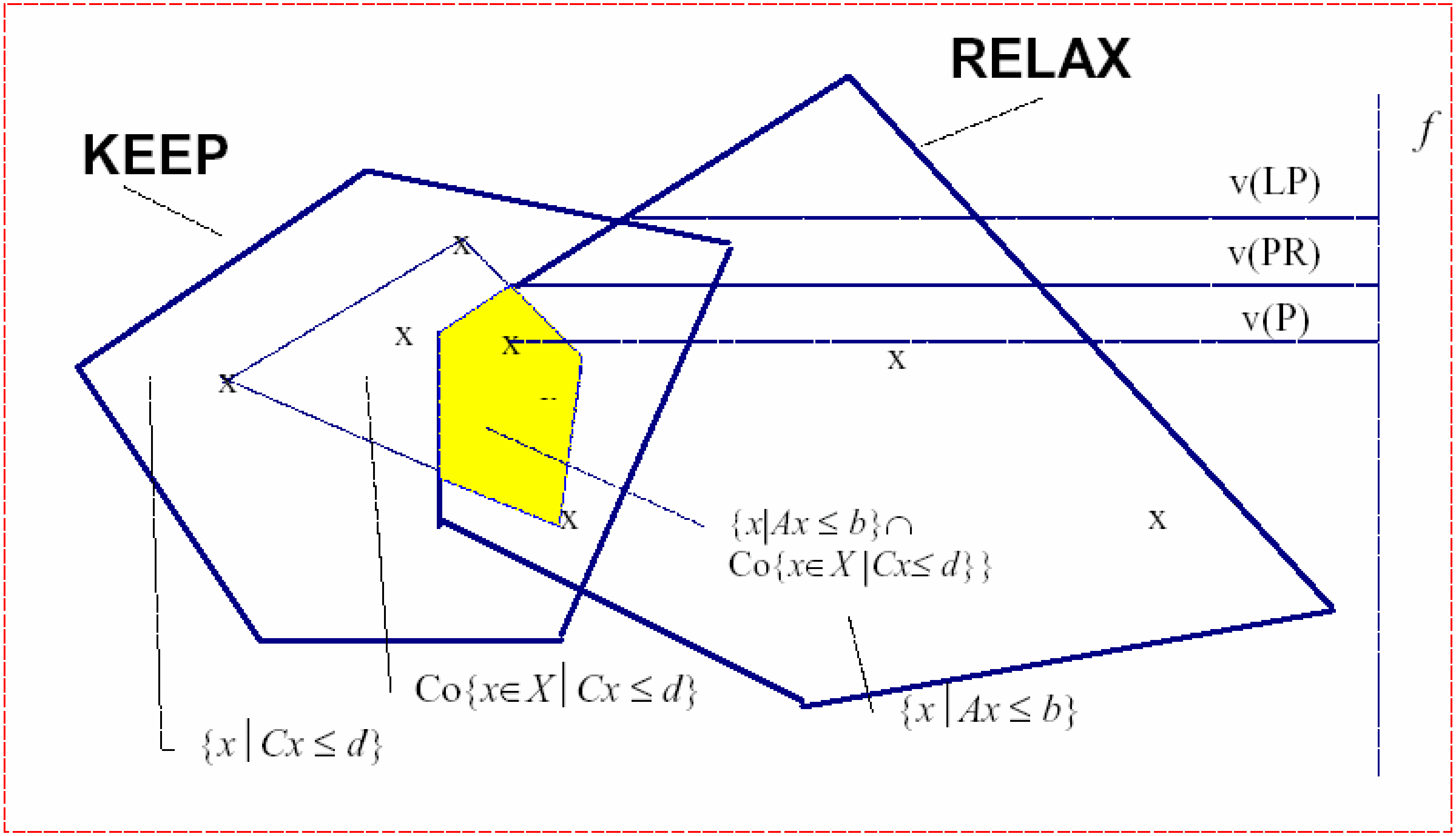
$$(LR) \qquad \mathrm{Max}_{\lambda \geq 0}\ v(LR_\lambda)$$



Interprétation géométrique de la relaxation lagrangienne (Geoffrion 1974):

Il existe un problème primal, (PR), équivalent à (LR), i.e., tel que

$$v(LR) = v(PR)$$



RELAX
KEEP
f
v(LP)
v(PR)
v(P)
x
x
x
x
x
x
x
$\{x | Ax \le b\} \cap$
$\mathrm{Co}\{x \in X | Cx \le d\}\}$
$\mathrm{Co}\{x \in X | Cx \le d\}$
$\{x | Ax \le b\}$
$\{x | Cx \le d\}$

Conséquences de l'interprétation géométrique :

v(LP) £ v(LR) =v(PR) £ v(P)

Propriété d'Intégralité:

si l'enveloppe convexe des points entiers des contraintes que l'on garde est égale au polyèdre des contraintes que l'on garde

autrement dit, si

$$Co\{x \in X \mid Cx \leq d\} = \{x \in R^n \mid Cx \leq d\}$$

Conséquences

(malheureuses !)

de la propriété d'intégralité:

v(LR)=v(PR)=v(LP)

Quelques Rappels

Génération de coupes (ou plans sécants, de Kelley)

(LR) $\quad \text{Max}_{\lambda \geq 0}\, v(LR_\lambda)$

$= \quad \text{Max}_{\lambda \geq 0}\ \text{Min}_{k=1,\ldots,K} \quad \{f\ x^k + \lambda(b\text{-}Ax^k\ )\}$

$= \quad \text{Max}_{\lambda\geq 0,\ \eta}\ \{\eta \mid \eta \leq f\ x^k + \lambda(b\text{-}Ax^k\ ),\ k=1,\ldots,K\}.$

A chaque itération k, on engendre une coupe (ou contrainte) de

la forme $\quad \eta \leq f\ x^{(k)} + \lambda(b-Ax^{(k)}),$

en résolvant le sous-problème lagrangien ($LR_\lambda{}^k$) de solution $x^{(k)}$.

Ces coupes sont ajoutées à celles engendrées lors des itérations précédentes pour former le problème-maître

$$(MP^k) \qquad \text{Max}_{\lambda\geq 0,\, \eta} \{\eta \mid \eta \leq f x^{(h)} + \lambda(b-Ax^{(h)}),\ h = 1,\dots, k\},$$

dont la solution fournit l'itéré suivant λ^{k+1}. En général, on a

$$v(LR_\lambda{}^k) \leq v(LR) \leq v(MP^k) \ .$$

Le processus se termine lorsque $v(MP^k) = v(LR_\lambda{}^k)$. Cette valeur est la valeur optimale du dual lagrangien (LR) :

$$v(LR) = v(MP^k) = v(LR_\lambda{}^k) \ .$$

Décomposition Lagrangienne

Un programme en nombres entiers

(P): $\quad \textbf{Min } fx$

$$\text{s.c. } Ax \le b$$

$$\boxed{Cx \le d}$$

$$x \in X$$

on crée des doubles des variables:

Min fx

s.t. Ax £ b

Cy £ d

x Î X

y Î X

x = y

et on dualise les contraintes de copie:

$$
\begin{array}{lll}
\textbf{Min} & \mathbf{fx} - \mu\mathbf{(x-y)} & \\
\textbf{s.t.} & \mathbf{Ax} \le \mathbf{b} & \\
 & \mathbf{Cy} \le \mathbf{d} & \\
 & \mathbf{x} \in \mathbf{X} & \\
 & \mathbf{y} \in \mathbf{X} &
\end{array}
$$

on obtient un problème qui se décompose en deux sous-problèmes, l’un en x, l’autre en y:

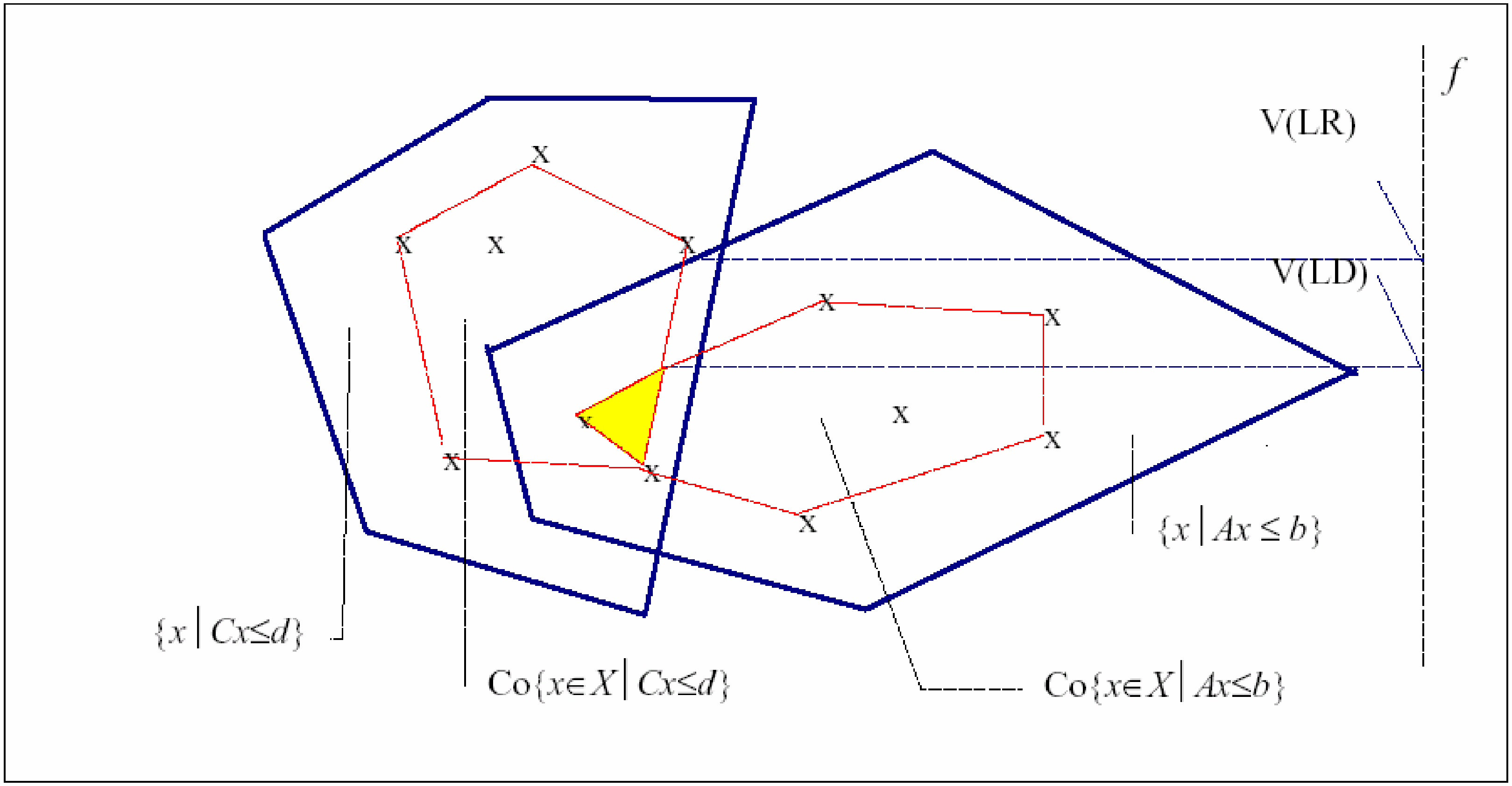
Min fx - m**x**	**+**	**Min**	**m****y**
s.t. Ax £ **b**		**s.t.**	**Cy** £ **d**
x Î **X**			**y** Î **X**

Première différence avec la relaxation classique: on garde toutes les contraintes.

Interprétation Géométrique de la Décomposition Lagrangienne (Guignard, Kim 1987):

Il existe un problème primal, (PD), équivalent à (LD), i.e.,

v(LD) = v(PD)



f
V(LR)
V(LD)
$\{x \mid Ax \le b\}$
$\{x \mid Cx \le d\}$
$\mathrm{Co}\{x \in X \mid Cx \le d\}$
$\mathrm{Co}\{x \in X \mid Ax \le b\}$

Conséquences de la <u>propriété d' intégralité (PI)</u>:

Si l'un des sous-problèmes a la PI, alors

$$v(LD)=v(PR)=v(LR_2)$$

Si les deux sous-problèmes ont la PI, alors

$$v(LD)=v(PD)=v(LR_1)=v(LR_2)=v(LP)$$

Substitution Lagrangienne (Maculan et Reinoso 1988)

Un programme en nombres entiers

(P): $\mathbf{Min}\ \ fx$

$$\text{s.c. } Ax \leq b$$

$$\boxed{Cx \leq d}$$

$$x \in X$$

on substitue, par exemple:

$$
\begin{aligned}
\textbf{Min} \quad & \mathbf{fx} \\
\textbf{s.t.} \quad & \mathbf{Ax} \le \mathbf{b} \\
& \boxed{\mathbf{Dy} \le \mathbf{d}} \\
& \mathbf{x} \in \mathbf{X} \\
& \mathbf{y} \in \mathbf{X'} \\
& \mathbf{Cx} = \mathbf{Dy}
\end{aligned}
$$

et on dualise la contrainte de copie:

$$
\begin{array}{lll}
\text{Min} & fx - \mu(Cx - Dy) & \\
\text{s.t.} & Ax \le b & \\
 & Dy \le d & \\
 & x \in X & \\
 & y \in X' &
\end{array}
$$

Quelques questions:

(1) Est-il possible d'obtenir une borne “forte” (i.e., une borne meilleure que la borne v(LP)) même si les problèmes lagrangiens sont simples à résoudre?

(2) Est-il possible de décomposer les plans sécants lorsque le problème lagrangien se décompose en plusieurs sous-problèmes?

(3) Est-il possible d'obtenir une borne aussi forte que par la décomposition lagrangienne en dualisant des copies agrégées plutôt que des copies individuelles des variables?

Aussi surprenant que cela puisse paraître, dans les trois cas, la réponse est

OUI!

Situations illustrant les trois cas.

Cas 1:

(1) Il est possible d'obtenir une borne "forte" (i.e., une borne meilleure que la borne v(LP)) même si les problèmes lagrangiens sont simples à résoudre.

Quand?

Quand le problème lagrangien a la propriété de linéarisation entière.

(LR_λ) Max $\{ fx + gy \mid A_i x_i \le p_i y_i, \forall i, x \in X, By \le b, y_i = 0 \text{ or } 1, \forall i\}$

separées, une par i *seulement sur y*

où $X = \Pi_i (X_i)$ peut contenir des contraintes d'intégralité sur les variables.

x_i peut être un vecteur.

(i) ignorer initialement les contraintes sur les variables entières, $By \le b$**.**

(LR_λ) Max $\{ fx + gy \mid A_i x_i \leq p_i y_i, \forall i, x \in X, \quad y_i = 0 \text{ or } 1, \forall i\}$

separées, une par *seulement sur y*

où $X = \Pi_i (X_i)$ peut contenir des contraintes d'intégralité sur les variables.

x_i peut être un vecteur.

(ii) le problème se décompose en un problème pour chaque i:

(LR^i_λ) Max $\{ f_i x_i + g_i y_i \mid A_i x_i \leq p_i y_i, x_i \in X_i, y_i = 0 \text{ or } 1, \forall i\}$

où y_i joue le rôle d'un paramètre 0-1:

pour $y_i = 0$, $x_i = 0$, et $f_i x_i + g_i y_i = 0$.

pour $y_i = 1$, résoudre $(LR^i_\lambda \mid y_i = 1)$:

$$v_i = \text{Max}\{ f_i x_i + g_i \mid A_i x_i \leq p_i, x_i \geq 0\}.$$

v_i est la **contribution** de $y_i = 1$ dans la f. éco.

(LR_λ) $\quad \text{Max } \{ fx + gy \mid A_i x_i \le p_i y_i, \forall i, x \in X, \; By \le b, \; y_i = 0 \text{ or } 1, \forall i\}$

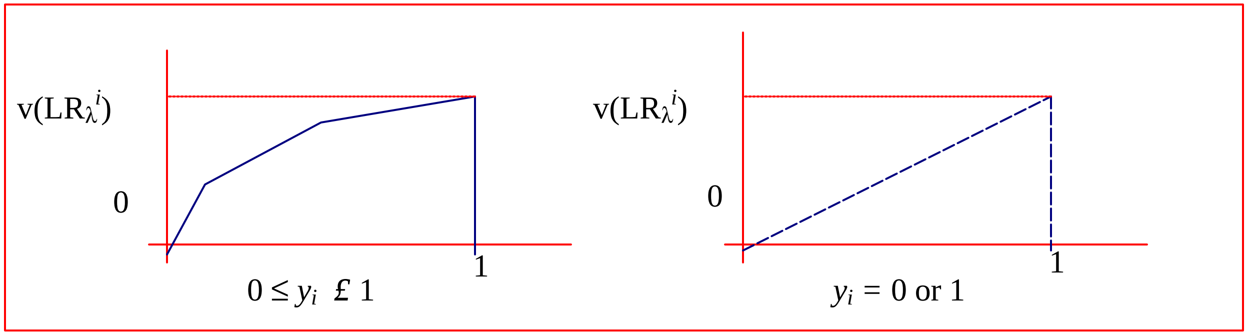
separées, une par i *seulement sur y*

où $X = \Pi_i (X_i)$ peut contenir des contraintes d'intégralité sur les variables.

x_i peut être un vecteur.

(iii) **remplacer** $v(LR_\lambda)$ par $v(PL_\lambda)$ où (PL_λ) est

$(PL_\lambda) \qquad \text{Max } \{\sum_i v_i y_i \mid By \le b, \; y_i = 0 \text{ ou } 1, \forall i\}.$



Utilise la condition d'intégralité sur y_i:

$y_i \in \{0,1\}$.

Exemple: PRODOPT (de Matta, Guignard, 1994)

Min $\quad \sum_{j,i,t} (\sum_s c_{ijts} + q_{ij} z_{ijt})$

s.c. $\quad \boxed{\sum_{j,t} x_{ijts} = 1,} \quad \forall i,s$

$\sum_s d_{is} x_{ijts} = p_{ij} y_{ijt}, \ \forall i,j,t, \qquad x_{ijts} \le y_{ijt}, \ \forall i,j,t,s$

$\sum_i y_{ijt} = 1, \qquad \forall j,t$

$z_{ijt} \ge y_{ijt} - y_{ij,t-1}, \quad \forall i,j,t$

$0 \le x_{ijts} \le 1, \ y_{ijt} \in \{0,1\}, \ z_{ijt} \ge 0$

(1) (LR_λ) se décompose en un problème par machine j.

(2) exactement un y_{ijt} est égal à 1, pour tout j et t:

en période t, la machine j passe à exactement un produit i. Donc si $y_{ijt}=1$, on peut trouver le x_{ijts} correspondant en résolvant un problème de pseudo sac-à-dos continu avec comme contrainte:

$$\sum_s d_{is} x_{ijts} = p_{ij} \quad \Rightarrow \text{ valeur } v_{ijt} \text{ pour } y_{ijt}=1.$$

Sous-problème pour la machine j:

<u>problème de plus court chemin acyclique</u> (recherche d'une suite de changements, de période en période, d'un produit à un autre, ou au même produit).

Un noeud correspond à l'affectation de la machine j en période t à la production du produit i.

Un arc du noeud (l,j,t–1) au noeud (i,j,t) a un coût $v_{ijt} + q_{lj}$ (ou v_{ijt} si l=i).

Cas 2:

(4) Il est possible de décomposer les plans sécants lorsque le problème lagrangien se décompose en plusieurs sous-problèmes.

Quand?

toujours!

$$(MP^k) \qquad \text{Max}_{\lambda\geq 0,\, \eta} \{\eta \mid \eta \leq f\, x^{(h)} + \lambda(b - Ax^{(h)}),\ h=1,\ldots,k\}$$

$$= \qquad \text{Max}_{\lambda\geq 0,\, z} \{z - \lambda b \mid z \leq (f - \lambda A)\, x^{(h)},\ h=1,\ldots,k\},$$

est remplacé par

$$(MPD^k) \qquad \text{Max}_{\lambda\geq 0,\, zl} \{\textstyle\sum_l z_l - \lambda b \mid z_l \leq (f - \lambda A)^l\, x_l^h,\ h=1,\ldots,k\}.$$

où l est le numéro du sous-problème, et x_l^h est une solution lagrangienne du $l^{\text{ème}}$ sous-problème à l'itération h.

Ceci permet de mélanger des solutions partielles trouvées à des itérations différentes : en anglais, « mix-and-match ». Peut s'utiliser aussi dans une méthode de faisceaux.

Cas 3:

(5) Il est possible d'obtenir une borne aussi forte que par la décomposition lagrangienne en dualisant des copies agrégées plutôt que des copies individuelles des variables.

Quand?

En particulier quand le problème lagrangien a une structure similaire à celle d'un problème d'implantation avec capacités.

Considérons le problème général suivant:

Min		$f(x,y,s,t)$
s.t.	(D)	$(x,y,s) \in S_{x,y,s}$
	(T)	$(y,t) \in S_{y,t}$
	(C)	$d_{Nj}x_{Nj} \le a_j y_j \ \forall j$
	(N)	$0 \le x \le e$
	(I)	$y_j \in \{0,1\}$

$$\left.\begin{array}{ll} (C) & d_{Nj}x_{Nj} \le a_j y_j \ \forall j \\ (N) & 0 \le x \le e \\ (I) & y_j \in \{0,1\} \end{array}\right\} \quad (B) \quad x_{Nj} \le y_j e \ \forall j$$

LD:

Min $f(x,y,s,t)$

s.t.

(D)	$(x,y,s) \in S_{x,y,s}$
(B)	$x_{Nj} \le y_j e \ \forall j$
(N)	$0 \le x \le e$
(I)	$y_j \in \{0,1\}$

(N)	$0 \le x' \le e$
(I)	$y'_j \in \{0,1\}$
(T)	$(y',t') \in S_{y,t}$
(C)	$d_{Nj} x'_{Nj} \le a_j y'_j \ \forall j$

$x = x'$ and $y = y'$

LS:

Min $f(x,y,s,t)$

s.t.

(D)	$(x,y,s) \in S_{x,y,s}$
(B)	$x_{Nj} \leq y_j e \ \forall j$
(N)	$0 \leq x \leq e$
(I)	$y_j \in \{0,1\}$

(N)	$0 \leq x' \leq e$
(I)	$y'_j \in \{0,1\}$
(T)	$(y',t') \in S_{y,t}$
(C)	$d_{Nj} x'_{Nj} \leq a_j y'_j \ \forall j$

$d_{Nj} x_{Nj} = d_{Nj} x'_{Nj}$ and $y = y'$

Résultat:

Les deux bornes lagrangiennes sont égales !

(Chen, Guignard, DAM, 1998)

Première application: CPLP

Min $\Sigma_{i,j}\ c_{ij}x_{ij} + \Sigma_i\ f_i y_i$

s.c.

(D)	$\Sigma_j\ x_{ij} = 1,$	$\forall\ i$		
(C)	$\Sigma_i\ d_i x_{ij} \leq a_j y_j,$	$\forall\ j$	(KP)	
(B)	$x_{ij} \leq y_j,$	$\forall i,j$		
(N)	$0 \leq x_{ij} \leq 1,$	$\forall\ i,j$		
(I)	$y_j \in \{0,1\},$	$\forall j$		(APLP)
(T)	$\Sigma_j\ a_j y_j\ \geq \Sigma_i\ d_i$			

sous-problèmes identiques dans les deux cas,

cependant

pour (LD), il y a $|i| \times |j| + |j|$ multiplicateurs

pour (LS), il n'y a que $2|j|$ multiplicateurs

Seconde application: PRODOPT

Produit i

Machine j

Production en période t

Demande en période s

x_{ijts} = % de la demande d_{is} de i en période s qui est produit en période t sur la machine j

y_{ijt} = 1 si on produit i sur la machine j en période t

z_{ijt} = 1 si on commence à produire le produit i sur la machine j en début de période t

Min $\sum_{j,i,t} (\sum_s c_{ijts} + q_{ij} z_{ijt})$

s.c.

$$\boxed{\begin{array}{l} \sum_s d_{is} x_{ijts} = p_{ij} y_{ijt} \\ \\ \sum_i y_{ijt} = 1 \\ \\ z_{ijt} \geq y_{ijt} - y_{ij,t-1} \end{array}} \quad \Rightarrow \quad \boxed{\begin{array}{l} \sum_{j,t} x_{ijts} = 1 \\ \\ x_{ijts} \leq y_{ijt} \end{array}} \quad (SPLP_i)$$

(ShP_j)

$$0 \leq x_{ijts} \leq 1, \; y_{ijt} \in \{0,1\}, \; z_{ijt} \geq 0$$

Conclusion

Lors de la construction d’une relaxation lagrangienne, on devrait se poser les questions suivantes:

(1) Est-ce que le problème a la propriété de linearisation entière? Dans ce cas il se peut que le problème lagrangien soit très facile à résoudre (même s'il n'a pas la PI).

(2) Est-ce que le problème lagrangien est décomposable? en ce cas on peut découper et mélanger les plans sécants dans une méthode de génération de coupes, de colonnes ou de faisceaux.

(3) Peut-on utiliser une substitution lagrangienne (ou une décomposition lagrangienne agrégée)? Il se peut que la borne soit (presque) aussi bonne (et beaucoup moins coûteuse) que par la décomposition lagrangienne.

~~~~~~~
~~~~~~~