

ANGEL CORBERÁN

MÉTODOS DE GENERACIÓN DE COLUMNAS

UNIVERSITAT DE VALÈNCIA

MÉTODOS DE GENERACIÓN DE COLUMNAS

Introducción

Una de las ideas clave en optimización es la de descomposición. Por ejemplo, consideremos el PLE

$$\max_{x \in X} cx$$

donde X es una región convexa que puede escribirse como la intersección de 2 o más conjuntos con estructura $X = \bigcap_{k=1}^K X^k$, por algún $K \geq 1$. Este caso aparece con frecuencia en la literatura. Más particular es incluso el caso en que las restricciones tienen la forma:

$$\begin{aligned} A^1 x^1 + A^2 x^2 + \dots + A^K x^K &= b \\ D^1 x^1 &\leq d_1 \\ D^2 x^2 &\leq d_2 \\ &\vdots \\ D^K x^K &\leq d_K \\ x^1 \in \mathbb{Z}_+^{n_1}, \dots, x^K \in \mathbb{Z}_+^{n_K} \end{aligned}$$

de forma que los conjuntos $X^k = \{x^k \in \mathbb{Z}_+^{n_k} : D^k x^k \leq d^k\}$ son independientes por $k=1, 2, \dots, K$ y únicamente las restricciones "conjuntas" $\sum_{k=1}^K A^k x^k = b$ relacionan de forma conjunta los diferentes conjuntos de variables.

Dada una función objetivo $\max \sum_{k=1}^K c^k x^k$, dos aproximaciones que nos permiten beneficiarnos de una estructura así son:

- generación de cortes, en la que intentaríamos generar desigualdades válidas para cada subconjunto X^k , $k=1 \dots K$, y
- relajación lagrangiana, en la que dualizaríamos las restricciones "conjuntas" (o globales) para obtener un problema que podríamos particionar en

K subproblemas distributos a resolver.

En este tema introduciremos una 3ª forma de explotar la estructura de PLE's de la forma anterior.

- Supondremos a lo que rije que cada uno de los X^k , $k=1 \dots K$, es acotado.
- Los problemas que deseamos resolver aquí son PLE's con un enorme n° de variables, donde las columnas están, a menudo, descritas de forma implícita como los vectores de incidencia de ciertos subconjuntos de un conjunto (trucks, subconjuntos de clientes, etc).
- Primeros vemos cómo tales (enormes) formulaciones de un PLE, llamados "Problemas Maestros (Master)", aparecen. A continuación estudiaremos cómo resolver la relajación lineal de esos problemas Maestros y relacionaremos lo potente de esa relajación con la obtenida por relajación lagrangiana o con el uso de planos de corte. Finalmente, estudiaremos qué hacer cuando la solución lineal no es entera y debemos realizar "enumeración", lo que nos lleva a la generación de columnas en PLE o algoritmos llamados de Branch-and-Price.

Reformulación Dantzig-Wolfe de un PLE

Consideremos el problema en la forma:

$$\begin{aligned}
 (\text{PLE}) \quad z = \max \quad & \sum_{k=1}^K c^k x^k \\
 & \sum_{k=1}^K A^k x^k = b \\
 & x^k \in X^k, \quad k=1 \dots K
 \end{aligned}
 \quad \left. \vphantom{\sum_{k=1}^K} \right\}$$

donde $X^k = \{x^k \in \mathbb{Z}_+^{n_k} : D^k x^k \leq d_k\}$ para $k=1 \dots K$.

Como hemos supuesto que los X^k son acotados, esto nos permite asumir que cada X^k tiene un enorme, pero finito, conjunto de puntos $\{x^{k,t}\}_{t=1}^{T_k}$.

Así, tenemos:

$$X^k = \{x^k \in \mathbb{R}^{n_k} : x^k = \sum_{t=1}^{T_k} \lambda_{k,t} x^{k,t}, \sum_{t=1}^{T_k} \lambda_{k,t} = 1, \lambda_{k,t} = 0,1 \quad \forall t=1, \dots, T_k\}$$

Ahora, reemplazando x^k , llegamos a un equivalente Problema Master PLE

$$\begin{aligned} (\text{PLEMaster}) \quad z = \max \quad & \sum_{k=1}^K \sum_{t=1}^{T_k} (c^k x^{k,t}) \lambda_{k,t} \\ & \sum_{k=1}^K \sum_{t=1}^{T_k} (A^k x^{k,t}) \lambda_{k,t} = b \\ & \sum_{t=1}^{T_k} \lambda_{k,t} = 1 \quad \forall k=1, \dots, K \\ & \lambda_{k,t} = 0,1 \quad \forall t=1, \dots, T_k \text{ y } \forall k=1, \dots, K \end{aligned}$$

Resolución del PL Master

Aquí usaremos un algoritmo de generación de columnas para resolver la Rel. Lineal del PLEMaster, llamada PLMaster:

$$\begin{aligned} (\text{PLMaster}) \quad z^{\text{PLM}} = \max \quad & \sum_k \sum_t (c^k x^{k,t}) \lambda_{k,t} \\ & \sum_k \sum_t (A^k x^{k,t}) \lambda_{k,t} = b \\ & \sum_t \lambda_{k,t} = 1 \quad k=1, \dots, K \\ & \lambda_{k,t} \geq 0 \quad t=1, \dots, T_k \quad k=1, \dots, K \end{aligned}$$

donde hay una columna $\begin{pmatrix} c^k x \\ A^k x \\ e_k \end{pmatrix}$ por cada $x \in X^k$.

Luego usaremos v.d.duals $\{\pi_i\}_{i=1}^m$ asociadas a las restricciones globales y las

v.d.als $\{\mu_k\}_{k=1}^K$ asociados a las restricciones de convexidad.

- La idea es resolver el (PLM) con el algoritmo (primal) del simplex. Sin embargo, la etapa de calcular el coste reducido (pricing step) de una columna para que entre en la base debe ser modificada a causa del enorme número de variables (columnas). En lugar de "pricing" las columnas una a una, el problema de encontrar la columna con el mayor ~~coste~~ reducido (menor $\pi - c_j$) es el mismo en conjunto de K problemas de optimización.

Inicialización: Supongamos que conocemos un subconjunto de columnas (al menos 1 por cada k). Así tenemos un Problema Lineal Marko Restringido

$$(PLMR) \quad \left. \begin{aligned} \tilde{z}^{PLM} &= \max \tilde{c} \tilde{z} \\ \tilde{A} \tilde{z} &= \tilde{b} \\ \tilde{z} &\geq 0 \end{aligned} \right\}$$

donde $\tilde{b} = \begin{pmatrix} b \\ 1 \end{pmatrix}$, $\tilde{A}$ está formado por el conjunto de columnas disponible y es una abstracción de

$$\begin{pmatrix} A^1 x^{1,1} & \dots & A^1 x^{1,T_1} & A^2 x^{2,1} & \dots & A^2 x^{2,T_2} & \dots & A^K x^{K,1} & \dots & A^K x^{K,T_K} \\ 1 & \dots & 1 & & & & & & & \\ & & & 1 & \dots & 1 & & & & \\ & & & & & & \dots & 1 & \dots & 1 \end{pmatrix}$$

y $\tilde{c}, \tilde{z}$ son los correspondientes coeficientes y variables.

La resolución de (PLMR) proporciona una solución óptima primal $\tilde{z}^*$ y una d. óptima dual $(\pi, \mu) \in \mathbb{R}^m \times \mathbb{R}^K$.

Propiedad Primal : Cualquier solución primal de (PLMR) lo es del (PLM).

En particular $\tilde{z}^*$ es una sol. primal de (PLM) y, por lo tanto, tenemos:

$$\tilde{z}^{PLM} = \tilde{c} \tilde{h}^* = \sum_{i=1}^m \pi_i b_i + \sum_{k=1}^K \mu_k \leq z^{PLM}$$

Test de optimalidad para (PLM) : Necesitamos chequear si (π, μ) es dual primal para (PLM). Esto comprende el chequear cada columna, es decir ^{por} cada k y para cada $x \in X^k$ si el coste reducido $c^k x - \pi A^k x - \mu_k \leq 0$. $\longrightarrow$

En lugar de examinar cada punto separadamente, tomaremos ~~de~~ examinando cada punto todos los puntos en X^k implícitamente resolviendo el siguiente subproblema:

$$z_k = \max_{x \in X^k} (c^k - \pi A^k) x - \mu_k$$

Criterio de parada :

Si $z_k = 0 \quad \forall k=1 \dots K$, la solución (π, μ) es dual primal para (PLM) y, por lo tanto,

$$z^{PLM} \leq \sum_{i=1}^m \pi_i b_i + \sum_{k=1}^K \mu_k.$$

Como el valor de la solución primal primal $\tilde{z}$ iguala la de este coste superior $\rightarrow \tilde{z}$ es óptima de (PLM).

Generación de una nueva columna :

Si $z_k > 0$ para algún k , la columna correspondiente a la solución óptima $\tilde{x}^k$ del subproblema tiene coste reducido positivo.

Introduciremos la columna $\begin{pmatrix} c^k & \tilde{x}^k \\ A^k & \tilde{x}^k \\ e_k \end{pmatrix}$

para obtener un nuevo (PLMR) que puede ser reoptimizado fácilmente. (Usando el algoritmo primal del simplex).

Una Cota Dual (Cota Superior) :

Del subproblema tenemos que

$$z_k \geq (c^k - \pi A^k)x - \mu_k \quad \forall x \in X^k$$

Por lo tanto, $(c^k - \pi A^k)x - \mu_k - z_k \leq 0 \quad \forall x \in X^k.$

Haciendo $z = (z_1, \dots, z_K)$, tenemos que $(\pi, \mu + z)$ es dual viable $\rightarrow$ para el (PLM). Por lo tanto :

$$\tilde{z}^{\text{PLM}} \leq z^{\text{PLM}} \leq \pi b + \underbrace{\sum_{k=1}^K \mu_k}_{\tilde{z}^{\text{PLM}} \text{ (el restringido)}} + \sum_{k=1}^K z_k$$

Estas consideraciones conducen directamente a un algoritmo para el (PLM) que termina cuando $z_k = 0 \quad \forall k=1-K$. Sin embargo, como en la Rel. Lagrangiana, es posible terminar antes.

Un Criterio de Parada Alternativo :

Si las soluciones a los subproblemas $(\tilde{x}^1, \dots, \tilde{x}^K)$ satisfacen las restricciones globales (compartes) originales $\sum_{k=1}^K A^k \tilde{x}^k = b$, entonces $(\tilde{x}^1, \dots, \tilde{x}^K)$ es óptima.

La razón es por que $z_k = (c^k - \pi A^k)\tilde{x}^k - \mu_k \rightarrow$

$$\sum_k c^k \tilde{x}^k = \sum_k \pi A^k \tilde{x}^k + \sum_k \mu_k + \sum_k z_k = \pi b + \sum_k \mu_k + \sum_k z_k$$

Con lo que la solución primal viable tiene el mismo valor que la cota superior (anterior) a z^{PLM} .

Ejemplo : El TSP Simétrico resuelto por generación de columnas.

Consideramos aquí la aplicación del algoritmo anterior a la resolución del PL Master de un problema que contiene un único subproblema.

Sea el TSP Simétrico, que puede plantearse como:

$$\left. \begin{array}{l} \min \sum_{e \in E} c_e x_e \\ \sum_{e \in \delta(i)} x_e = 2 \quad i \in V \\ x \in X^1 \end{array} \right\}$$

$$\text{donde } X^1 = \left\{ x \in \mathbb{B}_+^m : \sum_{e \in \delta(1)} x_e = 2, \sum_{e \in E(s)} x_e \leq |S|-1, \forall S \subset V \setminus \{1\}, \right. \\ \left. \sum_{e \in E} x_e = n \right\}$$

es el conjunto de vectores de incidencia de los 1-árboles.

(Recordemos que un 1-árbol es un árbol sobre el grafo inducido por los vértices $\{2, \dots, n\}$ mas 2 aristas incidentes con el vértice 1. Un tour del TSP es un 1-árbol que tiene 2 aristas incidentes con cualquier nodo)

Representaremos por $G^t = (V, E^t)$ el t -ésimo 1-árbol y supondremos que hay T_1 1-árboles. Como aquí $K=1$, podemos escribir:

$$X^1 = \left\{ x \in \mathbb{R}_+^m : x = \sum_{t=1}^{T_1} \lambda_t x^t, \sum_{t=1}^{T_1} \lambda_t = 1, \lambda_t = 0,1 \quad \forall t=1, \dots, T_1 \right\}$$

donde x^t son los vectores de incidencia de los 1-árboles.

Entonces, podemos escribir

$$x_e = \sum_{t=1}^{T_1} \lambda_t x_e^t, \text{ donde } x_e^t = \begin{cases} 1 & \text{si } e \in E^t \\ 0 & \text{en otro caso} \end{cases}$$

$$\text{con lo que } x_e = \sum_{t: e \in E^t} \lambda_t$$

Por lo tanto, las restricciones sobre el grado de los vértices se pueden escribir como:

$$\sum_{e \in \delta(i)} x_e = \sum_{e \in \delta(i)} \sum_{t: e \in E^t} h_t = \sum_t d_i^t \cdot h_t = 2 \quad \forall i \in V$$

donde d_i^t es el grado del vértice i en el t -árbol G^t .

Entonces, el (PLE Master) es:

$$\begin{aligned} \text{(PLEM)} \quad & \min \sum_{t=1}^{T_1} (c x^t) h_t \\ & \left. \begin{aligned} \sum_{t=1}^{T_1} d_i^t h_t &= 2 \quad \forall i \in V \\ \sum_{t=1}^{T_1} h_t &= 1 \\ h_t &= 0,1 \quad \forall t=1 \dots T_1 \end{aligned} \right\} \end{aligned}$$

y su correspondiente relajación lineal:

$$\begin{aligned} \text{(PLM)} \quad & \min \sum_{t=1}^{T_1} (c x^t) h_t \\ & \left. \begin{aligned} \sum_{t=1}^{T_1} d_i^t h_t &= 2 \quad \forall i \in V \\ \sum_{t=1}^{T_1} h_t &= 1 \\ h_t &\in R_+^{T_1} \end{aligned} \right\} \end{aligned}$$

Asociadas a las restricciones de grado tenemos las variables duales $\{u_i\}_{i=1}^w$ y la variable dual μ asociada a la restricción de convexidad.

El (único) subproblema asociado es:

$$z_1 = \min_{x \in X^1} \left\{ \sum_{e \in E} (c_e - u_i - u_j) x_e - \mu \right\}$$

ya que el 1-árbol G^t tiene coste reducido

$$cx^t - \sum_{i \in V} d_i^t u_i - \mu \quad (= \text{coeficiente de } h_t - \text{v.duales} \times \text{columna de } h_t)$$

y tenemos

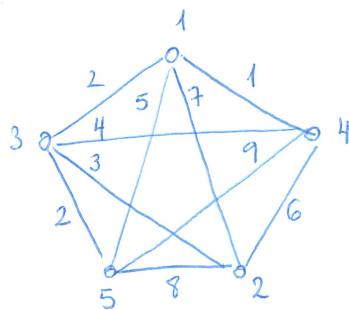
$$\begin{aligned} cx^t - \sum_{i \in V} d_i^t u_i - \mu &= cx^t - \sum_{i \in V} u_i \sum_{e \in \delta(i)} x_e^t - \mu = \\ &= \sum_{e \in E} (c_e - u_i - u_j) x_e^t - \mu \end{aligned}$$

donde x_e^t por $e \in E$ son las variables del 1-árbol G^t y $e = (i, j)$ por $e \in E$

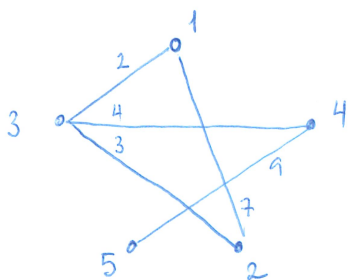
• Notad que como estamos trabajando con 1-árboles, $d_i^t = 2 \quad \forall i \rightarrow$ la 1ª ecuación del (PLM) es el doble de la ecuación de convexidad $\rightarrow$ Podemos eliminar la restricción de convexidad. (y por lo tanto μ)

Aplicación: Sea un TSP en distancias:

$$C_e = \begin{pmatrix} & 7 & 2 & 1 & 5 \\ & & 3 & 6 & 8 \\ & & & 4 & 2 \\ & & & & 9 \\ & & & & & \end{pmatrix}$$



Comenzamos con un PLM Restringido con 7 columnas, correspondientes a un tour $(1, 5, 4, 3, 2, 1)$ de coste 28 y 6 1-árboles escogidos aleatoriamente:



1-árbol de coste 25

notar que $d_1^t = 2, d_2^t = 2, d_3^t = 3, d_4^t = 2, d_5^t = 1$

(PLMR):

$$\left. \begin{array}{l} \min \quad 28 h_1 + 25 h_2 + 21 h_3 + 19 h_4 + 22 h_5 + 18 h_6 + 28 h_7 \\ \quad 2 h_1 + 2 h_2 + 2 h_3 + 2 h_4 + 2 h_5 + 2 h_6 + 2 h_7 = 2 \\ \quad 2 h_1 + 2 h_2 + 2 h_3 + 1 h_4 + 1 h_5 + 2 \quad + 3 \quad = 2 \\ \quad 2 h_1 + 3 h_2 + 2 h_3 + 3 h_4 + 2 h_5 + 3 \quad + 1 \quad = 2 \\ \quad 2 h_1 + 2 h_2 + 3 h_3 + 3 h_4 + 3 h_5 + 1 \quad + 1 \quad = 2 \\ \quad 2 h_1 + 1 h_2 + 1 h_3 + 1 h_4 + 2 h_5 + 2 \quad + 3 \quad = 2 \end{array} \right\}$$

$$h \geq 0$$

(noted qe hemos quitado la rest. de convexidad por el comentario anterior)
(y, por tanto, X, u)

La sol. óptima lineal es: $h = (0, 0, \frac{1}{4}, 0, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ de coste 22.5 y
con sol. dual óptima $u = (\frac{151}{8}, -1, \frac{-11}{2}, \frac{-5}{4}, 0)$.

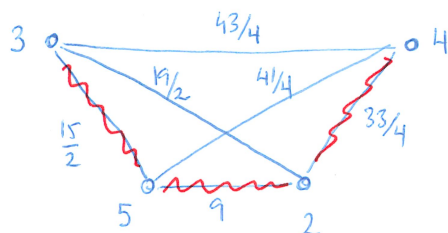
La matriz de costes reducidos para el subproblema es:

$$\begin{pmatrix} \bullet & -\frac{87}{8} & -\frac{91}{8} & -\frac{133}{8} & -\frac{111}{8} \\ & \bullet & \frac{19}{2} & \frac{33}{4} & 9 \\ & & \bullet & \frac{43}{4} & \frac{15}{2} \\ & & & \bullet & \frac{41}{4} \\ & & & & \bullet \end{pmatrix}$$

(por ejemplo $c_{12} - u_1 - u_2 = 7 - \frac{151}{8} - (-1) = -\frac{87}{8}$)

Y tenemos qe resolver el problema

$$\left. \begin{array}{l} z_1 = \min \sum_{e \in E} (\text{coste reducido aista } e) x_e \\ \quad x \in X' \text{ (los 1-árboles)} \end{array} \right\}$$



El 1-árbol óptimo es $x_{24} = x_{25} = x_{35} = 1$ junto a las 2 aristas incidentes con el vértice 1 de menor coste: $x_{14} = x_{15} = 1$. Su coste es $z_1 = -23/4$

Por lo tanto $22.5 + z_1 = 16.75 \leq z \leq \underbrace{22.5}_{z^{PLMR}}$

- Comenzamos una nueva iteración introduciendo este 1-árbol como una nueva columna en el PLM Restringido con coste $\uparrow$ z_{Cex} y grado $(2, 2, 1, 2, 3)$.

La nueva solución óptima lineal es $h = (0, 0, \frac{1}{3}, 0, 0, \frac{1}{3}, 0, \frac{1}{3})$ con coste 20.333 y sol. dual $u = (\frac{65}{6}, \frac{1}{3}, -\frac{5}{3}, \frac{2}{3}, 0)$.

La nueva matriz de costes reducidos para el subproblema es:

$$\begin{pmatrix} \bullet & -\frac{25}{6} & -\frac{43}{6} & -\frac{21}{2} & -\frac{35}{6} \\ \bullet & \frac{13}{3} & 5 & \frac{23}{3} \\ \bullet & \bullet & 5 & \frac{11}{3} \\ \bullet & \bullet & \bullet & \frac{25}{3} \\ \bullet & \bullet & \bullet & \bullet \end{pmatrix}$$

El 1-árbol óptimo es $x_{13} = x_{14} = x_{23} = x_{24} = x_{35} = 1$ con $z_1 = -\frac{14}{3}$.

La costa inferior de $20.333 - \frac{14}{3} = 15.667$ es peor que la anterior.

Por lo tanto, ahora tenemos $16.75 \leq z \leq \underbrace{20.333}_{z^{PLM}}$

- De nuevo introducimos este 1-árbol como una nueva columna en el PLM Restringido con coste 14 y grado $(2, 2, 3, 2, 1)$.

La nueva solución óptima lineal es $h = (0, 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2})$ con coste 18 y sl. dual $u = (13, 0, -4, 0, 0)$.

La matriz de costes reducidos correspondiente es

$$\begin{pmatrix} 0 & -6 & -7 & -12 & -8 \\ & 0 & 7 & 6 & 8 \\ & & 0 & 8 & 6 \\ & & & 0 & 9 \\ & & & & 0 \end{pmatrix}$$

El 1-árbol óptimo es $x_{14} = x_{15} = x_{23} = x_{24} = x_{35} = 1$ con $z_1 = -1$.

La coste inferior a z^{PLM} sube a $18 - 1 = 17$. ($17 \leq z^{PLM} \leq 18$)

Como este 1-árbol es un tour, por el criterio de parada alternativo, es óptimo. (Alternativamente, se podría comprobar que el coste real es 17).

Reformulación Valor

del PLineal Master

¿Cuán fuerte es la relajación lineal del PLMaster? ¿Existe alguna esperanza de que puede resolver el PL original?

Proposición :

$$z^{PLM} = \max \left\{ \sum_{k=1}^K c^k x^k : \sum_{k=1}^K A^k x^k = b, x^k \in \text{conv}(X^k) \text{ para } k=1 \dots K \right\}$$

Demostreci : Obvia, ya que (PLM) puede obtenerse a partir del PL original sustituyendo

$$x^k = \sum_{t=1}^{T_k} x^{k,t} \lambda_{k,t}, \quad \sum_{t=1}^{T_k} \lambda_{k,t} = 1, \quad \lambda_{k,t} \geq 0 \quad \forall t=1 \dots T_k, \text{ y esto es equivalente a}$$

sustituir $x^k \in \text{conv}(X^k)$.

Como discutimos en la Introducción, cuando el PL puede descomponerse, la Rel. Lag. y los algoritmos de planos de corte son 2 aproximaciones alternativas.

Específicamente, sea $z(D)$ el valor del dual lagrangiano cuando las restricciones conjuntas $\sum_{k=1}^K A^k x^k = b$ son relajadas y sea z^{WT} el valor obtenido cuando planos de corte se han añadido a la Rel. Lineal del PL usando un algoritmo exacto de separación para cada uno de los conjuntos $\text{row}(X^k)$, $k=1 \dots K$.

El resultado siguiente muestra que las 3 aproximaciones son, en cierto sentido, equivalentes, ya que conducen a las mismas cotas. Está basado en la Proposición anterior, un resultado de Rel. Lag. y en el hecho de que un algoritmo de separación exacto para $\text{conv}(X^k)$, genera implícitamente $\text{conv}(X^k)$.

Teorema : $z^{PLM} = z(D) = z^{WT}$

Como los subproblemas que se resuelven tanto en la Generación de Columnas como con el Dual Lagrangiano son problemas de optimización sobre X^k , la generación de columnas puede considerarse como un algoritmo para resolver el Dual

Consideremos el PLE siguiente:

$$(P) \quad \left. \begin{array}{l} \max \quad cx \\ Ax \leq b \\ Dx \leq d \\ x \geq 0 \text{ y entera} \end{array} \right\}$$

Se puede demostrar para $X = \{x \in \mathbb{Z}_+^n : Dx \leq d\}$ cualquier, pero supongamos que es acotado. Así X tendrá un n° enorme pero finito de puntos $\{x^1, \dots, x^T\}$

Recordemos que $(PR_L): \quad \left. \begin{array}{l} \max \quad cx + L(b - Ax) \\ x \in X \end{array} \right\}$
para $L \geq 0$

y que $z(D) = \min_{L \geq 0} z(PR_L)$

Teorema: $z(D) = \left. \begin{array}{l} \max \quad cx \\ Ax \leq b \\ x \in \text{conv}(X) \end{array} \right\}$

Demostración:

$$z(D) = \min_{L \geq 0} z(PR_L) = \min_{L \geq 0} \left\{ \max_{x \in X} cx + L(b - Ax) \right\} =$$

$$= \min_{L \geq 0} \left\{ \max_{t=1,2,\dots,T} cx^t + L(b - Ax^t) \right\} =$$

$$= \min \left\{ \begin{array}{l} \eta \\ \eta \geq cx^t + L(b - Ax^t) \quad \forall t \\ L \geq 0 \quad (L \in \mathbb{R}_+^T) \\ \eta \in \mathbb{R}^1 \end{array} \right\} = \text{plantando el dual} =$$

$$= \left\{ \begin{array}{l} \max \quad \sum_{t=1}^T \mu_t (cx^t) \\ \sum_{t=1}^T \mu_t (Ax^t - b) \leq 0 \\ \sum_{t=1}^T \mu_t = 1 \\ \mu \in \mathbb{R}_+^T \end{array} \right\}$$

Most, haciendo $x = \sum_{t=1}^T \mu_t x^t$, con $\sum_t \mu_t = 1$, $\mu_t \in \mathbb{R}_+^T$, tenemos

$$z(D) = \max \left. \begin{array}{l} cx \\ Ax \leq b \\ x \in \text{conv}(X) \end{array} \right\}$$

Notar que en (1), podemos escribir el problema como:

$$\left. \begin{array}{l} \min \quad 1 \cdot \eta + 0 \cdot L \\ \eta + L(Ax^t - b) \geq cx^t \quad \forall t=1-T \\ L \geq 0 \quad (\text{~~no~~ } L \in \mathbb{R}_+^T) \\ \eta \text{ cualquier} \end{array} \right\} \leftarrow \mu_t \text{ v. duales}$$

$$\text{su dual es:} \quad \left. \begin{array}{l} \max \quad \sum_{t=1}^T \mu_t (cx^t) \\ \sum_t \mu_t = 1 \\ \sum_t \mu_t (Ax^t - b) \leq 0 \\ \mu_t \geq 0 \quad \forall t \end{array} \right\} \leftarrow (2)$$

y (2) se puede escribir como:

$$Ax \leq \sum_t \mu_t b = b \underbrace{\sum_t \mu_t}_1 = b$$

$$\text{donde } x = \sum_{t=1}^T \mu_t x^t$$

Logramos en el qe las variables duales π son actualizadas usando Prog. Lineal para resolver el PLM Restringido. Este será el método de actualización de multiplicadores qe, en comparación con el subgradiente, es mucho más complicado.

Por otra parte, si usamos el método de planos de corte, aunque potencialmente la cota obtenida sea la misma, debemos resolver problemas de separación en $\text{conv}(X^k)$ en lugar de resolver problemas de optimización.

Como la complejidad teórica de los problemas de optimización y separación para $\text{conv}(X^k)$ es la misma, la elección de un método u otro dependerá de la dificultad relativa de resolver ambos problemas así como de la convergencia en la práctica de la generación de columnas y de los algoritmos de planos de corte.

GENERACIÓN DE COLUMNAS PARA UN PLE BINARIO

Si al acabar el algoritmo de generación de columnas la solución óptima $\tilde{L} = (\tilde{L}_1, \dots, \tilde{L}_K)$ del PLM no es entera, entonces aún no hemos resuelto el PLEM.

Sin embargo, $z^{\text{PLM}} \geq z$, lo qe sugiere la posibilidad de usar estas cotas superiores en un B&B. Ahora presentaremos un algoritmo para PLE's 0-1 qe usa esta cota, llamado algoritmo de Generación de Columnas para PLE o Branch-and-Rice.

Consideremos de nuevo el problema (P) original:

$$(P) \quad \begin{cases} z = \max \sum_{k=1}^K c^k x^k \\ \sum_{k=1}^K A^k x^k = b \\ D^k x^k \leq d^k \quad \text{para } k=1 \dots K \\ x^k \in B^{u_k} \quad \text{para } k=1 \dots K \end{cases}$$

y su reformulación:

$$\begin{aligned}
 (\text{PLEM}) \quad \left\{ \begin{array}{l}
 z = \max \sum_{k=1}^K \sum_{t=1}^{T_k} (c^k x^{k,t}) l_{k,t} \\
 \sum_k \sum_t (A^k x^{k,t}) l_{k,t} = b \\
 \sum_t l_{k,t} = 1 \quad k=1 \dots K \\
 l_{k,t} = 0,1 \quad t=1, \dots, T_k, \quad k=1, \dots, K
 \end{array} \right.
 \end{aligned}$$

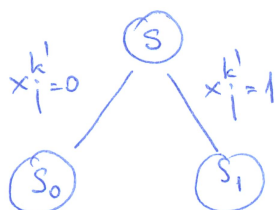
cuya relajación lineal tiene sol. óptima $\tilde{L}$.

Como los puntos $x^{k,t} \in X^k$ son vectores 0-1 diferentes, notemos que

$$\tilde{x}^k = \sum_{t=1}^{T_k} \tilde{L}_{k,t} x^{k,t} \in \{0,1\}^{u_k} \iff \tilde{L} \text{ es entero (pues } \sum \tilde{L}_{k,t} = 1)$$

Por lo tanto, si $\tilde{L}$ no es entero $\rightarrow \exists k', j$ tal que la correspondiente variable 0-1 $x_j^{k'}$ tiene valor lineal $\tilde{x}_j^{k'}$ que es fraccionario y sobre el que podemos ramificar.

Esto origina el esquema de ramificación siguiente:



donde S se ha separado en $S_0 = S \cap \{x: x_j^{k'} = 0\}$ y $S_1 = S \cap \{x: x_j^{k'} = 1\}$ (esquema típico de cualquier algoritmo de BdB).

Es fundamental estar seguros de que será posible resolver los nuevos PLM's en dificultad. Por ello necesitamos definir los nuevos PLEM's asociados con S_0 y S_1 y los nuevos subproblemas.

Como $x_j^k = \sum_{t=1}^{T_k} \lambda_{k,t} x_j^{k,t} \in \{0,1\}$, $x_j^k = \delta \in \{0,1\}$ implica que $x_j^{k,t} = \delta \quad \forall k,t$ en $\lambda_{k,t} > 0$. (Recordemos que $\sum_{t=1}^{T_k} \lambda_{k,t} = 1$).

Por lo tanto el (PLEM) en el nodo $S_i = S \cap \{x = x_j^{k'} = i\}$ para $i=0,1$ es:

$$\begin{aligned}
 (\text{PLEM}(S_i)) \quad \left\{ \begin{aligned}
 z(S_i) &= \max \sum_{k \neq k'} \sum_t (c^k x^{k,t}) \lambda_{k,t} + \sum_{t: x_j^{k',t} = i} (c^{k'} x^{k',t}) \lambda_{k',t} \\
 \sum_{k \neq k'} \sum_t (A^k x^{k,t}) \lambda_{k,t} + \sum_{t: x_j^{k',t} = i} (A^{k'} x^{k',t}) \lambda_{k',t} &= b \\
 \sum_t \lambda_{k,t} &= 1 \quad \forall k \neq k' \\
 \sum_{t: x_j^{k',t} = i} \lambda_{k',t} &= 1 \\
 \lambda_{k,t} &= 0,1 \quad \forall t, \forall k
 \end{aligned} \right.
 \end{aligned}$$

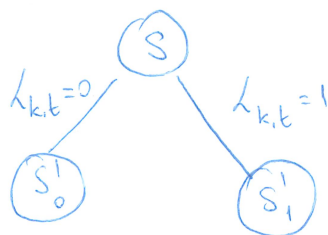
que tiene la misma forma que el (PLEM) original excepto que un conjunto de columnas son excluidas en cada rama y la solución anterior al (PLM) es ahora inaplicable.

Volviendo a los subproblemas de generación de columnas, el subproblema no cambia si $k \neq k'$. Sin embargo, por el subproblema k' y $i=0,1$ tenemos:

$$z_{k'}(S_i) = \max \left\{ \begin{aligned}
 &(c^{k'} - \pi A^{k'}) x - \mu_{k'} \\
 &x \in X^{k'} \\
 &x_j = i
 \end{aligned} \right\}$$

que es muy similar al subproblema original.

Nota: Otra posibilidad es ramificar sobre alguna variable $\lambda_{k,t}$ fraccionaria, fijándola a 0 o 1, respectivamente, como en la figura siguiente:



Sin embargo, notemos que en la rama de la izquierda lo que estamos haciendo es excluir una única columna (correspondiente a la t -ésima solución del subproblema k), de forma que el problema resultante es prácticamente idéntico al problema original. Esto implica que el árbol de enumeración tiene la característica no deseable de ser muy desequilibrado. Además, a menudo es difícil imponer la condición $L_{k,t}=0$ y, así, impedir que la misma solución sea generada de nuevo como solución óptima después de la ramificación.

Nota 2: Una ventaja potencial de la generación de columnas es que las soluciones óptimas al (PLRM) son a menudo enteras, o casi enteras. En el primer caso obtenemos una solución perfectamente entera y en el segundo una solución que "redondeada" con un heurístico puede convertirse en perfecta.

PROBLEMAS IMPLÍCITOS DE PARTICIONAMIENTO/EMPAQUETAMIENTO

Una clase importante de PLE's 0-1 que pueden descomponerse en los problemas de packing y partitioning.

Dado un qto finito $M = \{1, \dots, m\}$, existen K conjuntos de subconjuntos posibles que están descritos de forma implícita. El problema consiste en encontrar un empaquetamiento o partición de M , consistente en algunos de esos subconjuntos, con valor máximo.

En términos del PLE original :

$$(PLE) \quad \left. \begin{aligned} z &= \max \sum_{k=1}^K c^k x^k \\ \sum_{k=1}^K A^k x^k &= b \\ x^k &\in X^k \quad k=1..K \end{aligned} \right\}$$

hacemos $x^k = (y^k, w^k)$ con $y^k \in \{0,1\}^m$ el vector de incidencia del subconjunto k de M , $c^k = (e^k, f^k)$, $A^k = (I, 0)$ y $b=1$.

Podemos pensar en las variables w^k como variables auxiliares necesarias para definir si el subconjunto con vector de incidencia y^k es primal y definir el valor de la f. objetivo (primitivamente no lineal) del correspondiente subconjunto.

Así, tendríamos la formulación :

$$\left. \begin{aligned} z &= \max \sum_{k=1}^K (e^k y^k + f^k w^k) \\ \sum_{k=1}^K y^k &\leq 1 \\ (y^k, w^k) &\in X^k \quad k=1, \dots, K \end{aligned} \right\}$$

Ahora, si $(y^{k,t}, w^{k,t})$ corresponde a la t -ésima sol. primal de X^k y $\lambda_{k,t}^{k,t}$ es la correspondiente variable, obtenemos el equivalente (PLEM):

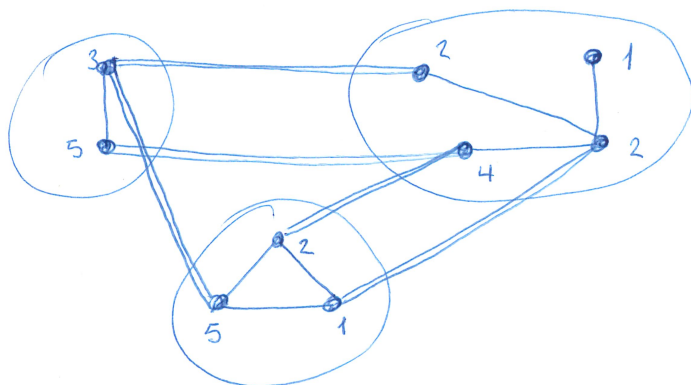
$$\left\{ \begin{aligned} z &= \max \sum_{k=1}^K \sum_{t=1}^{T_k} (e^k y^{k,t} + f^k w^{k,t}) \lambda_{k,t} \\ \sum_k \sum_{t: y^{k,t}=1} \lambda_{k,t} &= 1 \quad \forall i \in M \\ \sum_{t=1}^{T_k} \lambda_{k,t} &\leq 1 \quad k=1..K \\ \lambda_{k,t} &\in \{0,1\} \quad \forall t=1, \dots, T_k, \quad \forall k=1, \dots, K \end{aligned} \right.$$

Un problema de este tipo es el siguiente:

CLUSTERING : Dado un grafo $G = (V, E)$, costs $c_e \forall e \in E$,

pesos de los vértices $d_i \forall i \in V$ y una capacidad del cluster C , deseamos dividir el conjunto de vértices V en K (prácticamente alguno varios) clusters satisfaciendo la propiedad de que la suma de los pesos de los vértices en cada cluster no excede la capacidad C y de forma que se minimice la suma de los costs de las aristas entre clusters (maximice la suma de los costs de las aristas dentro de los clusters).

La figura siguiente muestra una solución ^{prácticamente} para una instancia con 3 clusters y $C=9$. Las aristas en trazo grueso son las que van entre clusters:



El problema puede formularse como:

$$\left. \begin{aligned} \max \quad & \sum_{k=1}^K \sum_{e \in E} c_e w_e^k \\ & \sum_{k=1}^K y_i^k \leq 1 \quad \forall i \in V \\ & (w_e^k, y_i^k) \in X^k \quad \forall k=1, \dots, K \end{aligned} \right\}$$

$$\text{donde } X^k = \left\{ (w_e^k, y_i^k) \in B^m \times B^n : w_e^k \leq y_i^k; w_e^k \leq y_j^k; w_e^k \geq y_i^k + y_j^k - 1 \right. \\ \left. \text{por } e = (i, j) \in E; \sum_{i \in V} d_i y_i^k \leq C \right\}$$

$$\text{con } y_i^k = \begin{cases} 1 & \text{si el vértice } i \text{ está en el cluster } k \\ 0 & \text{—} \end{cases}$$

$$\text{y } w_e^k = \begin{cases} 1 & \text{si la arista } e \text{ tiene sus 2 vértices incidentes en el cluster } k \\ 0 & \text{—} \end{cases}$$

Otro ejemplo es :

CVRP (Capacitated Vehicle Routing Problem) : Dado un grafo $G=(V,E)$, un vértice depósito 0, costs $c_e \forall e \in E$, K vehículos idénticos de capacidad C , y demandas de los clientes $d_i, \forall i \in V \setminus \{0\}$, deseamos encontrar un conjunto de rutas (ciclos) para cada vehículo de forma que :

- cada ruta contenga al depósito
- las rutas conjuntamente contengan todos los vértices
- las rutas sean disjuntas (excepto respecto del depósito)
- la demanda total de un ruta no exceda C
- el coste total sea mínimo.

Particionamiento con subconjuntos idénticos

El problema de Clustering y el CVRP tienen la propiedad de que los clusters o vehículos son intercambiables (independiente de k). Esto significa que la numeración de los subconjuntos es arbitraria e **intercambiar** 2 conjuntos cualesquiera nos conduce a la misma solución.

Vemos cómo especializar el algoritmo de Generación de Columnas para que tenga en cuenta esta simetría.

Como $X^k = X$, $(e^k, f^k) = (e, f)$ y $T_k = T \forall k$, podemos hacer

$$L_t = \sum_{k=1}^K L_{k,t} \quad \text{y el (PLEM) toma la forma :}$$

$$\left. \begin{aligned} \max \quad & \sum_{t=1}^T (e y^t + f w^t) L_t \\ & \sum_{t: y_i^t=1} L_t = 1 \quad \forall i \in M \\ & \sum_{t=1}^T L_t \leq K \\ & L \in B^T \end{aligned} \right\}$$

Sílo hay un subproblema de generación de columnas.

Representando por $\{\pi_i\}_{i \in M}$ las v. duales asociadas en la relajación lineal así como por μ , el subproblema es:

$$z = \max \{ (e - \pi)y + \omega - \mu : (y, \omega) \in X \}$$

y el (PLM) puede resolverse como hemos visto.

¿Qué ocurre si la solución $\tilde{z}$ del PLM no es entera?

Ahora no es evidente cómo recuperar las variables originales x^k o las variables $L_{k,t}$, por lo tanto debemos modificar el esquema de ramificación.

Consideramos 2 posibilidades:

(i) Si $\sum_{t=1}^T L_t = \alpha \notin \mathbb{Z}$, entonces formar 2 ramas con $\sum_t L_t \leq \lfloor \alpha \rfloor$ y $\sum_t L_t \geq \lceil \alpha \rceil$, respectivamente.

(ii) Notemos que si tomamos 2 elementos de M , o bien ambos aparecen juntos en un subconjunto, o no. Así, escogemos un par de elementos (filas) i y j de M por lo que

$$0 < \sum_{t: y_i^t = y_j^t = 1} L_t < 1$$

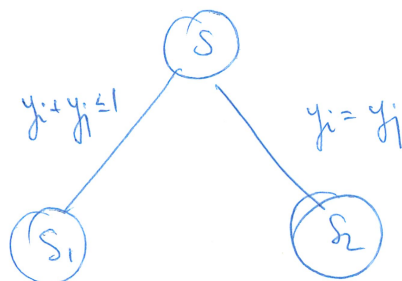
y formamos 2 ramas con $\sum L_t = 1$ y $\sum L_t = 0$, respectivamente.

En el primer caso (i) imponemos que i y j estén en el mismo subconjunto y en el segundo (ii) que estén en subconjuntos diferentes.

En (i) todas las columnas correspondientes a subconjuntos α conteniendo a i o a j pero no a ambos son eliminadas del Master y la restricción $y_i = y_j$ se añade al subproblema para asegurar que cualquier nueva columna generada no corresponda a un subconjunto conteniendo a i pero no a j , o viceversa. En (ii) se eliminan del Master las columnas conteniendo a

i y j a la vez, y la restricción $y_i + y_j \leq 1$ le añade al subproblema.

Así pues, añadiendo las restricciones $y_i = y_j$ o $y_i + y_j \leq 1$ en cada subproblema nos permite ramificar como sigue:



El siguiente resultado nos dice que este segundo esquema de branching es suficiente.

Proposición : Si $\tilde{z} \notin B^T$, existen filas $i, j \in M$ tales que

$$0 < \sum_{k: y_i^k = y_j^k = 1} z_k < 1$$

(Tal pareja de filas no es difícil de encontrar).