0.1 MÉTODO DE BENDERS

Adaptado de Jensen Paul A., *Benders' Method*,

`www.me.utexas.edu/~jensen/`

Este método sirve para resolver un problema de optimización (lineal) mixta, algunas variables son enteras, las otras son reales.

$$\begin{aligned} \max \quad & z(x,y) = c^{\mathrm{T}}x + f^{\mathrm{T}}y \\ & Ax + Gy \le b \\ & x \in \mathbb{Z}_+^n, \quad y \in \mathbb{R}_+^p \ , \end{aligned} \tag{P_M}$$

con $A \in \mathbb{R}^{m\times n}$ (con frecuencia $A \in \mathbb{Z}^{m\times n}$), $G \in \mathbb{R}^{m\times p}$. Los demás son vectores columna del tamaño adecuado.

En lo que sigue se utiliza la notación $\xi = (\xi_1, \xi_2, ..., \xi_n) = [\xi_1 \ \ \xi_2 \ \ ... \ \ \xi_n]^{\mathrm{T}}$ y entonces $\xi^{\mathrm{T}} = [\xi_1 \ \ \xi_2 \ \ ... \ \ \xi_n]$. Dicho de otra forma, un vector en $\mathbb{R}^n$ es exactamente lo mismo que una matriz columna de n filas.

Sean $\mathcal{F}_M$ el conjunto factible, x^* y y^* los vectores óptimos y $z^* = z(x^*, y^*)$ el valor máximo.

Dado un $x \in \mathbb{Z}_+^n$ un vector fijo. El problema resultante es un problema de OL (optimización lineal).

$$\begin{aligned} \max \quad & \zeta_x(y) = c^{\mathrm{T}}x + f^{\mathrm{T}}y \\ & Gy \le b - Ax \\ & y \in \mathbb{R}_+^p \ . \end{aligned} \tag{P_x}$$

En este problema, $c^{\mathrm{T}}x$ es constante y se puede quitar para resolver el problema. Sea $\mathcal{F}_{P_x}$ el conjunto factible de este problema. Si y_x^* es el maximizador (depende de x), entonces el valor máximo de ζ_x es

$$\zeta_x^* = f^{\mathrm{T}}y_x^* + c^{\mathrm{T}}x.$$

Obviamente

$$\zeta_x^* \le z^* \quad \forall x.$$

El problema dual de P_x es

$$\begin{array}{lll} \min & w(u) = (b-Ax)^{\mathrm{T}}u & \\ & G^{\mathrm{T}}u \geq f & (D_x) \\ & u \in \mathbb{R}^m_+ \ . & \end{array}$$

Sean u_x^* el minimizador y w_x^* el valor mínimo. Entonces

$$\zeta_x^* = w_x^* + c^{\mathrm{T}}x.$$

Aunque este problema depende de x, el conjunto factible no depende de x. Sea $\mathcal{F}_D$ el conjunto factible del problema (D_x), para cualquier x.

Sea $E = \{u^1, u^2, ..., u^r\}$ el conjunto de puntos extremos de $\mathcal{F}_D$. Si $w_x^* = -\infty$ (óptimo no acotado), entonces, por dualidad, $\mathcal{F}_{P_x} = \emptyset$. Si suponemos que el valor w_x^* es finito, entonces se debe obtener en un punto extremo. Se concluye que para resolver D_x basta con estudiar el valor de w en los puntos extremos.

$$w_x^* = \min_{1\leq i\leq r}\{(b-Ax)^{\mathrm{T}}u^i\}$$

Entonces

$$\zeta_x^* = c^{\mathrm{T}}x + \min_{1\leq i\leq r}\{(b-Ax)^{\mathrm{T}}u^i\}$$

Para obtener z^*, se debe hacer variar x,

$$\begin{aligned} z^* &= \max_{x\in\mathbb{Z}^n_+}\{\zeta_x^*\} \\ z^* &= \max_{x\in\mathbb{Z}^n_+}\{\ c^{\mathrm{T}}x + \min_{1\leq i\leq r}\{(b-Ax)^{\mathrm{T}}u^i\}\ \} \\ z^* &= \max_{x\in\mathbb{Z}^n_+}\{\ \min_{1\leq i\leq r}\{c^{\mathrm{T}}x + (b-Ax)^{\mathrm{T}}u^i\}\ \} \end{aligned}$$

Sea

$$\eta = \min_{1\leq i\leq r}\{c^{\mathrm{T}}x + (b-Ax)^{\mathrm{T}}u^i\},$$

entonces

$$\eta \leq c^{\mathrm{T}}x + (b-Ax)^{\mathrm{T}}u^i\ , \quad i = 1, ..., r.$$

Esto indica que z^* se puede obtener como el valor máximo del problema

$$\begin{array}{lll} \max & \varphi(x,\eta) = \eta & \\ & \eta \leq c^{\mathrm{T}}x + (b-Ax)^{\mathrm{T}}u^i, \quad i = 1, ..., r, & (P'_M) \\ & x \in \mathbb{Z}^n_+ . & \end{array}$$

Sea $K = \{u^1, u^2, ..., u^k\} \subseteq E$. Al quitar las últimas $r-k$ restricciones de (P'_M) se obtiene

$$\begin{aligned} \max \quad & \varphi(x,\eta) = \eta \\ & \eta \le c^{\mathrm{T}}x + (b-Ax)^{\mathrm{T}}u^i, \quad i = 1, ..., k, \\ & x \in \mathbb{Z}^n_+ . \end{aligned} \tag{P'_k}$$

Dejando las variables x_1, x_2, ..., x_n, η a la izquierda y las constantes a la derecha, el problema se puede reescribir

$$\begin{aligned} \max \quad & \varphi(x,\eta) = \eta \\ & \hat{M}^k \begin{bmatrix} x \\ \eta \end{bmatrix} \le p^k \\ & (x,\eta) \in \mathbb{Z}^n_+ \times \mathbb{R}, \end{aligned} \tag{P'_k}$$

donde $\hat{M}^k$ es una matriz $k \times (n+1)$, b es un vector columna $k \times 1$,

$$\hat{M}^k = \begin{bmatrix} {u^1}^{\mathrm{T}}A - c^{\mathrm{T}} & 1 \\ {u^2}^{\mathrm{T}}A - c^{\mathrm{T}} & 1 \\ \vdots & \\ {u^k}^{\mathrm{T}}A - c^{\mathrm{T}} & 1 \end{bmatrix}, \quad p^k = \begin{bmatrix} b^{\mathrm{T}}u^1 \\ b^{\mathrm{T}}u^2 \\ \vdots \\ b^{\mathrm{T}}u^k \end{bmatrix}.$$

P'_k es un problema mixto, hay una variable real. Una manera de resolverlo es mediante el método de bisección que será presentado en la siguiente sección.

Como (P'_k) es una relajación de (P'_M) entonces

$$z^* \le \eta^*_k .$$

En resumen, se tienen dos cotas para z^*,

$$\zeta^*_x \ \le \ z^* \ \le \ \eta^*_k .$$

Hasta ahora sea ha supuesto que $\mathcal{F}_D \neq \emptyset$. En caso contrario, si $\mathcal{F}_D = \emptyset$, entonces, por dualidad, (P_M) tiene óptimo no acotado o el conjunto factible $\mathcal{F}_M$ es vacío.

Para evitar que el problema (D_x) tenga óptimo no acotado (w^*_x no sea finito) se puede agregar la restricción

$$u_1 + u_2 + \cdots + u_m \le L \quad \text{(valor positivo grande)}.$$

ALGORITMO DE BENDERS

datos: c, f, A, G, b, ε, $MAXIT$
si $\mathcal{F}_D = \emptyset$ **ent**
    $\mathcal{F}_{P_M} = \emptyset$ ó
    $z^* = \infty$
    parar
fin-si
obtener $u^1 \in \mathcal{F}_D$
$\underline{z} = -\infty$
$\bar{z} = \ \infty$
para $k = 1, ..., MAXIT$
    (x^k, η_k) solución del problema P'_k
    $\bar{z} = \min\{\bar{z}, \eta_k\}$
    calcular el vector de costos para el problema D_{x^k}
    sean u^{k^*} el minimizador y w_k el valor mínimo
        del problema lineal D_{x^k}
    $u^{k+1} = u^{k^*}$
    $\zeta_k = c^{\mathrm{T}} x^k + w_k$
    $\underline{z} = \max\{\underline{z}, \zeta_k\}$
    si $\bar{z} - \underline{z} \leq \varepsilon$ **ent**
        $x^* = x^k$
        $y^* =$ solución dual de D_{x^k}
        parar
    fin-si
fin-para

Esta versión simplificada del método de Benders, presenta dificultades con algunos problemas. Una versión más completa y robusta del método se puede encontrar en Nemhauser G.L y Wolsey L.A., *Integer and Combinatorial Optimization*, Wiley, New York, 1999, pág. 412.

Ejemplo:

$$
\begin{array}{ll}
\max & z = 15x_1 + 25x_2 + 30x_3 + \ \ 5y_1 + 8y_2 + \ \ 9y_3 + 8y_4 \\
& \quad 11x_1 + 12x_2 + 13x_3 + 14y_1 + 4y_2 + \ \ 6y_3 + 7y_4 \leq \ 40 \\
& \quad 20x_1 + 22x_2 + 25x_3 + 16y_1 + 6y_2 + 10y_3 + 8y_4 \leq \ 74 \\
& \hfill x \in \mathbb{Z}^3_+ \\
& \hfill y \in \mathbb{R}^4_+ \ .
\end{array}
$$

Al resolver la relajación lineal se obtiene

$$x = (0, 0, 2.5455), \quad y = (0, 1.7273, 0, 0), \quad z = 90.1818.$$

La restricciones para D_x son :

$$\begin{bmatrix} 14 & 16 \\ 4 & 6 \\ 6 & 10 \\ 7 & 8 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \geq \begin{bmatrix} 5 \\ 8 \\ 9 \\ 8 \\ -100 \end{bmatrix}$$

$$u \in \mathbb{R}^2_+ .$$

La última restricción corresponde a $u_1 + u_2 \leq 100$.

Un punto factible para D_x es $u = (0,\ 1.3333)$, las cotas iniciales son $\underline{z} = -\infty$ y $\bar{z} = 90.8018$.

Para el problema P'_1:

$$\hat{M}^1 = \left[\ 11.6667 \quad 4.3333 \quad 3.3333 \quad 1\ \right], \quad p^1 = \left[\ 98.6667\ \right].$$

La solución es:

$$x = (0, 0, 0), \quad \eta = 98.6667.$$

La cota superior sigue siendo:

$$\bar{z} = 90.1818.$$

El vector de costos para D_x es

$$b - Ax = (40,\ 74).$$

La solución de D_x es

$$u = u^2 = (2, 0), \quad w = 80.$$

Entonces

$$c^{\mathrm{T}} x = 0, \quad \zeta_x = 80, \quad \text{luego}\ \ \underline{z} = 80.$$

Para el problema P'_2:

$$\hat{M}^2 = \begin{bmatrix} 11.6667 & 4.3333 & 3.3333 & 1 \\ 7 & -1 & -4 & 1 \end{bmatrix}, \quad p^2 = \begin{bmatrix} 98.6667 \\ 80 \end{bmatrix}.$$

La solución es:
$$x = (0, 0, 3), \quad \eta = 88.6667\,.$$
La nueva cota superior es:
$$\bar{z} = 88.6667$$
El vector de costos para D_x es
$$b - Ax = (1, -1).$$
La solución de D_x es
$$u = (0, 100), \quad w = -100.$$
Entonces
$$c^{\mathrm{T}}x = 90, \quad \zeta_x = -10, \quad \text{luego} \;\; \underline{z} = 80.$$
Para el problema P_3':
$$\hat{M}^3 = \begin{bmatrix} 11.6667 & 4.3333 & 3.3333 & 1 \\ 7 & -1 & -4 & 1 \\ 1985 & 2175 & 2470 & 1 \end{bmatrix}, \quad p^2 = \begin{bmatrix} 98.6667 \\ 80 \\ 7400 \end{bmatrix}.$$
La solución es:
$$x = (0, 0, 2), \quad \eta = 88\,.$$
La nueva cota superior es:
$$\bar{z} = 88.$$
El vector de costos para D_x es
$$b - Ax = (14,\ 24).$$
La solución de D_x es
$$u = (2, 0), \quad w = 28.$$
Entonces
$$c^{\mathrm{T}}x = 60, \quad \zeta_x = 88, \quad \text{luego} \;\; \underline{z} = 88.$$

Como las dos cotas coinciden, entonces se tiene el óptimo, $x^* = x = (0, 0, 2)$. Para y se utiliza la solución del dual del último D_x. En este caso $y^* = (0, 3.5, 0, 0)$

0.2 Solución de P_k'

Sea $\eta_k^* = \eta_k$ el valor máximo de la función objetivo. En esta sección, para simplificar la notación, no se tiene en cuenta el índice (subíndice o superíndice) k y el problema P_k' se escribe en una forma ligeramente diferente. El vector columna de unos se denotara por e. Entonces η^* es valor óptimo del problema

$$\begin{aligned} \max \;\; & \eta \\ M\,x + \eta\, e & \leq p \\ x & \in \mathbb{Z}_+^n \\ \eta & \in \mathbb{R}. \end{aligned} \tag{P_k'}$$

Hay que empezar con dos valores $\alpha_0 < \beta_0$ tales que el problema entero obtenido al hacer $\eta = \alpha_0$ tenga por lo menos un punto (entero) factible y el problema entero resultante de fijar $\eta = \beta_0$ no tenga puntos factibles. Necesariamente $\eta^* \in [\alpha_0, \beta_0]$. Se calcula el punto medio μ_0. Entonces $\eta^* \in [\alpha_0, \mu_0]$ o $\eta^* \in [\mu_0, \beta_0]$. Uno de estos dos intervalos será el nuevo intervalo $[\alpha_1, \beta_1]$. Así se continúa hasta obtener un intervalo $[\alpha_t, \beta_t]$ suficientemente pequeño.

De manera más precisa, denotaremos por P_η' el problema de factibilidad entero que depende de η y por $\mathcal{F}_\eta$ el conjunto factible:

$$\begin{aligned} \text{obtener } & x \\ M\,x & \leq p - \eta e \\ x & \in \mathbb{Z}_+^n\,, \end{aligned} \tag{P_η'}$$

$$\mathcal{F}_\eta = \{x \in \mathbb{Z}_+^n \;:\; M\,x \leq p - \eta e\}.$$

Para resolver P_η' basta con utilizar un método de optimización entera, por ejemplo bifurcación y acotamiento, con $c = 0$.

datos: α_0, β_0, ε , tales que $\mathcal{F}_{\alpha_0} \neq \emptyset$, $\mathcal{F}_{\beta_0} = \emptyset$
$t = 1$
mientras $\beta_t - \alpha_t \geq \varepsilon$
    $\mu_t = (\alpha_t + \beta_t)/2$
    resolver el problema de factibilidad entera P'_{μ_t}
    si $\mathcal{F}_{\mu_t} \neq \emptyset$
        $\alpha_{t+1} = \mu_t$
        $\beta_{t+1} = \beta_t$
    sino
        $\alpha_{t+1} = \alpha_t$
        $\beta_{t+1} = \mu_t$
    fin-si
    $t = t + 1$
fin-mientras
$\eta^* \approx \alpha_t$
$x^k \in \mathcal{F}_{\alpha_t}$

Encontrar los valores α_0 y β_0 se puede hacer por medio de ensayo y error. El esquema de ensayo y error, en lugar de ser ciego, puede ser guiado para lograr mayor rapidez. Sea (x^L, η_L) la solución del siguiente problema de OL (pero con la variable η no restringida):

$$\begin{aligned} \max \ \ & \eta \\ M\,x + \eta\,e &\leq p \\ x &\in \mathbb{R}^n_+ \\ \eta &\in \mathbb{R}\,. \end{aligned} \qquad (P'_L)$$

Si $x^L \in \mathbb{Z}^n_+$, entonces $\eta^* = \eta_L$ y no hay que hacer nada más. En caso contrario se resuleve el problema entero P'_{η_L}. Si $\mathcal{F}_{\eta_L} \neq \emptyset$, entonces $\eta^* = \eta_L$ y no hay que hacer nada más.

Aunque el paso anterior parece superfluo, puede ser útil si P'_L tiene varios vectores solución x^L, algunos enteros otros no.

Cuando $\mathcal{F}_{\eta_L} = \emptyset$, entonces ya se tiene $\beta_0 = \eta_L$. Se requiere ahora encontar el valor α_0. Para eso es necesario contruir parte de una sucesión real $\{\tau_i\}$ con las siguiente características:

- $\tau_0 = \beta_0$,

- $\tau_{i+1} < \tau_i, \ \ \forall i,$
- $\lim\limits_{i\to\infty} \tau_i = -\infty.$

Sea $\delta > 0$. Ejemplos sencillos de tales sucesiones son: $\tau_i = \tau_0 - i\delta$ y $\tau_i = \tau_0 - 2^i\delta$. Esta segunda sucesión avanza más rápidamente.

Los términos de la sucesión se construyen hasta encontrar el primer valor τ_κ tal que

$$\mathcal{F}_{\tau_\kappa} \neq \emptyset.$$

Entonces

$$\begin{aligned} \alpha_0 &= \tau_\kappa, \\ \beta_0 &= \tau_{\kappa-1}. \end{aligned}$$

Estos dos valores sirven para empezar el proceso de bisección, sin embargo se puede mejorar α_0 para obtener un intervalo más pequeño. Como para α_0 hay puntos enteros factibles, sea x^{α_0} uno de estos puntos (el obtenido al resolver el problema de factibilidad). Se busca aumentar el valor de η hasta dejar sin holgura una de las desigualdades. Sea s el vector (no negativo) de holguras para el valor α_0,

$$s = p - \alpha_o e - Mx^{\alpha_0}.$$

Sea

$$\sigma = \alpha_0 + \min\{s_1, s_2, ...\}.$$

Entonces α_0 se puede llevar hasta este valor, o sea,

$$\alpha_0' = \sigma.$$

De manera semejante, si durante el proceso de bisección, para $\eta = \mu_t$, se encuentra un nuevo punto entero (diferente del último obtenido), el valor μ_t se puede aumentar hasta anular una holgura.

Ejemplo.

$$\begin{aligned} \max \quad \eta \\ 11.6667x_1 + 4.3333x_2 + 3.3333x_3 + \eta \leq 98.6667 \\ (x, \eta) \in \mathbb{Z}_+^3 \times \mathbb{R}. \end{aligned}$$

Inicialmente se resuelve el problema de OL

$$\begin{aligned} \max \quad & \eta \\ 11.6667x_1 + 4.3333x_2 + 3.3333x_3 + \eta &\leq 98.6667 \\ (x, \eta) &\in \mathbb{R}^3_+ \times \mathbb{R}. \end{aligned}$$

Su solución es

$$(x, \eta) = (0,\ 0,\ ,0,\ 98.6667).$$

Como $x \in \mathbb{Z}^3_+$, entonces $\eta^* = 98.6667$.

Ejemplo.

$$\begin{aligned} \max \quad & \eta \\ 11.6667x_1 + 4.3333x_2 + 3.3333x_3 + \eta &\leq 98.6667 \\ 7x_1 - \qquad x_2 - \qquad 4x_3 + \eta &\leq 80 \\ (x, \eta) &\in \mathbb{Z}^3_+ \times \mathbb{R}. \end{aligned}$$

Inicialmente se resuelve el problema de OL

$$\begin{aligned} \max \quad & \eta \\ 11.6667x_1 + 4.3333x_2 + 3.3333x_3 + \eta &\leq 98.6667 \\ 7x_1 - \qquad x_2 - \qquad 4x_3 + \eta &\leq 80 \\ (x, \eta) &\in \mathbb{R}^3_+ \times \mathbb{R}. \end{aligned}$$

Su solución es

$$(x, \eta) = (0,\ 0,\ 2.5455,\ 90.1818).$$

Como x no es entero, se resuelve el problema de factibildad entero con $\eta = 90.1818$

$$\begin{aligned} 11.6667x_1 + 4.3333x_2 + 3.3333x_3 &\leq \quad 8.4848 \\ 7x_1 - \qquad x_2 - \qquad 4x_3 &\leq -10.1818 \\ x &\in \mathbb{Z}^3_+\,. \end{aligned}$$

No hay puntos factibles, entonces ya se tiene $\beta_0 = 90.1818$.

Con $\tau_1 = 89.6818$ (términos independientes: 8.9848 y -9.6818) tampoco hay solución.

Con $\tau_2 = 88.6818$ (términos independientes: 9.9848 y -8.6818) tampoco hay solución.

Con $\tau_3 = 86.6818$ (términos independientes: 11.9848 y -6.6818), se obtiene $x = (0, 0, 2) \in \mathcal{F}_{\tau_3}$. Entonces $\alpha_0 = 86.6818$ y $\beta_0 = 88.6818$. Estos valores sirven para empezar la bisección, pero se puede mejorar α_0 para hacer el intervalo más pequeño.

El vector de holguras es $s = (5.3182,\ 1.3182)$, $\sigma = 88$ y $\alpha_0 = 88$. Para este valor el punto $x = (0, 0, 2)$ sigue siendo factible.

En la siguiente tabla están algunos resultados del proceso de bisección. El símbolo $\surd$ indica que, para $\eta = \mu_t$, el conjunto factible es no vacío.

t	α_t	β_t	μ_t		x_1	x_2	x_3	μ'_t
0	88.0000	88.6818	88.3409	$\surd$	0	0	3	88.6667
1	88.6667	88.6818	88.6742	$\emptyset$				
2	88.6667	88.6742	88.6705	$\emptyset$				
3	88.6667	88.6705	88.6686	$\emptyset$				
4	88.6667	88.6686	88.6676	$\emptyset$				

El vector x es el último obtenido, $x = (0, 0, 3)$, y $\eta \approx 88.6667$. En realidad $\eta = 88.6666...$

Para ilustrar mejor el aumento de μ_t para anular una holgura, con los mismos datos del ejemplo anterior, supongamos que el intervalo inicial es $\alpha_0 = 40.1818$ y $\beta_0 = 90.1818$, con el punto entero inicial $x = (0, 0, 1)$.

El vector de holguras es $s = (55.1515,\ 43.8182)$, $\sigma = 84$ y $\alpha_0 = 84$. Para este valor el punto $x = (0, 0, 1)$ sigue siendo factible.

t	α_t	β_t	μ_t		x_1	x_2	x_3	μ'_t
0	84.0000	90.1818	87.0909	$\surd$	0	0	2	88.0000
1	88.0000	90.1818	89.0909	$\emptyset$				
2	88.0000	89.0909	88.5455	$\surd$	0	0	3	88.6667
3	88.6667	89.0909	88.8788	$\emptyset$				
4	88.6667	88.8788	88.7727	$\emptyset$				
5	88.6667	88.7727	88.7197	$\emptyset$				
6	88.6667	88.7197	88.6932	$\emptyset$				
7	88.6667	88.6932	88.6799	$\emptyset$				
8	88.6667	88.6799	88.6733	$\emptyset$				
9	88.6667	88.6733	88.6700	$\emptyset$				
10	88.6667	88.6700	88.6683	$\emptyset$				
11	88.6667	88.6683	88.6675	$\emptyset$				

El vector x es el último obtenido, $x = (0, 0, 3)$, y $\eta \approx 88.6667$.