INTRODUCCIÓN AL MÉTODO DE PLANOS DE CORTE Y CENTRO ANALÍTICO (ACCPM) PARA LA SOLUCIÓN DE PROBLEMAS DE OPTIMIZACIÓN NO DIFERENCIABLE

INTRODUCTION TO ANALYTIC CENTER CUTTING PLANE METHOD (ACCPM) BY THE SOLUTION OF NONDIFFERENTIABLE OPTIMIZATION PROBLEMS

Héctor Andrés, López Ospina

Matemático, Msc. Matemática Aplicada, Profesor Auxiliar, Facultad de Ingeniería, Universidad de la Sabana, Chía, Colombia, hector.lopez1@unisabana.edu.co

Fecha de recepción: Abril 11 de 2007 Fecha de aprobación: Agosto 21 de 2007

RESUMEN

Este trabajo tiene por objetivo presentar una descripción del método de planos de corte que se basa en el cálculo del centro analítico conocido como ACCPM (Analytic Center Cutting Plane Method). Se presentan las generalidades de los métodos de planos de corte definiendo el concepto de oráculo y programa principal, y las subrutinas que dichos procesos contienen, tales como los cortes de factibilidad y optimalidad, conjunto de localización y centro analítico. Finalmente se desarrollan dos ejemplos que tienen por objetivo aclarar la teoría presentada.

PALABRAS CLAVES: Optimización convexa, ACCPM, métodos numéricos, Planos de corte.

ABSTRACT

This work has like objective to present a description of the analytic to center cutting plane method- ACCPM. The concepts of oracle and main program are defined and the subroutines that these processes contain, such as the cuts of feasibility and optimality, location set and analytic center. Finally, two examples are developed.

KEY WORDS: Convex optimization, ACCPM, numerical methods, Cutting planes.

INTRODUCCIÓN

El método de planos de corte basado en el cálculo del centro analítico (ACCPM) fue creado por Goffin, Vial y Haurie en 1992 [2] y tiene por objetivo resolver problemas de optimización que pueden ser formulados de la siguiente manera

$$\min f(y)
y \in Y,
l \le y \le h,$$
(1)

donde $f: \Re^n \to \Re$ es una función convexa, Y es un subconjunto convexo de \Re^n y l y h son dos vectores que representan restricciones de caja para las variables de decisión.

1. MÉTODOS DE PLANOS DE CORTE

En principio los algoritmos de planos de corte fueron creados con el objetivo de solucionar problemas de programación matemática que fuesen descritos así:

$$\min c^T x \\ x \in X, \tag{2}$$

donde $X \subseteq \Re^{n+1}$, es un conjunto convexo y acotado. Estos métodos construyen una aproximación lineal del problema anterior "mejorándola" en cada iteración. Sea P_0 una aproximación poliédrica de x ($P_0 \subseteq x$) y x_0 el punto óptimo de la función c^Tx en P_0 , la formulación general de un algoritmo de planos corte para resolver el problema anterior es:

El problema del algoritmo anterior es el hecho de tener que solucionar un problema de programación lineal en cada iteración, por lo cual los algoritmos más recientes de métodos de planos de corte difieren del anterior, principalmente por la manera de seleccionar el nuevo punto x^{k+1} , siendo éste el aspecto de mayor importancia ya que entre mejor sea el corte definido por x^{k+1} , más rápido convergerá el algoritmo.

1.1. MÉTODOS BASADOS EN CENTROS

Entre los métodos de planos de corte, se encuentran los métodos basados en centros. Estos métodos encuentran el valor de x^{k+1} por medio del cálculo del centro de un conjunto convexo y compacto llamado conjunto de localización L. El conjunto de localización esta formado por la intersección de los semiespacios generados por la aproximación lineal de la región factible y por una cota superior de la función objetivo. Dicho conjunto se puede expresar como un poliedro acotado:

$$L = \{x : Ax \le b, c^t x \le \overline{z}\}$$

Por la definición de L, se tiene que dicho conjunto contiene la solución óptima del problema (2). Los métodos basados en centros difieren en la forma de definir el punto central del conjunto de localización, entre los más conocidos se encuentran: el método del centro de gravedad, método volumétrico y el ACCPM, [1] y [4].

2. EL ORÁCULO Y LOS PLANOS DE CORTE

Los métodos de planos de corte se basan en la interacción de dos procedimientos como se describe a continuación: el oráculo y el programa principal.

El programa principal trabaja sobre una relajación lineal de la región factible del problema original de optimización (1), calculando en cada iteración del método un nuevo punto central. Además, controla la convergencia del proceso.

El oráculo es un procedimiento o algoritmo que tiene como dato de entrada el punto central obtenido y retorna uno o varios planos de corte al programa principal. Estos planos son de dos tipos: cortes de optimalidad o cortes de factibilidad, dependiendo de la naturaleza del punto. La siguiente figura explica lo descrito anteriormente.

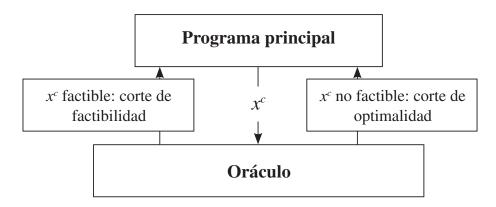


Figura 1. Fases de los métodos de planos de corte

2.1. CORTES DE FACTIBILIDAD

Dado un punto y^j enviado al oráculo, se dice que no es factible para el problema (1) cuando $y^j \notin Y$ pero cumple con las restricciones de caja. El oráculo produce entonces un corte de factibilidad.

Un corte de factibilidad es un semiespacio, definido por un hiperplano de separación $H := \{y : w^T (y - y^j) + y = 0\}$, de tal forma que

$$Y \subset \{w^T(y - y^j) + \gamma \le 0\} \text{ y } y^j \in \{w^T(y - y^j) + \gamma > 0\}$$

Definamos $g(y; y^j) := w^T(y - y^j) + \gamma$. Debido a la convexidad de la región factible Y, es posible obtener una aproximación poliédrica \overline{Y} de Y:

$$Y \subset \overline{Y} := \{y : g(y; y^j) \le 0, y^j \in J\}$$

donde $J := \{y^s : y^s \notin Y, l \le y \le h, s \in [1, ..., k]\}$, es decir, el conjunto de puntos generados por el programa principal, que no pertenecen a la región factible Y, obtenidos hasta la k -ésima iteración.

Por lo tanto, un corte de factibilidad generado por el punto y^{j} se representa de la forma:

$$g(y; y^j) \le 0, \quad \forall y \in Y$$

En el caso que la región factible se defina de la siguiente manera:

$$Y = \{y : g(y) \le 0\}$$

Donde $g: \Re^n \to \Re$ es una función convexa, los cortes de factibilidad se definen por medio del subgradiente. Sea $y^j \in J$, $\xi \in \partial g(y^j)$, por la convexidad de la función g se tiene que:

$$g(y) \ge g(y^j) + \xi^T(y - y^j), \forall y \in Y$$

Obteniendo así el corte de factibilidad:

$$g(y^j) + \xi^T(y - y^j), \le 0$$

2.2. CORTES DE OPTIMALIDAD

Dado un punto factible y^j enviado del programa principal al oráculo generará un corte conocido como corte de optimalidad.

Por la convexidad de f, si ξ es un subgradiente de f en y^j se cumple la siguiente desigualdad:

$$f(y) \ge f(y^i) + \xi^T(y - y^i), \xi \in \partial f(y^i), \forall y \in Y, l \le y \le h.$$

 $\text{Definase la función: } \overline{f}(y) = \max_{\mathbf{y}^i \in I} \ \{ f(y^i) + \xi^T(\mathbf{y} - y^i), \, \xi \in \partial f(y^i) \},$

donde $I = \{y^s : y^s \in Y, l \le y^s \le h, s \in [1, ..., k]\}$, es decir, el conjunto de puntos generados por el programa principal que pertenecen a Y, hasta la k-ésima iteración.

La función \overline{f} es una aproximación lineal a trozos de la función objetivo. Si se omite la restricción en el problema (1) (y \in *Y*), este problema de optimización se escribe:

$$\min \overline{f}(y)$$

$$l \le y \le h$$

El anterior problema es una relajación del problema original y es equivalente a:

$$\frac{\min z}{f(y) - z \le 0}$$

$$l \le y \le h$$

Tomando $f(y; y^i) := f(y^i) + \xi^T(y - y^i), \xi \in \partial f(y^i)$ y utilizando el conjunto I, problema descrito anteriormente se describe así

$$\min z$$

$$f(y; y^i) \le z, y^i \in I$$

$$l \le y \le h$$

Por lo tanto un corte de optimalidad se define por el siguiente semiespacio:

$$f(y; y^i) \le z$$

2.3. RELAJACIÓN LINEAL

Basándonos en los cortes de factibilidad y optimalidad podemos definir el siguiente problema de programación lineal, que es una relajación del problema de optimización (1):

$$\min \overline{f}(y)$$
$$y \in \overline{Y}$$
$$l \le y \le h$$

O bien, utilizando los conjuntos I, J

$$\min z$$

$$f(y; y^i) \le z, \forall y^i \in I$$

$$g(y; y^j) \le 0, \forall y^j \in j$$

$$l \le y \le h$$

De esta forma se trabajará con una función objetivo lineal (aproximación poliédrica de f) sobre una aproximación "lineal" de Y.

2.4. CONJUNTO DE LOCALIZACIÓN

Si $I \neq \phi$, es posible obtener una cota superior z de la solución óptima del problema (1) tomando:

$$\overline{z}_i := \min_{y^i \in I} \{ f(y^i) \}$$

Si se tiene en cuenta la unión de todos los cortes y desigualdades obtenidos anteriormente, se define el conjunto de localización L_z de la siguiente manera:

$$z \ge f(y^{j}) + \xi^{T}(y - y^{j}), \forall i \in I$$

$$0 \ge (w^{k})^{T}(y - y^{k}) + \gamma_{k}, \forall k \in j$$

$$\overline{z} \ge z$$

$$b \ge By$$

El primer conjunto de restricciones recibe el nombre de restricciones de optimalidad, luego se encuentran las restricciones de factibilidad. El tercer conjunto de restricciones define la cota superior del problema de optimización y por último, se describe de forma matricial las restricciones de caja. Es importante anotar que $L_{\overline{z}}$ contiene la solución del problema original de optimización.

Por otro lado, es posible asociar a cada conjunto de restricciones descritas anteriormente sus respectivas variables duales $\alpha_i \ge 0$, $\mu_k \ge 0$, $\nu \ge 0$, $\rho \in \Re$ que satisfacen la siguiente desigualdad:

$$z \ge \sum_{i \in I} \alpha_i (f(y^i) - \xi^T y^i) + \sum_{k \in J} \mu_k (\gamma_k - (w^k)^T y^k) + b^T y$$

Para todo (z, x) que pertenece al conjunto de localización. La expresión del lado derecho de la anterior expresión es una cota inferior del problema original de optimización (1). Dicha cota se nota \underline{z}_i . Dadas las cotas superior e inferior es posible definir una brecha o salto de dualidad: $d_{g,i} = \overline{z}_i - \underline{z}_i$. Este valor es muy importante debido a que con él se construye el criterio de parada. El método se detiene cuando $d_{g,i} \le \theta$.

2.5. CENTRO ANALÍTICO

Tomando x = (y, z), el conjunto de localización se puede escribir de la siguiente forma:

$$L_z = \{x \in \Re^{n+1} : A^T x \le b\}$$

El conjunto anterior es un poliedro acotado y para dicho conjunto es posible definir su centro analítico como el único punto que soluciona el siguiente problema:

$$\arg\min\left\{-\sum_{i=1}^{n+1}\log\left(s_{i}\right): s=b-A^{T}x\right\}$$

En la ecuación anterior, se penalizan los puntos cercanos a la frontera, es decir, las variables de holgura (s_i) que tiendan a cero. Por otro lado, el centro analítico es el único punto solución a un sistema de ecuaciones no lineales (condiciones de optimalidad de primer orden).

$$A^{T}x + s = b$$
$$Au = 0$$
$$Us = 1$$

Donde *U* es la matriz diagonal correspondiente al vector de variables duales *u*. También es posible definir el centro analítico del poliedro dual.

$$\{Au = 0, u \ge 0\}$$

Dicho centro analítico es la única solución al siguiente problema:

$$\arg \min \left\{ b^{T} u - \sum_{k=1}^{k} \log (u_{k}) : Au = 0, \ u \ge 0 \right\}$$

El cálculo del centro analítico se basa en métodos de punto interior y los métodos más conocidos son:

- Método primal: encuentra las direcciones con el polítopo acotado generado por el conjunto de localización.
- Método Dual: trabaja con el poliedro de variables duales.
- Método primal-dual: combinación de los métodos anteriores
- Método primal proyectivo: encuentra las direcciones por medio de proyecciones de las variables generadas por el conjunto de localización.

Una descripción detallada de dichos métodos y su convergencia se encuentra en [4] y en [6].

2.6. ALGORITMO ACCPM

El algoritmo ACCPM se formula de manera general de la siguiente forma:

Igoritmo ACCPM

Entrada:

Definir el nivel de tolerancia para la función objetivo: $\theta > 0$ Definir las restricciones de caja: $l, h \in \Re^n$

Inicialización:

Calcular \bar{z} y \underline{z} ;

Conjunto de localización inicial: $L_{\bar{z}} = \{(y, z) \in \Re^{n+1} : z \le z \le \bar{z}, l \le y \le h\}$

Calcular el centro analítico inicial de $L_{\overline{z}}$: $(y^c, z^c) = \left(\frac{1+h}{2}, \frac{\overline{z}+\underline{z}}{2}\right)$

Mientras $\overline{z} - \underline{z} \ge \theta$ hacer

Oráculo

Si $y^c \in Y$ entonces Añadir corte de optimalidad: $z \ge f(y; y^c)$; $\overline{z} = \min \{\overline{z}, f(y^c)\}$ Sino

Añadir corte de factibilidad: $g(y; y^c) \le 0$ fin si

Programa principal

Actualizar el conjunto de localización; Encontrar el nuevo centro analítico Calcular una cota inferior z_0 ; $\underline{z} = \max \{z_0, \underline{z}\}$

fin mientras

Existe una biblioteca del método desarrollado por Péton y Víal en C++ [5]. Dependiendo de la aplicación que se desee programar, el usuario debe implementar el oráculo de su respectivo problema. El análisis de convergencia del algoritmo ACCPM se encuentra en [1].

3. EJEMPLOS

Existen una gran cantidad de problemas de optimización que han sido resueltos por medio ACCPM [1], [3]. A continuación se presentan dos ejemplos sencillos que tienen como objetivo aclarar la utilización del método.

3.1. UN PROBLEMA DE OPTIMIZACIÓN l_p

El problema a trabajar, se representa matemáticamente de la siguiente manera:

$$\min \|y - y^1\|.$$
s.a. $\|y - y^2\| \le r$

A continuación se presenta una forma sencilla de programar el oráculo:

```
Oráculo
\xi = 0, (Definición del subgradiente inicial por medio del vector 0)
        Si y \in Y entonces
           Corte de optimalidad
               Si |y_i - y_i| = \max \{y_j - y_j|: j = 1, ..., n\}
                   Caso 1: (y_i - y_i^1 \ge 0)
                   \xi = \xi + e_i (donde e_i es el vector de ceros con 1 en la i-ésima posición)
                   Caso 2: (y_i - y_i^{\dagger} < 0)
                   \xi = \xi - e_i
                Fin si
             Corte: f(y) + \xi^T y \le z
        Sino
            Corte de factibilidad
               Para i = 1, ..., n hacer
                   Caso 1: (y_i - y_i^2 \ge 0)
                   \xi = \xi + e_i (donde e_i es el vector de ceros con 1 en la i-ésima posición)
                   Caso 2: (y_i - y_i^2 < 0)
                   \xi = \xi - e_i
                Fin Para
             Corte: g(y) + \xi^T y \le z
      Fin sino
```

Tomando $y' = (4,2), y^2 = (1,1), r = 1, \bar{z} = 10, \underline{z} = 0$ y las restricciones de caja K = [-3,3] X [-3,3]. El problema de optimización quedaría expresado de la siguiente manera:

$$\min f(y) = ||y - (4,2)||.$$

$$s.a.g(y) = ||y - (1,1)||_1 - 1 \le 0$$

$$y \in K$$

La región factible es $Y = \{y \in \Re^2: g(y) \le 0.$

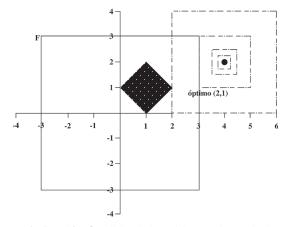


Figura 2. Región factible del problema de optimización

El primer centro analítico es (y,z)=(0,0,5).

Primera iteración:

Como $y^l = (0,0) \notin Y$ entonces el oráculo retorna un corte de factibilidad, así:

$$g(y; y^{1}) = 1 + (-1 - 1) \begin{pmatrix} y_{1} - 1 \\ y_{2} - 1 \end{pmatrix} \le 0$$
$$g(y; y^{1}) = 1 - y_{1} - y_{2} \le 0$$

Luego, el primer conjunto de localización obtenido es:

$$L_{\bar{z}}^{1} = \{ (y_{1}, y_{2}, z) \in \Re^{3} : (y_{1}, y_{2}) \in K, z \leq \bar{z}, 1 - y_{1} - y_{2} \leq 0 \}$$

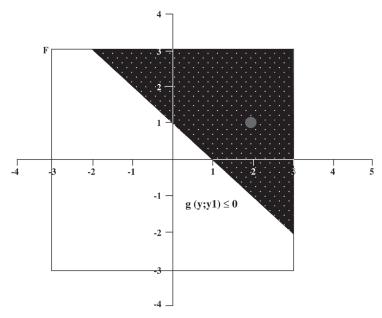


Figura 3. Primer conjunto de localización y su centro analítico

El centro analítico de L_z^1 es (1.333, 1.333,5). Además z = 0

Segunda iteración:

Como $y^2 = (1.333, 1.333)$ es un punto factible, se obtiene un corte de optimalidad; Es decir, el oráculo retorna el siguiente corte:

$$f(y; y^2) \le z$$

$$f(y; y^2) = f(y^2) + (-1 \ 0) \left(\frac{y^1 - 4}{y^2 - 2} \right) \le z$$

$$f(y; y^2) = 4 - y_1 \le z$$

Con lo anterior $\overline{z} = 2.667$.

Actualizando el conjunto de localización se obtiene que:

$$L_{\overline{z}}^2 = L_{\overline{z}}^1 \cap \{(y_1, y_2, z) \in \Re^3 : 4 - y_1 \le z, z \le 2.667\}$$

Como el último corte hallado no se puede graficar en \Re^2 , se hace una aproximación notando que $z \le \overline{z} = 2.6667$. Es decir, la gráfica del corte corresponde a una curva de nivel en z = 2.6667.

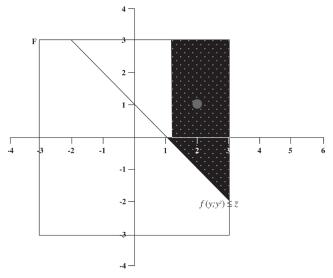


Figura 4. Segundo conjunto de localización y su centro analítico

El centro analítico de L_z^2 es (2.445,1.188,1.33) y $\underline{z} = 0$.

Tercera iteración:

En este caso $y^3 = (2.445, 1.188)$ no es factible, así pues el oráculo retorna un corte de factibilidad, de la siguiente manera:

$$g(y; y^3) = y_1 + y_2 - 3 \le 0$$

Actualizando el conjunto de localización se obtiene:

$$L_z^{\frac{3}{2}} = L_z^{\frac{2}{2}} \cap \{ (y_1, y_2, z) \in \Re^3 : y_1 + y_2 - 3 \le 0 \}$$

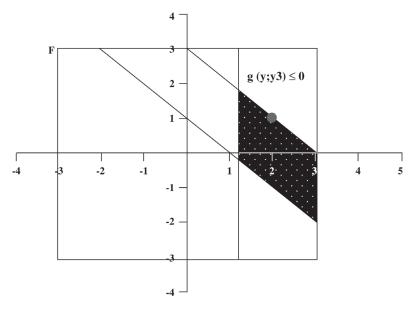


Figura 5. Tercer conjunto de localización y su centro analítico

El centro analítico de L_z^3 es (2,0,1.333).

Cuarta iteración:

 $(2,0) \notin Y$, luego se obtiene el siguiente corte de factibilidad:

$$y_1 - y_2 \le 1$$

Actualizando el conjunto de localización:

$$L_{\overline{z}}^{4} = L_{\overline{z}}^{3} \cap \{(y_{1}, y_{2}, z) \in \Re^{3} : y_{1} - y_{2} \le 1\}$$

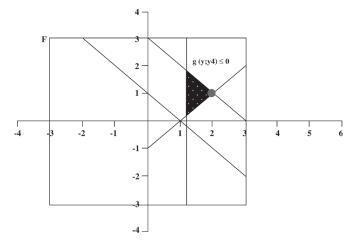


Figura 6. Primer conjunto de localización y su centro analítico

Del gráfico anterior se deduce que los próximos cortes serán de optimalidad. Después de varias iteraciones se obtiene el óptimo y = (2,1) y z = 2.

3.1.1 Problemas de localización de una planta (problema de Fermat)

Estos problemas de localización buscan encontrar un punto $y \in \Re^n$ tal que la siguiente función f(y) tenga valor mínimo:

$$f(y) = \sum_{i=1}^{m} ||y - y^{i}||$$

Donde $y^i \in \Re^n$ son m puntos dados.

Para este ejemplo sea m=3 y n=2. Los tres puntos son: $y^1=(0,0), y^2=(6,1), y^3=(4,2)$ que forman el siguiente triángulo:

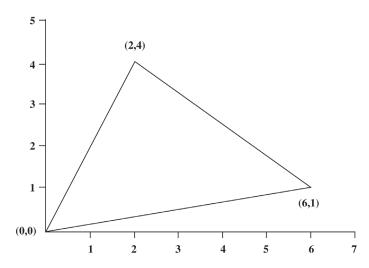


Figura 7. Región factible para el problema de Fermat

Como dicho problema de optimización no tiene restricciones naturales, basta con tomar una caja adecuada. Obviamente, la solución se encuentra en el interior del triángulo formado por los puntos. En este caso las restricciones de caja podrían ser $K = [0,6] \times [0,6]$ que a la vez será la región factible: luego así, solo se obtendrán cortes de optimalidad.

Se puede tomar como cotas los siguientes valores:

$$\underline{z} = \max \{ \|y^{j} - y^{i}\| : i, j = 1, 2, 3 \} \text{ y}$$

$$\overline{z} = \min_{1 \le i \le 3} \{ \sum_{j=1, j \ne i}^{3} \|y^{i} - y^{j}\| \}$$

Así pues z = 6.1 y $\overline{z} = 9.5$.

Un subgradiente para la función objetivo puede ser:

$$\xi = \sum_{i=1}^{3} \xi^{i},$$

donde,

$$\xi^{i} = \begin{cases} \frac{(y^{j} - y^{i})}{\|y^{j} - y^{i}\|}, & si \ y \neq y^{i} \\ 0, & si \ y = y^{i} \end{cases}$$

Cuando m=3 y n=2, este problema es conocido como problema de Fermat (ver [4]). En este caso, la solución puede ser descrita de manera precisa. La solución óptima esta dada por el punto interior al triángulo, tal que el ángulo formado entre este punto y las rectas que lo unen con dos extremos del triángulo es 120°. De esta manera se obtiene el óptimo, $y^* = (2.39401, 2.16395)$ y $z^* = 8,8941058$.

Como la región factible Y = K, entonces el primer centro analítico es (3, 3, 7.8).

Primera iteración

$$y^{1} = (3,3)$$
 genera el siguiente corte $f(y;y^{1}) = 9.262 + (3.325 \ 18.525) \begin{pmatrix} y_{1} - 3 \\ y_{2} - 3 \end{pmatrix} \le z$. Es decir $f(y;y^{1}) = 3.32544818y_{1} + 18.524811y_{2} - 55.574433 \le z$.

Además, $\overline{z} = 9.262$. Este valor aproxima gráficamente el siguiente corte:

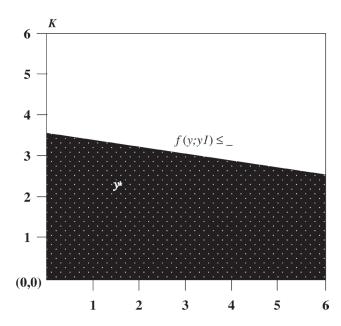


Figura 8. Primer conjunto de localización y su centro analítico

El primer conjunto de localización es:

$$L^1_{\bar{z}} = \{(y_1, y_2, z) \in \Re^3 : z \le 9.262, 3.325y_1 + 18.525y_2 - 55.574 \le z\}$$

El centro analítico de $L_{\bar{z}}^1$ es (4.874, 1.901, 6.947) y \underline{z} = 6.1.

Segunda iteración:

 $y^1 = (4.874, 1.901)$ genera el siguiente corte $34.109y_1 + 3.780y_2 - 163.213 \le z$ y $\bar{z} = 9.262$. Obteniendo la siguiente aproximación gráfica:

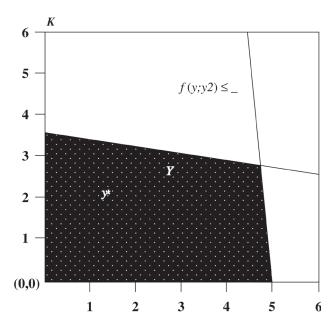


Figura 9. Segundo conjunto de localización y su centro analítico

El nuevo conjunto de localización es

$$L_{\bar{z}}^2 = L_{\bar{z}}^1 \cap \{(y_1, y_2, z) \in \Re^3 : 34.109y_1 + 3.78y_2 - 163.213 \le z\}$$

El centro analítico de $L_{\bar{z}}^2$ es (1.69, 3.11, 9.214).

REFERENCIAS BILBLIOGRÁFICAS

[1] DU MERLE, O.; Points intérieurs et plans coupants: mise en uvre et développement d'une méthode pour l'optimisation convexe et la programmation linéaire structureée de grand taille; Suiza, 1995. Tesis de doctorado, Universidad de Ginebra.

- [2] GOFFIN, J.L.; HAURIE A.; VIAL J.; (1992), Decomposition and nondiferentiable optimization with the projective algorithm, En Management Science, 38-2, 284-302.
- [3] GOFFIN, J.L.; VIAL, J; (1999), Convex Nondifferentiable Optimization: A survey focused the analytic center cutting plane method. Reporte técnico de logilab. Universidad de Ginebra. 1999
- [4] LÓPEZ, H.; *Método de planos de corte y centro analítico para optimización no diferenciable*. Bogotá, 2002, 65 p. Trabajo de grado (Matemático). Universidad Nacional de Colombia. Facultad de Ciencias.
- [5] PETÓN O., Vial J.P. (2001), *A tutorial of ACCPM. Version 2.01*, Reporte Técnico, HEC/logilab. Universidad de Ginebra. 2001.
- [6] VIAL J.P.; (1998), *Analytic center of polytope*, Manuscrito, Universidad de Ginebra, Suiza.