UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA

FACULTAD DE INGENIERÍA

“Optimización global: Una introducción a la minimización cóncava”

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN INGENIERÍA

INGENIERÍA DE SISTEMAS – INVESTIGACIÓN DE OPERACIONES

P R E S E N T A :

CLAUDIA ORQUÍDEA LÓPEZ SOTO

TUTOR:

DRA. IDALIA FLORES DE LA MOTA

2008

JURADO ASIGNADO:

Presidente: Dr. Servio Tulio Guillen Burguete

Secretario: Dr. Juan Manuel Estrada Medina

Vocal: Dra. Idalia Flores De la Mota

1er. Suplente: M. en I. O. Ma. Del Carmen Hernández Ayuso

2do. Suplente: Dr. Javier Ramírez Rodríguez

Lugar o lugares donde se realizó la tesis:

FACULTAD DE INGENIERÍA

TUTOR DE TESIS:

Dra. Idalia Flores De la Mota

____

FIRMA

Optimización global: Una introducción a la minimización cóncava

Claudia Orquídea López Soto

En memoria de
Raúl Soto y Soto

Agradecimientos

A mi abuelita *por darme su amor, su apoyo en todos mis proyectos y por que a ella le debo lo que soy.*
A mi mamá *por darme la vida y por estar siempre presente demostrándome su cariño.*
A Carlos y mis hermanos*: Alejandro, Arantxa y Andrea por ser parte de mi vida y hacerla más plena.*
A mis tíos *Blanca y Paco por su cariño, apoyo y cuidados recibidos desde niña.*
A Tess *por su cariño aunque físicamente se encuentre lejos.*

A mis amigos y entrañables compañeros de la maestría: ***Abigail, Alejandra y Daniel*** *por su amistad, su confianza y por todos los momentos compartidos.*
A Germán *por su tiempo, su amistad y colaboración en los detalles finales del presente trabajo.*

A los miembros de mi jurado*: Dr. Servio Tulio Guillen Burguete, Dr. Juan Manuel Estrada Medina, Dra. Idalia Flores De la Mota, M. en I.O. Ma. del Carmen Hernández Ayuso, Dr. Javier Ramírez Rodríguez por sus correcciones e imprescindibles observaciones.*

En especial a la Dra. Idalia por su tiempo y confianza durante la realización de este trabajo.
A Carmen Hernández Ayuso por su amistad, su tiempo y apoyo incondicional.

A CONACYT *por la beca que me fue otorgada para la realización de mis estudios de maestría.*

Por último y no por eso menos importante a ***Carlos Erwin*** *por tenerme paciencia, por impulsarme a seguir adelante, por estar a mi lado en todo momento y sobre todo por su AMOR y comprensión. Te amo pecoso!*

Claudia Orquídea López Soto

Contenido

Resumen

Se estudiaron algunos de los diferentes algoritmos que son usados para obtener la solución global de problemas de minimización cóncava. Dichos métodos se clasifican de acuerdo a su forma de proceder y éstos son: Métodos de corte, Métodos de aproximación sucesiva y Métodos de partición sucesiva. Los algoritmos presentados son los desarrollados por Reiner Horst y Hoang Tuy principalmente. Cada algoritmo está acompañado de un ejemplo que muestra detalladamente su respectivo procedimiento. Se presenta una aplicación al algoritmo de aproximación exterior (metodología que se encuentra dentro de los llamados métodos de aproximación sucesiva) a través del problema de extracción de agua de un acuífero, estudiado y presentado por Karatzas y Pinder. Como otra aplicación, se muestra la implementación del algoritmo de ramificación y acotamiento (metodología que se encuentra dentro de los llamados métodos de partición sucesiva) a un problema de flujo en redes con costos cóncavos desarrollado por Fontes, Hadjiconstantinou y Christofides.

Abstract

Some of the different algorithms which are used to obtein the global solution to problems of concave minimization were studied. These methods are clasified acording to their way to work and these are: Cutting methods, succesive approximation methods and succesive partition methods. The presented algorithms are mainly developed by Reiner Horst and Hoang Tuy. Each algorithm has an example which shows in detail its respective procedure. An application to the outer approximation algorithm (methodology which is considered a succesive approximation method) is presented using the problem of extracting water from an aquifer studied by Karatzas and Pinder. As an application to the branch and bound algorithm (methodology which is considered a succesive partition method) is presented throgh the problem of network flows with concave costs developed by Fontes, Hadjiconstantinou y Christofides.

Introducción

Muchos son los problemas en los cuales hace su aparición la optimización global; modelos económicos, cargos fijos, finanzas, redes y transportes, bases de datos, diseño de chips, procesamiento de imágenes, diseño mecánico y nuclear, diseño y control en ingeniería química, biología molecular e ingeniería ambiental [7]. El principal objetivo de la optimización global es encontrar el mínimo (máximo) global del problema, esto puede resultar difícil si tratamos de usar las técnicas de programación no lineal, pues sólo encontraremos óptimos locales, que no necesariamente son globales. Es por esto que se hace imprescindible el desarrollo de técnicas que nos ayuden a encontrar la solución óptima global. En este trabajo se estudiará una rama de la optimización global conocida como minimización cóncava y los métodos que estudiaremos son: los métodos de corte, los métodos de aproximación sucesiva y los métodos de partición sucesiva.

Los métodos de corte son una herramienta que tienen su origen en la solución de problemas de programación lineal, fueron introducidos por Gomory en la década de los sesentas, pues buscaba la solución entera de un problema lineal, más adelante se adaptarían para encontrar solución a problemas de optimización convexa. Pero es Hoang Tuy uno de los primeros en establecer problemas de programación cóncava global, de hecho el construye lo que se conoce como el corte de Tuy. La idea de estos métodos es construir cortes (restricciones) que excluyan una parte de la región factible en donde sabemos no se encuentra la solución óptima.

Los métodos de aproximación sucesiva reciben dicho nombre puesto que se caracterizan por construir a través de su desarrollo una sucesión de soluciones que gradualmente se aproximan a la solución del problema original. Estos problemas en general son más sencillos de resolver que el original, usualmente son resueltos secuencialmente hasta que es encontrada una solución exacta o una aproximación óptima global [3].

Los métodos ramificación y acotamiento son los más conocidos sobre todo en el área de programación entera; consisten básicamente en la búsqueda de puntos que minimizan (maximizan según sea el caso) a la función objetivo eliminando de manera adecuada regiones del espacio factible que resultan ser poco prometedoras. Este método fue propuesto en 1960 por A.H. Land y A.G. Doing.

El algoritmo a grandes rasgos consiste de dos procesos, el primero es el conocido como *Ramificación*: proceso mediante el cual se divide la región factible en subregiones, como el algoritmo es recursivo, a partir de cada región dividida es posible seguir aplicándole el método de tal manera que en cada iteración se construye un árbol de posibles soluciones, el segundo proceso es el llamado *Acotamiento*: herramienta que genera cotas inferiores y superiores para probar si se ha llegado a la solución óptima dentro de una subregión factible, este proceso evita que se realice una búsqueda exhaustiva.

El presente trabajo tiene por objetivo dar un panorama general de los métodos más comunes que se usan en la solución de problemas de minimización cóncava, para cada algoritmo desarrollado se presentan ejemplos detallados para la mejor comprensión del mismo. Constan de cinco capítulos y un apéndice, el cual contiene definiciones y resultados importantes.

En el Capítulo 1 se presentan el problema de minimización cóncava. Se dan algunas aplicaciones que tiene el problema a diferentes tipos de problemas de optimización, así como algunas aplicaciones que se consideran indirectas a otras áreas de optimización como son la programación entera, programación bilineal. Así mismo, se presentan algunos de los artículos más recientes que se han desarrollado en los cuales se utiliza alguno de los métodos que se estudian en este trabajo.

En el Capítulo 2 se estudian los métodos de corte, comenzamos por establecer la teoría que nos ayuda a construir el algoritmo general de cortes para después estudiar los algoritmos de corte puro, el algoritmo de corte de cara y finalmente el algoritmo de corte y división.

El Capítulo 3 consta básicamente de dos secciones, en la primera se aborda el problema de aproximación exterior, es presentado el algoritmo en su versión más general, posteriormente se desarrolla el algoritmo de aproximación exterior para problemas con restricciones lineales; en la segunda sección se estudia el método de aproximación interior, el cuál se construye a través de la ayuda de otro problema denotado como (DG) y se desarrolla el algoritmo de anexión poliedral.

El Capítulo 4 está formado de cuatro secciones. En la primer sección se presenta al algoritmo de ramificación y acotamiento en su forma general, en las siguientes tres secciones se desarrolla un algoritmo ligeramente diferente de los otros. En la segunda sección se estudia el algoritmo cónico, en la tercer sección se estudia al algoritmo simplicial y finalmente en la última sección se estudia al algoritmo rectangular.

En el Capitulo 5 se presenta dos aplicaciones de algunos de los métodos estudiado a lo largo de todo el trabajo. La primer aplicación que se ve es al problema de extracción de agua de un acuífero basados en [9]. El problema consiste en determinar el número de pozos, su ubicación y tasa de bombeo óptima al menor costo posible, para lo cual, los autores del artículo en cuestión usan el método de aproximación exterior. La segunda aplicación que se presenta es al área de teoría de redes, pues en el artículo escrito por[6] desarrollan la implementación del algoritmo de ramificación y acotamiento para el problema de flujo en redes con costos cóncavos donde existe un único nodo fuente sin límite de capacidad.

1

Antecedentes

La minimización cóncava forma parte de la optimización global. La minimización cóncava tiene aplicaciones en diversas áreas del conocimiento tales como economía, ingeniería, administración de negocios, física, etc. El objetivo principal de la optimización global es encontrar la mejor solución posible a problemas que pueden ser expresados a través de modelos matemáticos. Los modelos matemáticos que se presentan en el área de optimización constan de una función objetivo y la mayor parte de las veces dicha función se encuentra sujeta a ciertas restricciones. Los problemas de minimización cóncava que se resuelven a través de los métodos que veremos se denotan del siguiente modo:

$$\begin{array}{ll} \textit{glob}\,\text{mín} & f(x) \\ \text{s.a} & \\ & x \in D \end{array} \qquad \text{(P)}$$

donde $D \subseteq \mathbb{R}^n$ es un conjunto no vacío, convexo cerrado y $f(x)$ es una función real cóncava definida sobre un conjunto $A \subseteq \mathbb{R}^n$ que contiene a D. El objetivo del problema (P) es encontrar el valor mínimo que la función f puede alcanzar sobre D, si es que éste existe.

Es muy común que al conjunto de restricciones que generan a D se defina como un sistema de igualdades o desigualdades lineales o no lineales (este trabajo sólo nos enfocaremos en la ecuaciones o inecuaciones lineales). De aquí, que podemos suponer que D tiene la siguiente forma:

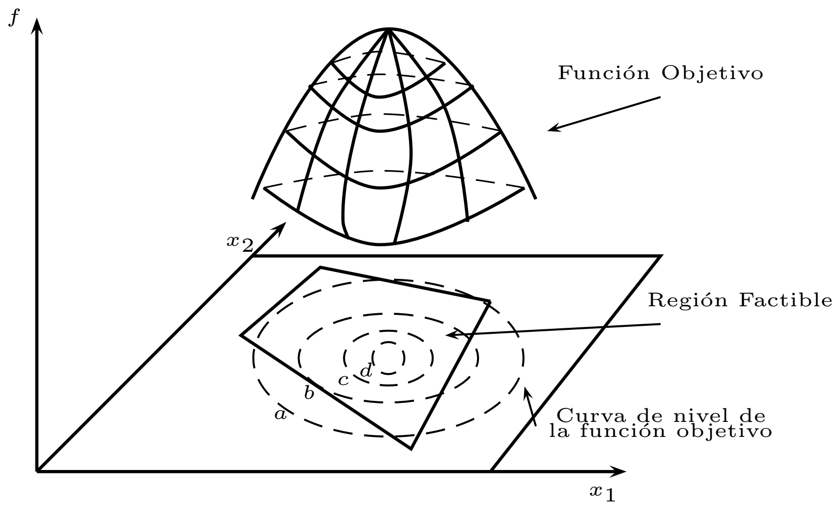
$$D = \{x \in A | g_i(x) \leq 0, \quad i = 1, 2, ..., m\}, \tag{1.1}$$

donde para cada $i = 1, 2, ..., m$, $g_i : A \to \mathbb{R}^n$. Observemos que si multiplicamos ambos lados de la desigualdad anterior por -1, es igualmente válido suponer que D toma la siguiente forma

$$D = \{x \in A | g_i(x) \geq 0, \quad i = 1, 2, ..., m\}, \tag{1.2}$$

donde $g_i : A \to \mathbb{R}^n$, $i = 1, 2, ..., m$.

La siguiente gráfica es la representación geométrica del tipo de problemas que deseamos resolver; la región de soluciones factibles es un conjunto convexo y la función objetivo es cóncava.



donde a, b, c, d son números reales tales que $a < b < c < d$

Encontrar solución a este tipo de planteamientos no siempre resulta en tarea sencilla, para hacerlo debemos tomar en cuenta el tipo de función que representa la función objetivo y las restricciones, ya que si al menos uno de estos elementos es de tipo no lineal, hay diferentes algoritmos que les dan solución. Lo que veremos aquí está enfocado al estudio de optimización global a través de métodos deterministas considerados ya clásicos tales como; "‘métodos de corte"’ y "‘métodos de ramificación y acotamiento"’, los cuales serán usados como una herramienta en los métodos principales: *Métodos de enumeración de puntos extremos, Métodos de aproximación sucesiva y Métodos de partición sucesiva.*

1.1. Propiedades básicas de los problemas de minimización cóncava

Recordemos que en el problema (P) supusimos que $D \subseteq \mathbb{R}^n$ es un conjunto no vacío, cerrado y convexo, sea f una función cóncava real [1] definida sobre un conjunto convexo $A \subseteq \mathbb{R}^n$ que contiene a D. A estos supuestos los llamaremos *supuestos básicos* del problema (P), los cuales supondremos que se satisfacen.

[1]Para conocer las definiciones, ver el apéndice

En general en los problemas de programación no lineal, D es un conjunto no vacío, por lo que matemáticamente cada instancia del problema P se encuentra dentro de una de las siguientes categorías:

1. Al menos una solución global existe.

2. El valor del mínimo global v es finito pero no hay ningún punto en D que lo satisfaga.

3. El problema es no acotado, es decir, $v = -\infty$.

Los problemas que caen en la segunda opción es debido a que; A no es un conjunto abierto y f es discontinua en al menos un punto de la frontera de D o bien, porque D es no acotado. En instancias prácticas del problema (P) esto no sucede. Para situaciones típicas, alguna de las propiedades matemáticas de los problemas de minimización cóncava son usados para mostrar que el mínimo global, si es finito, puede ser alcanzado. Para ciertas clases de problemas de optimización global, la búsqueda del mínimo global cuando existe, se centra en la frontera de la región factible.

Propiedad 1. *Supongamos que se satisfacen los supuestos básicos para el problema (P), además (D) contiene al menos un punto extremo. Entonces si el problema (P) tiene una solución óptima, tiene una solución óptima global la cual es un punto extremo de D.*

Observemos que esta propiedad es muy importante, pues justamente como en programación lineal, la búsqueda de la solución óptima para el problema (P) puede ser restringida al conjunto de puntos extremos de D, siempre que D sea no vacío. Veamos también, que D podría ser no acotado o bien f podría ser discontinua en la frontera de D. Particularmente esta propiedad es útil cuando D es un poliedro convexo, puesto que estos conjunto tienen un número finito de puntos extremos, aún cuando son no acotado. En general, nosotros estaremos trabajando sobre este tipo de conjuntos. Para poder usar esta propiedad lo primero que tenemos que verificar es que el problema tenga al menos una solución óptima global. Usando el teorema de Weierstrass [2] cuando f es continua sobre A y D es compacto entonces se cumple que hay al menos una solución óptima global. Usando el hecho de que un conjunto no vacío convexo compacto debe tener un punto extremo y la siguiente propiedad

Propiedad 2. *Si $A = \mathbb{R}^n$ o A es un subconjunto convexo abierto de $\mathbb{R}^n$, entonces, f es una función continua sobre A.*

es posible establecer la propiedad que a continuación se presenta.

[2]Ver Apéndice

Propiedad 3. *Supongamos que se satisfacen los supuestos básicos para el problema* (P)*, además supongamos que* A *es igual a* $\mathbb{R}^n$ *o a un conjunto convexo abierto en* $\mathbb{R}^n$ *y que* (D) *es acotado. Entonces el problema* (P) *debe tener una solución óptima global que es un punto extremo de* D*.*

Propiedad 4. *Supongamos que* (D) *es un poliedro posiblemente no acotado. Entonces si* f *es acotada inferiormente sobre* D*, el problema* (P) *tiene al menos una solución óptima global.*

Esta propiedad establece que una función cóncava acotada inferiormente sobre un poliedro siempre alcanzará su valor mínimo sobre el poliedro en algún punto de éste. Esto se satisface aún cuando la función sea discontinua sobre la frontera del poliedro o que el poliedro sea no acotado o ambas opciones. Podemos observar, que sucede una situación análoga en programación lineal, cuando D es un poliedro no vacío y P es no acotado, entonces debe haber una solución óptima global.

1.2. Aplicaciones directas

Debido a la variedad de problemas que pueden ser resueltos a través de minimización cóncava, ya sea de manera directa o bien mediante una transformación puedan ser expresados como tal, es que encontramos una gran cantidad de referencias que muestran que es un problema que ha sido objeto de estudio desde los años sesentas y continua su investigación.

Algunos problemas que se resuelven de manera directa son:

El problema de cargo fijo. Son problemas cuya función objetivo es de la forma $f(x) = \Sigma_{i=1}^{n} f_i(x_i)$ para cada $x \in A$, donde A es un conjunto abierto convexo, donde para cada $i = 1, ..., n$ f_i es una función cóncava de una sola variable x_i sobre un intervalo apropiado en R. Tales funciones son separables y en varias de las aplicaciones f_i es una función sobre $\{x_i \in R | x_i \geq 0\}$ de la forma

$$f_i(x_i) = \begin{cases} 0 & si\ x_i = 0 \\ a_i + w_i(x_i) & si\ x_i > 0 \end{cases}$$

donde $a_i > 0$ y w_i es una función continua cóncava sobre $\{x_i \in \mathbb{R} | x_i > 0\}$. Los problemas de cargo fijo se encargan de elegir niveles x_i con $i = 1, ..., n$ de n actividades que minimizan el costo total $f(x) = \Sigma_{i=1}^{n} f_i(x_i)$ de n actividades sujeto a que $x \in D \subset \{x \in \mathbb{R}^n | x > 0\}$. Para cada actividad i, $a_i > 0$ representa el costo fijo de la actividad y $w_i(x_i)$ es la variable de costo fijo de la actividad. Para cada i no se incurre en ningún costo si $x_i = 0$ pero en cuanto $x_i > 0$ el costo tiene un valor de a_i. Así, cada función f_i es discontinua en 0 e indefinida para actividades de nivel negativos. [3]

El siguiente es un ejemplo de la estructura que presenta un problema de cargo fijo: Una cierta compañía desea comprar un número específico de unidades de un cierto bien. Se cuenta con ofertas de n vendedores y cada uno de los cuales no puede cubrir la demanda total. Cada vendedor ofrece sus ofertas que indican los precios como una función de la cantidad comprada. El problema consiste en escoger una cantidad a comprar a cada vendedor de tal manera que se minimice el costo total de la compra. Ahora, en particular consideremos que se desean adquirir $150,000$ unidades de un bien, se cuenta con tres vendedores cuya ofertas se muestran en la siguiente tabla. Consideremos a x_i como el número de unidades que se ofrecen del bien por el vendedor $i = 1, 2, 3$.

vendedor	costo fijo	precio por unidad	cantidad de bienes por unidad
1	3,520.20	51.20	$0 < x_1 \leq 50,000$
2	82,810.00	52.10	$0 < x_2 \leq 20,000$
		51.10	$20,000 < x_2 \leq 60,000$
		50.10	$60,000 < x_2 \leq 80,000$
		49.10	$80,000 < x_2 \leq 100,000$
3	0	60.50	$0 < x_3 \leq 50,000$
		59.00	$50,000 < x_3 \leq 80,000$

Sea $f_i(x_i)$ el costo de x_i unidades ofrecidas por el vendedor i, con $i = 1, 2, 3$. Entonces a partir de los datos que tenemos podemos construir el modelo matemático que corresponde al problema de cargo fijo:

$$f_1(x_1) = \begin{cases} 0 & x_1 = 0 \\ 3,520.20 + 51.20x_1 & 0 < x_1 \leq 50,000 \end{cases}$$

$$f_2(x_2) = \begin{cases} 0 & x_2 = 0 \\ 82,810.00 + 52.10x_2 & 0 < x_2 \leq 20,000 \\ 102,810.00 + 51.10x_2 & 20,000 < x_2 \leq 60,000 \\ 162,810.00 + 50.10x_2 & 60,000 < x_2 \leq 80,000 \\ 242,810.00 + 49.10x_2 & 80,000 < x_2 \leq 100,000 \end{cases}$$

$$f_3(x_3) = \begin{cases} 60.50x_3 & 0 \leq x_3 \leq 50,000 \\ 75,000.00 + 59.00x_3 & 50,000 < x_3 \leq 80,000 \end{cases}$$

Notemos que las funciones construidas $f_i(x_i)$ son cóncavas y lineales por pedazos, esto quiere decir que dependiendo del intervalo en el que se encuentre la variable de decisión, será diferente la función objetivo; por ejemplo, si $x_2 \in (20,000, 60,000]$ la función a considerar será $102,810.00 + 51.10x_2$, la cual se obtuvo del siguiente modo: se tiene un cargo fijo de $82,810.00$ más un costo variable; el costo variable en este caso se construye de dos precios diferentes que corresponden a los dos primeros intervalos que debemos tomar en cuenta para su correcto cálculo. Veamos que x_2 puede tomar valores superiores a los $20,000$ hasta los $60,000$, por haber tomado x_2 el valor de $20,000$ tendrá un costo de 52.10 mientras que el costo del intervalo actual es de 51.10. Haciendo las siguientes operaciones $82,810.00+52.10(20,000)+51.10x_2-51.10(20,000)$ es como obtenemos la función objetivo deseada $102,810.00+51.10x_2$. Observemos que en esta operación, se resta la cantidad de $51.10(20,000)$ puesto que las veinte mil unidades ya fueron cobradas a un costo de 52.10. La obtención de las respectivas funciones dependiendo de los intervalos se hace de manera análoga. Así el modelo es representado por:

$$\begin{array}{rl} \text{mín} & f_1(x_1)+f_2(x_2)+f_3(x_3) \\ \text{s.a} & x_1+x_2+x_3=150,000 \\ & 0 \leq x_1 \leq 50,000 \\ & 0 \leq x_2 \leq 100,000 \\ & 0 \leq x_3 \leq 80,000 \end{array}$$

Problemas multiplicativos. Sea $A = R^n$ y sea f en el problema (P) definida para cada $x \in R^n$ por

$$f(x) = \Pi_{j=1}^{p} f_j(x)$$

donde $p > 2$ y para cada $j = 1, 2, ..., p$ $f_j : R^n \to R$ satisface que $f_j(x) \geq 0$ para todo $x \in D$. Entonces el problema es llamado un problema multiplicativo. Cuando $f_j(x^0) = 0$ para algún $j \in \{1, 2, ..., p\}$ y algún $x^0 \in D$ entonces el valor mínimo global para el problema (P) es cero y x^0 es una solución óptima global. Entonces supondremos que para cada $j = 1, 2, ..., p$ $f_j(x) > 0$ para todo $x \in D$. Un problema multiplicativo (P) para el cual $f_j : R^n \to R$, $j = 1, ..., p$ es una función cóncava es llamado un problema multiplicativo cóncavo. De manera análoga podemos encontrar problemas multiplicativos lineales convexos. Cuando (P) es un problema multiplicativo cóncavo, la función $w_i : D \to R$ definida para cada $x \in D$ por

$$w_1(x) = lnf(x) = \Sigma_{j=1}^{p} lnf_j(x)$$

donde f está definida como $f(x) = \Pi_{j=1}^{p} f_j(x)$, es posible mostrar que es una función cóncava sobre D. También puede mostrarse que en este caso la función $w_2 : D \to R$ dada por

$$w_2(x) = [f(x)]^{1/p}$$

es también una función cóncava sobre D. Así, resolver un problema multiplicativo cóncavo (P) es equivalente a resolver los problemas de minimización cóncava

$$glob\, \text{mín}\, w_k(x), \text{ sujeto a } x \in D, k = 1, 2.$$

Problemas min-max Supongamos que el conjunto

$$\{(x, y) \in R^{n+m} | Ax + By \leq b, x, y \geq 0\}$$

es un conjunto convexo compacto diferente del vacío, donde A es una matriz de $p \times n$, B es una matriz de $p \times n$ y $b \in R^p$. Sea $c \in R^n$ y $d \in R^m$ vectores. Entonces el problema lineal min-max se establece como:

$$\min_x \left\{ \max_y [\langle c, x\rangle + \langle d, y\rangle \right\}$$
$$sujeto \quad a$$

$$Ax + By \leq b, \quad x, y \geq 0 \tag{1.3}$$

En el problema (P) si tomamos $A = D = \{x \in R^n | x \geq 0, By \leq b - Ax$ para al menos un $y \geq 0\}$ y f definida para $x \in D$ por

$$f(x) = \langle c, x\rangle + \max\{\langle d, y\rangle | By \leq b - Ax, y \geq 0\},$$

entonces es posible mostrar que f es una función de valor real sobre D, que D es un conjunto convexo compacto no vacío y minimizar f sobre D es equivalente a resolver el problema 1.3.

Problemas con economía escala. En los problemas de economía a escala el costo por unidad de una o más actividades decrece conforme el nivel de actividad aumenta. Los costos para tales actividades dan lugar a funciones cóncavas separables y no separables que han de ser minimizadas sobre algún dominio convexo.

Un fenómeno similar ocurre en varios problemas donde lo que se busca es maximizar la función de utilidad con valores marginales crecientes sujeto a restricciones. Problemas como estos pueden conducir a problemas de maximización convexa, o equivalentemente a problemas de minimización cóncava.

1.3. Aplicaciones indirectas

El tipo de problemas que se muestran en esta sección son aquellos que de manera inmediata su formulación no tiene la forma de un problema de minimización cóncava, el cuál nos hemos estado refiriendo como el problema (P), pero que puede ser reformulado como tal o bien puede ser resuelto a través de resolver una serie de problemas de minimización cóncava relacionados a éste.

Problemas de programación entera. Una amplia variedad de problemas de programación entera y sus aplicaciones pueden ser formulados como problemas equivalentes de minimización cóncava continua. En este punto la equivalencia es en el sentido de que el conjunto de soluciones óptimas global coincide. Sea $B =$

$\{0,1\}$, $C \subset \mathbb{R}^n$ un conjunto convexo y cerrado, $C \cap B^n \neq \varnothing$, $e = (1, ..., 1)^T \in \mathbb{R}^n$ y $E = \{x \in \mathbb{R}^n : 0 \leq x \leq e\}$. Supongamos que $f : \mathbb{R} \to \mathbb{R}$ es una función dos veces diferenciable sobre $A \supset E$. Entonces, es posible mostrar que existe $\mu_0 \in \mathbb{R}$ tal que para todo $\mu > \mu_0$, la función $f(x) + \mu x^T(e - x)$ es cóncava sobre E y el problema

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{s.a} & x \in C \cup B^n \end{array}$$

es equivalente a

$$\begin{array}{ll} \text{mín} & f(x) + \mu x^T(e - x) \\ \text{s.a} & x \in C \cup B^n \end{array}$$

[7]

Problemas de programación bilineal. Los problemas de programación bilineal han sido especialmente importantes en el desarrollo de la minimización cóncava debido a que la programación bilineal tiene numerosas aplicaciones en la vida real, además hay ciertos procedimientos usados para resolver problemas bilineales que son muy similares a los usados en problemas generales de minimización cóncava.

Matemáticamente la forma general del programa bilineal es:

$$\begin{array}{rl} \text{mín}\, f(x,y) = & \langle p, x\rangle + x^T C y + \langle q, y\rangle \\ \text{s.a} & x \in X,\ y \in Y \end{array}$$

donde X y Y son poliedros no vacíos en R^n y R^m respectivamente. C es una matriz de $n \times m$, $p \in R^n$ y $q \in R^m$. Debido al término mixto $x^T C y$ en la función objetivo, $f(x,y)$ no es convexa ni cóncava sobre $X \times Y$. Consideremos el problema anterior como la minimización sobre X de la función $v : X \to \mathbb{R}$ definida para cada $x \in X$ por

$$v(x) = \text{mín}_{y \in Y} f(x,y)$$

Es posible mostrar que ésta es una función sobre X y el programa bilineal puede ser resuelto minimizando v sobre X. [3]

Problemas de optimización d.c. (donde la función objetivo es una diferencia de funciones convexas) Una función de valor real g sobre un conjunto convexo $C \in \mathbb{R}^n$ es llamada *d.c.* (diferencia de convexos) sobre C cuando para cada $x \in C$ $g(x)$ puede ser expresado de la forma

$$g(x) = p(x) - q(x)$$

donde p y q son funciones convexas sobre C. El problema general de programación d.c. tiene la forma

$$\begin{aligned} \text{mín} \quad & g(x) \\ \text{s. a} \quad & h_j(x) \leq 0, \quad j = 1, ..., m \\ & x \in C \end{aligned}$$

donde $C \subset \mathbb{R}^n$ es un conjunto convexo no vacío y las funciones g y h_j, $j = 1, ..., m$ son d.c. sobre C. Generalmente problemas de programación d.c. con restricciones convexas pueden ser expresados como problemas de minimización cóncava. En particular, si g es d.c. sobre algún conjunto convexo abierto F contenido en C y las restricciones definen un conjunto no vacío convexo cerrado, entonces el problema anteriormente planteado puede ser escrito en la forma (P). Para ver esto, supongamos que estas condiciones se cumplen y que para cada $x \in F$, $g(x) = p(x) - q(x)$, donde p y q son funciones convexas sobre F. Sea D definido por

$$D = \{(x, t) \in C \times R | h_j(x) \leq 0, j = 1, ..., m; p(x) \leq t\}$$

Sea $A = F \times R$ y sea $f(x, t) = t - g(x)$ para cada $(x, t) \in F \times R$. Entonces, como las restricciones del problema planteado definen un conjunto convexo cerrado no vacío, D es también un conjunto cerrado convexo diferente del vacío. Más aún, f es una función real cóncava sobre un conjunto convexo A contenido en D que minimiza a f sobre D, en este caso esto es equivalente a resolver el problema general d.c. [3]

Problemas complementarios. Muchas aplicaciones indirectas ocurren en programación lineal, cuadrática y no lineal, frecuentemente en el contexto de resolver formulaciones primales duales o condiciones de optimalidad para estos problemas. Es posible encontrar numerosas aplicaciones directas de problemas complementarios, por ejemplo; el recubrimiento secundario de aceite, la competencia económica y varios tipos de problemas en ciencia, tecnología y economía.

Sea H un conjunto en R^n y sea $g, h : R^n \rightarrow R^m$. Un problema complementario es el problema de encontrar un vector $x \in H$ tal que

$$g(x) \geq 0, \quad h(x) \geq 0, \quad \langle g(x), h(x) \rangle = 0$$

Cuando al menos uno de dichos vectores existe, encontrar tal vector es equivalente a encontrar una solución óptima al problema de optimización global

$$\begin{aligned} \text{mín}\, u(x) = \quad & \Sigma_{i=1}^{m} \; \text{mín}\{g_i(x), h_i(x)\} \\ \text{s. a} \quad & g(x) \geq 0, h(x) \geq 0, \qquad x \in H \end{aligned}$$

[3]

1.4. Las tres aproximaciones

Esta sección está constituida por tres subsecciones, en cada una de las cuales se presentan los títulos y una breve descripción de algunos de los artículos cuyo estudio se basa en alguno de los diferentes tipos de métodos que se estudian en el presente trabajo. Tales métodos son: Métodos de Corte, Métodos de aproximación sucesiva y los Métodos de partición sucesiva.

Es importante mencionar que en el artículo *Methods for global concave minimization: A bibliographic survey* , [13] hacen una revisión de los artículos que hasta entonces se conocían sobre minimización cóncava, dando una breve descripción de las ideas principales de éstos, dentro de esta revisión se incluye a los métodos de enumeración de puntos extremos, los métodos de corte, ramificación y acotamiento usando aproximaciones a la función objetivo, aproximación exterior e interior, aproximación a la programación bilineal, métodos a gran escala de programación cóncava global. Podemos decir que dicho artículo representa un compendio de los inicios de la minimización cóncava, pero el interés sobre esta área a continuado hasta nuestros días.

1.4.1. Métodos enumerativos

Dentro de los métodos enumerativos se consideran: la enumeración de los puntos extremos y los diferentes métodos de aproximación a través de cortes como los algoritmos de corte puro, combinación de cortes, aproximación por coberturas de conos. Algunos de los artículos que desarrollan algoritmos basados en la enumeración de puntos extremos son *Enumerative techniques for solving some nonconvex global optimization problems* escrito por P.M. Pardalos (1987) en el cual se revisan diferentes métodos para resolver problemas de minimización no convexa a través de la enumeración de puntos extremos.

De los métodos enumerativos en el presente trabajo nos enfocaremos solamente en los métodos de corte y enseguida mencionaremos algunos de los trabajos recientes hechos sobre éste tema.

Métodos de corte

A continuación se presentan algunas referencias que basan su trabajo en el uso de métodos de corte.

1. *A nontangential cutting plane algorithm* , Siriphong Lawphongpanich (2002) El autor propone el uso de cortes no tangenciales en la solución del problema dual lagrangiano de programación convexa, ya que usualmente estos cortes son tangenciales a la función dual. En este artículo, el autor con su propuesta asegura la convergencia a la solución óptima. [10]

2. *Finitely convergent cutting planes for concave minimization* de Marcus Porembski (2001) en el cual propone una modificación a la cortadura cóncava que produce planos de corte más profundos asegurando la convergencia finita del algoritmo puro de planos de corte basado en tales cortes. [14]

3. *Cone adaptation strategies for a finite and exact cutting plane algorithm for concave minimization* por Marcus Porembski (2002), en este artículo

abordan la dificultad que se encuentra al construir los cortes de acuerdo al algoritmo y es que conforme el número de iteraciones aumenta, los cortes tienden a hacerse superficiales, por lo que proponen un procedimiento llamado cono de adaptación el cual profundiza los cortes cóncavos de tal manera que el corte resultante tiene al menos una cierta profundidad de $\delta > 0$, donde δ es independiente de cualquier iteración, lo cual obliga la convergencia finita del algoritmo basado en planos de corte. [15]

4. *Cutting plane method for multiple objective stochastic integer linear programming*, Moncef Abbas, Fatima Bellahcene (2006). El problema que se enfrenta es el de incorporar variables enteras en las restricciones de un problema de programación lineal estocástica multiobjetivo (MOSLP). Se transforma el problema en uno determinístico entero lineal multiobjetivo a través de suponer ciertas preferencias por parte del decisor, en base a esto es posible hacer uso de los métodos de corte para dar una solución eficiente. [1]

5. *A secuential cutting plane algorithm for solving convex NLP problems* Claus Still, Tapio Westerlund, (2006) Se propone un nuevo algoritmo para resolver problemas de programación no lineal convexa, basados en métodos de planos de corte. Una ventaja es que los subproblemas utilizados mantienen un tamaño fijo. El algoritmo propuesto indica, a través de experimentos numéricos ser más rápido que el algoritmo clásico de corte de Kelley. [17]

1.4.2. Los métodos de aproximación sucesiva

Los métodos más conocidos de aproximación sucesiva son: aproximación exterior, aproximación exterior a través de colapsar politopos, aproximación interior (anexión poliedral), estimación inferior sucesiva. Enseguida mencionamos algunos de los trabajos más recientes.

1. *A multiperiod approach to the solution of the groundwater management problems using an outer approximation method* de Alexander A Sipiliotopoulos, George P Karatzas, George F Pinder (2004). En este artículo los autores proponen un algoritmo que se basa en un modelo multiperiodo, que a diferencia de otros que resuelven el mismo problema usan programación dinámica. Esto con la idea de evitar que el número de variables de decisión crezca. Emplean el método de aproximación exterior a través de cortes para determinar la solución óptima. Este método fue probado en un acuífero hipotético y en un sitio real contaminado a gran escala, ilustrando así la efectividad del algoritmo. [16]

2. *An improved piecewise outer approximation algorithm for the global optimization of MINLP models involving concave and bilinear terms* de Maria Lorena Bergamini, Ignacio Grossmann, Nicolás Scenna, Pío Aguirre

(2008). En este artículo se presenta una nueva versión del algoritmo de optimización global de aproximación exterior hecho por Bergamini, Aguirre y Grossman en 2005, con el propósito de aumentar la convergencia en modelos de programación no lineal entera mixta que involucran términos bilineales y cóncavos. Los autores establecen que el trabajo propuesto se ha aplicado a diferentes problemas en el área de redes de agua, mostrando una reducción significativa en el costo computacional. [4]

1.4.3. Los métodos de partición sucesiva

En el área de optimización global los métodos de partición sucesiva mejor conocidos como de ramificación y acotamiento, son utilizados con algunas modificaciones según el tipo de problema que se este estudiando. Así, encontramos algoritmos de partición sucesiva simpliciales, cónicos y rectangulares. Algunos artículos recientes sobre éstos son:

1. *Finite exact branch and bound algorithms for concave minimization over polytopes* de Marco Locatelli y Nguyen V. Thoai (2000). En este artículo se garantiza que algoritmos simpliciales de ramificación y acotamiento a través de algunas modificaciones en lo referente a la reglas de búsqueda local introduciendo cortes y actualizando la función objetivo, alcanzan el óptimo en un número finito de pasos. [11]

2. *A branch and bound method for solving integer separable concave problems* escrito por O Barriento, R. Correa, P. Reyes y A. Valdebenito (2003). Como el título lo menciona, los autores proponen un algoritmo basado en el método de ramificación y acotamiento para resolver problemas enteros cóncavos separables, para lo cual se usan métodos de dualidad Lagrangeana para obtener cotas inferiores y superiores. Muestran que el problema dual de un programa cóncavo separable es un programa lineal, lo cual representa una gran ventaja al momento de solucionarlo, se identifica a un excelente candidato para probar en cada región que se genera en el proceso de ramificación y muestran una condición suficiente de optimalidad para dicho candidato. [12]

3. *A branch and bound algorithm for concave network flow problems* desarrollado por Dalila B.M.M. Fontes, Eleni Hadjiconstantinou, N. Christofides, donde proponen una implementación del método de ramificación y acotamiento para la solución óptima del problema de flujo en redes con costos cuya función es cóncava donde hay un único nodo fuente sin límite de capacidad, estos problemas son considerados NP-difíciles, aún en su caso más sencillo. [6]

2

Métodos de corte

Los métodos de corte están dentro de los llamados métodos enumerativos, los cuales son aplicados por ejemplo a problemas cuya región de soluciones factibles es un poliedro no vacío, acotado y la función objetivo está definida sobre un conjunto convexo abierto que contiene a la región de soluciones factibles. En base a esto y la propiedad 3 [1] sabemos que el mínimo global se encontrará en uno de los puntos extremos de la región factible. Esto da pie a establecer que bajo dichas condiciones el mínimo global se alcanzará enumerando adecuadamente los puntos extremos.

El presente capítulo está dividido en cuatro secciones; en la primer sección se estudia la construcción de los diferentes tipos de cortes que usaremos; la segunda sección está dedicada al desarrollo del algoritmo de corte puro, presentando además un ejemplo; la tercera sección trata del algoritmo de corte de cara y finalmente la cuarta sección aborda el algoritmo de corte y división.

En general, podemos decir que el papel que juegan los cortes es el de un hiperplano que separa una parte de la región de soluciones factibles sobre las cuales el valor de la función objetivo sabemos que no será mejor a uno ya encontrado.

2.1. Cortes cóncavos.

Consideremos el problema de minimizar una función continua $f : \mathbb{R}^n \to \mathbb{R}$ sobre un conjunto poliédrico D de dimensión completa. Supongamos que γ es el valor más pequeño de $f(x)$ tomado sobre todos los posibles puntos donde se ha hecho una evaluación en alguna iteración en el proceso de resolver el problema. Entonces podemos restringir nuestra investigación al subconjunto de D que consiste de los puntos x que satisfacen $f(x) < \gamma$. Dicho de otro modo,

[1] Ver Capítulo 1

podemos dejar a un lado todos los puntos $x \in D$ en el conjunto

$$G = \{x \in \mathbb{R}^n : f(x) \geq \gamma\} \tag{2.1}$$

Otra manera de trabajar esto mismo es añadiendo la restricción $f(x) < \gamma$ al conjunto factible D. Sin embargo, esto nos genera al nuevo conjunto factible

$$D^*(\gamma) = D \cap \{x : f(x) < \gamma\}$$

el cual en general no es un poliedro, sólo en casos excepcionales, por lo que su manejo sería difícil. Así que para preservar la estructura de poliedro del conjunto de restricciones, consideraremos una función afín $l(x)$ tal que la restricción $l(x) \leq 0$ no excluya ningún punto factible x con $f(x) < \gamma$, es decir, tal que:

$$D^*(\gamma) \subset \{x \in D : l(x) \geq 0\} \tag{2.2}$$

Definición 1. *Una desigualdad lineal $l(x) \geq 0$ que satisface 2.2 es llamada un corte $\gamma-$válido para (f, D).*

Usualmente tenemos un punto $z \in D$ tal que $f(z) > \gamma$ y nos gustaría tener un corte que lo eliminara, es decir, necesitamos que:

$$l(z) < 0.$$

Cuando $f(x)$ es una función convexa, entonces el conjunto $D^*(\gamma)$ es convexo, cualquier hiperplano $l(x) = 0$ separa estrictamente z de $D^*(\gamma)$ es un corte $\gamma-$válido que excluye z.

Nuestra situación es más complicada pues $f(x)$ es una función cóncava. Así que una manera de aprovechar lo que se tiene es trabajando sobre el conjunto $D^*(\gamma)$ (relativo a D), es decir, el conjunto $D \cap G$ el cual si es convexo. De este modo se aprovecha la convexidad para construir un hiperplano que separe estrictamente z de $D^*(\gamma)$.

Supongamos que tenemos un vértice x^0 de D que satisface $x^0 \in \operatorname{int} G$, es decir, $f(x^0) > \gamma$ y es no degenerado, es decir, hay exactamente n aristas de D que salen de x^0. Sean $u^1, u^2, ..., u^n$ las direcciones de estas aristas. Si la i-ésima arista es un segmento de linea que une x^0 con un vértice adyacente y^i, entonces podemos tomar $u^i = y^i - x^0$. Sin embargo, ciertas arista pueden ser no acotadas (los correspondientes vectores u^i son direcciones extremas de D). Como D tiene dimensión completa ($\dim D = n$) los vectores $u^1, u^2, ..., u^n$ son linealmente independientes y el cono con vértice en x^0 y generado por las semirectas que salen de x^0 en las direcciones $u^1, u^2, ..., u^n$ es un cono de dimensión n con exactamente n aristas tales que $D \subset K$ (de hecho K es el cono más pequeño con vértice en x^0 que contiene a D).

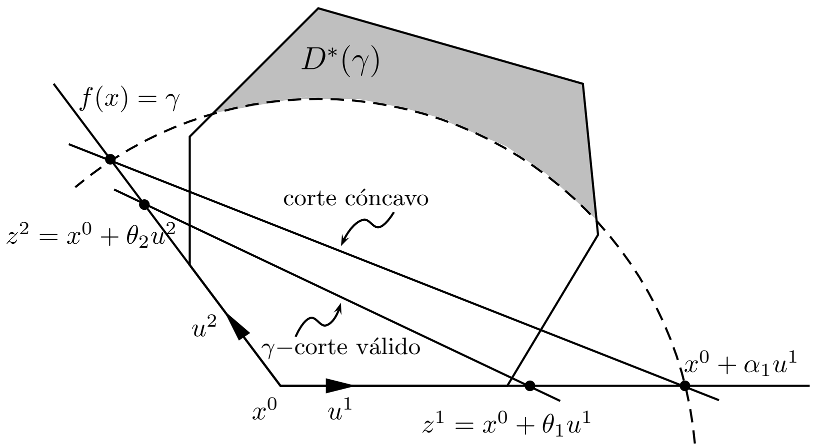


Figura 2.1: Cortes

Ahora para cada $i = 1, ..., n$ tomemos un punto $z^i \neq x^0$ sobre la i-ésima arista de K tal que $f(z^i) \geq \gamma$ (ver la figura 2.1) Estos puntos existen gracias al supuesto de que x^0 es un punto interior de G ($i.e, x^0 \in \text{int } G$). La matriz de $n \times n$

$$Q = (z^1 - x^0, z^2 - x^0, ..., z^n - x^0)$$

con $z^i - x^0$ como la i-ésima columna, es no singular debido a que sus columnas son linealmente independientes. Sea $e = (1, 1, ..., 1)$ el vector renglón de n unos.

Con base en lo anterior, construimos puntos z^i generadores de la matriz Q con respecto a x^0 para poder resolver un sistema de desigualdades que como establece el siguiente teorema, define un corte $\gamma-$válido.

Teorema 1. *Si $z^i = x^0 + \theta_i u^i$ con $\theta_i > 0$ y $f(z^i) \geq \gamma$ entonces la desigualdad lineal*

$$eQ^{-1}(x - x^0) \geq 1 \tag{2.3}$$

define una corte $\gamma-$ válido para (f, D).

Sea $U = (u^1, u^2, ..., u^n)$ una matriz de $n \times n$ no singular con columnas $u^1, u^2, ..., u^n$. Entonces $x \in K$ si y sólo si

$$x = x^0 + \Sigma_{i=1}^{n} t_i u^i$$

con $t = (t_1, t_2, ..., t_n)^T \geq 0$, es decir

$$K = \{x : x = x^0 + Ut, t \geq 0\} \tag{2.4}$$

$t = U^{-1}(x - x^0)$. Usando 2.4 y del teorema anterior podemos deducir la siguiente forma de un corte válido.

Corolario 1. *Para cualquier* $\theta = (\theta_1, \theta_2, ..., \theta_n) > 0$ *tal que* $f(x^0 + \theta_i u^i) \geq \gamma$ *la desigualdad lineal*

$$\Sigma_{i=1}^{n} \frac{t_i}{\theta_i} \geq 1 \tag{2.5}$$

donde $t = U^{-1}(x - x^0)$, *define un corte* $\gamma-$ *válido para* (f, D)

Definición 2. *Un corte* $\gamma-$ *válido que excluye una gran parte del conjunto* $D \cap G = \{x \in D : f(x) \geq \gamma\}$ *más que ningún otro, se dice que domina a éste, o que es más fuerte o profundo.*

Si $\theta_i \geq \theta'_i$ $(\forall i)$ entonces el corte $\Sigma_{i=1}^{n} \frac{t_i}{\theta'_i} \geq 1$ no puede dominar al corte 2.5. Entonces, el corte $\gamma-$ válido más fuerte del tipo de 2.5 se da cuando $\theta_i = \alpha_i$ $(i = 1, 2, ..., n)$, donde

$$\alpha_i = \sup\{\theta \in \mathbb{R}_+ : x^0 + \theta u^i \in G\} = \sup\{\theta \in \mathbb{R}_+ : f(x^0 + \theta u^i) \geq \gamma\} \tag{2.6}$$

Como este corte fue originado en programación cóncava y como el conjunto $D^*(\gamma)$ del cual es separado x^0 es cóncavo, nos referiremos a este corte profundo, como

$$\Sigma_{i=1}^{n} \frac{t_i}{\alpha_i} \geq 1$$

un corte cóncavo.

Definición 3. *Una corte de la forma 2.5 con* $\theta_i = \alpha_i$ *que satisface 2.6 es llamado un corte cóncavo* $\gamma-$ *válido para* (f, D) *construido en el punto* x^0.

Nótese que algunos valores de α_i definidos por 2.6 pueden ser iguales a $+\infty$. Esto sucede cuando u^i es la dirección de recesión del conjunto convexo G. Tomando $\frac{1}{\alpha_i} = 0$ cuando $\alpha_i = +\infty$, vemos que si $I = \{i : \alpha_i < +\infty\}$ entonces el corte cóncavo está dado por

$$\Sigma_{i \in I} \frac{t_i}{\alpha_i} \geq 1 \tag{2.7}$$

Su vector normal es $\pi = (\pi_1, \pi_2, ..., \pi_n)$ con $\pi = \frac{1}{\alpha_i}$ $(i \in I)$, $\pi_j = 0$ $(j \notin I)$, es decir, en este caso el hiperplano del corte cóncavo es paralelo a toda dirección u^j con $j \notin I$ como se puede ver en la figura 2.2. El corte 2.7 existirá sólo si I es no vacío, de otro modo, el cono K tendrá todas sus aristas contenidas en G, en cuyo caso $K \subset G$ y no hay un punto $x \in D$ con $f(x) < \gamma$.

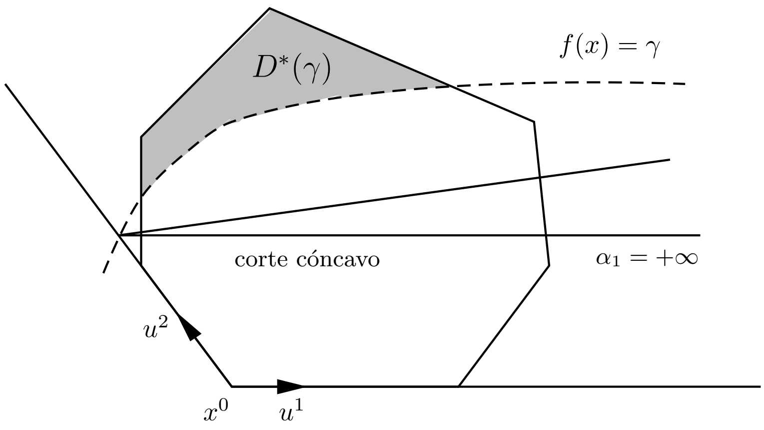


Figura 2.2: Cortes en el caso degenerado

Un caso especial que hay que considerar cuando de construir cortes se trata es que en ocasiones los vértices adyacentes a x^0 podrían ser $s > n$, es decir, estamos frente a un caso degenerado. Para dar solución a este problema 'Carvajal y Moreno (1972)' propusieron encontrar un vector normal π de un corte $\gamma-$válido como una solución básica al sistema de desigualdades

$$\pi u^i \geq \frac{1}{\alpha_i} \quad (i = 1, 2, ..., s) \tag{2.8}$$

donde $u^1, u^2, ..., u^s$ son las direcciones de las aristas de D que salen de x^0 con $(s > n)$ y α_i definido como en 2.6

Proposición 1. *Cualquier solución π de 2.8 genera un corte $\gamma-$válido para (f, D) a través de la siguiente desigualdad*

$$\pi(x - x^0) \geq 1$$

Si π es una solución básica de 2.8 entonces se cumple que n de las desigualdades de 2.8 se satisfagan en igualdad, es decir, el hiperplano correspondiente pasa a través de n puntos $z^i = x^0 + \alpha_i u^i$.

2.2. Un algoritmo de corte puro

El objetivo de esta sección es presentar el algoritmo de corte puro, el cual encuentra el óptimo global para un problema de minimización cóncava, para ello es necesario presentar la definición de una γ **-extensión**, el teorema que da una condición suficiente para la optimalidad global, el algoritmo y finalmente el teorema que garantiza la convergencia de éste.

El problema de programación cóncava que estudiaremos aquí es

$$\min \quad f(x) \tag{2.9}$$
$$s.a \quad Ax \leq b \tag{2.10}$$
$$x \geq 0 \tag{2.11}$$

donde A es una matriz de $m \times n$, x es un $n-$vector, b es un $m-$vector y $f : \mathbb{R}^n \to \mathbb{R}$ es una función cóncava. Supondremos que la región de soluciones factibles $D = \{x : Ax \leq b, x \geq 0\}$ es acotado con int$D \neq \varnothing$ y para cualquier número real α el conjunto de nivel $\{x : \mathbb{R}^n : f(x) \geq \alpha\}$ es acotado.

2.2.1. Cortes válidos

Consideremos a x^0 una solución básica factible cuyo valor en la función objetivo obtenido por algún método es el más pequeño hasta ahora, la pregunta es ¿cómo sabemos que x^0 es un óptimo global? Recordemos que la región de soluciones factibles sobre la que estamos trabajando es un politopo y sabemos que el mínimo global de $f(x)$ se alcanza en alguno de sus vértices, así que lo que haremos es minimizar $f(x)$ sobre el conjunto de vértices de D, además para cualquier politopo con vértices $u^1, u^2, ..., u^s$ el número $\min\{f(u^1), f(u^2), ..., f(u^s)\}$ representa una cota inferior para $f(D \cap P)$. Observemos que los puntos u^i podrían no ser puntos de D, aún así, puntos x fuera de D al ser evaluados en la función objetivo nos pueden brindar información acerca de los puntos que si son elementos de D. Con esto en mente, comencemos por la siguiente definición.

Definición 4. *Sea x^0 un vértice de D, sea $\gamma = f(x^0)$, para cualquier $x \in \mathbb{R}^n$ que satisfaga que $f(x) \geq \gamma$ el punto $x^0 + \theta(x - x^0)$ tal que*

$$\theta = \sup\{t : t \geq 0, f(x^0 + t(x - x^0)) \geq \gamma\} \tag{2.12}$$

*es llamada la γ-**extensión** de x con respecto a x^0.*

El siguiente es un ejemplo de como se obtienen las γ-extensiones:

Sea $D = \{x \in \mathbb{R}^2 : \frac{-1}{2}x_1 + x_2 \leq 1,\ x_1 + x_2 \leq 4,\ x_1 - x_2 \leq 2,\ x_1 \geq 0, x_2 \geq 0\}$ y sea $f(x_1, x_2) = -(x_1 - 1)^2 - (x_2 - 1)^2 + 2$. Consideremos al origen $x^0 = (0, 0)$ y a sus vértice adyacentes $(0, 1), (2, 0)$ donde $\gamma = \min\{f(0, 0), f(0, 1), f(2, 0)\} = f(0, 0) = 0$. Tomamos a $K = \mathbb{R}^2$ cuyas direcciones generadoras son $d_1 = e_1 = (0, 1)$, $d_2 = e_2 = (1, 0)$. Lo que buscamos es encontrar un par de punto en el plano tales que satisfagan la siguiente ecuación, $(-(x_1 - 1)^2 - (x_2 - 1)^2 + 2 = 0)$ la cual es la función objetivo igualada al valor de γ, que es lo mismo que $(x_1 - 1)^2 + (x_2 - 1)^2 = 2$. De este modo, las $\gamma-$extensiones y_1 y y_2 son la intersección de los ejes positivos con el círculo $(x_1 - 1)^2 + (x_2 - 1)^2 = 2$, así, $y_1 = (2, 0)$ y $y_2 = (0, 2)$

Ahora consideremos $y^1, y^2, ..., y^s$ vértices de D que son adyacentes a x^0 $(s \geq n)$, supondremos que $\text{mín}\{f(y^1), f(y^2), ..., f(y^s)\} \geq \gamma$ y para cada vértice tomemos su γ-extensión; $z^i = x^0 + \theta_i(y^i - x^0)$. Por otro lado por la proposición 1 sabemos que cualquier solución π al sistema de desigualdades lineales

$$\theta_i \pi(y^i - x^0) \geq 1 \quad con \quad i = 1, ..., s \tag{2.13}$$

genera un corte γ-válido para (f, D).

Así pues, con la solución π al sistemas de desigualdades lineales anterior, el siguiente teorema establece una condición suficiente para encontrar la optimalidad global de x^0.

Teorema 2. *Sea π una solución al sistema 2.13 entonces*

$$\pi(x - x^0) > 1 \ \ para\ todo\ x \in D\ tal\ que\ f(x) < \gamma. \tag{2.14}$$

Así, si

$$\text{máx}\{\pi(x - x^0) : x \in D\} \leq 1, \tag{2.15}$$

entonces x^0 es una solución óptima global del problema de programación cóncava.

El algoritmo que a continuación se presenta funciona básicamente del siguiente modo; Primero buscamos un vértice que sea un mínimo local de f sobre D, esto lo podemos hacer eligiendo un vértice de D, después tomamos vértices adyacentes a él, los cuales evaluamos en la función objetivo y el que sea el menor, ése será el mínimo local, que denotaremos como x^0. Enseguida construimos un corte $\gamma-$válido $\pi^0(x - x^0) \geq 1$ con respecto a x^0 y resolvemos el programa lineal $\text{máx}\, \pi^k(x - x^0)$ s.a $x \in D_0$, si la solución w^0 al evaluarla en la función objetivo del programa lineal es menor o igual a 1, entonces por el teorema anterior sabemos que x^0 será un mínimo global. En caso contrario, en base a dicho teorema, una mejor solución que x^0 debe ser buscada en la nueva región de soluciones factibles $D_1 = D_0 \cap \{x : \pi^0(x - x^0) \geq 1\}$. Así, empezamos desde w^0 a buscar un mínimo local x^1 de f sobre D_1, puede suceder que $f(x^1) < \gamma$. Así, por el teorema anterior, x^1 satisface la desigualdad $\pi^0(x - x^0) \geq 1$ de manera estricta, como x^1 es un vértice de D_1, debe serlo también de D_0, así las cosas el procedimiento se repite con x^1 y $\gamma_1 = f(x^1)$ en lugar de x^0y $\gamma = f(x^0)$. En cambio, si sucede que $f(x^1) \geq \gamma$ el proceso se repite pero con $x^0 \leftarrow x^1$, $D \leftarrow D_1$ hasta que γ no cambie. Siguiendo este procedimiento después de cada iteración obtenemos un vértice que es mejor en el sentido de que sea un mínimo, o bien se reduce la región de soluciones factibles y así irnos acercando al óptimo.

Algoritmo de corte puro

Inicialización:

Buscar un vértice x^0 que sea un mínimo local.
Sea $\gamma = f(x^0)$, $D_0 = D$
Iteración k=0,1,...:

1. En x^k construir un corte γ-válido π^k para (f, D).
2. Resolver el siguiente programa lineal

$$\text{LP}(x^k, \pi^k, D_k) \quad \text{máx}\, \pi^k(x - x^k) \;\; \text{s.a} \;\; x \in D_k$$

   Sea w^k una solución óptima básica de este programa lineal.

   Si $\pi^k(w^k - x^k) \leq 1$ entonces Fin, x^0 es un mínimo global.
   Sino ir a 3

3. Sea $D_{k+1} = D_k \cap \{x : \pi^k(x - x^k) \geq 1\}$. Empezando desde w^k encontrar un vértice x^{k+1} de D_{k+1} que es un mínimo local de $f(x)$ sobre D_{k+1}

   Si $f(x^{k+1}) \geq \gamma$ ir a la iteración $k+1$
   Sino ir a 4

4. Sea $\gamma \leftarrow f(x^{k+1})$, $x^0 \leftarrow x^{k+1}$, $D_0 \leftarrow D_{k+1}$ e ir a la iteración 0.

Teorema 3. *Si la sucesión $\{\pi^k\}$ es acotada, entonces el anterior algoritmo de corte es finito.*

Ejemplo. Consideremos el siguiente problema:

$$\begin{array}{lccc} \text{mín} & -(x-1.2)^2 & - & (y-0.6)^2 \\ \text{s.a} & & & \end{array}$$

$$\begin{array}{rcrl} -2x & + & y & \leq 1 \\ & & y & \leq 2 \\ x & + & y & \leq 4 \\ x & & & \leq 3 \\ \frac{1}{2}x & - & y & \leq 1 \\ x & , & y & \geq 0 \end{array}$$

Sea $x^0 = (0,0)$, $f(x^0) = -1.8 = \gamma$ $D_0 = D$ donde D es la región de soluciones factibles.

Iteración 0.

1. Construir un corte $\gamma-$válido π^0 para (f, D_0).
   dos vértices adyacentes a x^0 son $y^1 = (0,1)$, $y^2 = (2,0)$.
   $\text{mín}\{f(y^1) = -1.6, f(y^2) = -1\} = -1.6 > \gamma$
   $z^1 = x^0 + \theta_1(y^1 - x^0)$ donde θ =sup$\{t : t \geq 0, f(x^0 + t(x - x^0)) \geq \gamma\}$.
   De aquí que θ_1 =sup$\{0, 1, \frac{6}{5}\}$ y $z^1 = (0, \frac{6}{5})^t$
   $z^2 = x^0 + \theta_2(y^2 - x^0)$ Donde, θ_2 =sup$\{0, 1, \frac{6}{5}\}$ y $z^2 = (\frac{12}{5}, 0)^t$.
   Para generar el corte hay que obtener una solución para el siguiente sistema:

$$\begin{aligned}\theta_1\pi(y^1 - x^0) \geq 1\\ \theta_2\pi(y^2 - x^0) \geq 1\end{aligned}$$

sustituyendo tenemos:

$$\begin{aligned}\tfrac{6}{5}(x\ y)(0,1)^t \geq 1\\ \tfrac{6}{5}(x\ y)(2,0)^t \geq 1\end{aligned}$$

De aquí que $\pi^0 = (\frac{5}{12}, \frac{5}{6})$, así el corte es: $\frac{5}{12}x + \frac{5}{6}y \geq 1$

2. Resolver: $\text{máx}\, \pi^0(x - x^0)$ s.a $x \in D_0$
   La solución al problema de programación lineal es: $w^0 = (2,2)^t$ y su valor en la función objetivo es 2.5. Como $2.5 > 1$ ir a 3.

3. Sea $D_1 = D_0 \cap \{x : \pi^0(x - x^0) \geq 1\} = D_0 \cap \{x : \frac{5}{12}x + \frac{5}{6}y \geq 1\}$.
   Empezando en $w^0 = (2,2)^t$ buscamos en vértice que sea un mínimo local en D_1

$$w^0 = (2,2),\ f(2,2) = -2.6 \quad (\tfrac{1}{2}, 2)^t,\ f(\tfrac{1}{2}, 2)^t = -2.45$$
$$(3,1),\ f(3,1) = -3.4$$

   como $-3.4 < -1.8$ ir a 4.

4. Sea $\gamma = -3.4$ $x^0 = (3,1)^t$ $D_0 = D_1$ ir a la iteración 0.

Iteración 0.

1. En $x^0 = (3,1)^t$ construir un corte $\gamma-$válido π^0 para (f, D_0).
   dos vértices adyacentes a x^0 son $y^1 = (2,2)$, $y^2 = (3, \frac{1}{2})$.
   $\text{mín}\{f(y^1) = -2.6, f(y^2) = -3.25\} = -3.25 > \gamma$
   $z^1 = x^0 + \theta_1(y^1 - x^0)$ donde $\theta = \sup\{t : t \geq 0, f(x^0 + t(x - x^0)) \geq \gamma\}$.
   De aquí que $\theta_1 = \sup\{0, 1, \frac{7}{5}\}$ y $z^1 = (\frac{8}{5}, \frac{12}{5})^t$
   $z^2 = x^0 + \theta_2(y^2 - x^0)$ Donde, $\theta_2 = \sup\{0, 1, \frac{8}{5}\}$ y $z^2 = (3, \frac{1}{5})^t$.
   Para generar el corte hay que obtener una solución para el siguiente sistema:

$$\begin{aligned}\theta_1\pi(y^1 - x^0) \geq 1\\ \theta_2\pi(y^2 - x^0) \geq 1\end{aligned}$$

   sustituyendo tenemos:

$$\begin{aligned}\tfrac{7}{5}(x\ y)(-1,1)^t \geq 1\\ \tfrac{8}{5}(x\ y)(0,-\tfrac{1}{2})^t \geq 1\end{aligned}$$

   De aquí que $\pi^0 = (-\frac{55}{28}, -\frac{5}{4})$, así el corte es: $-\frac{55}{28}x - \frac{5}{4}y \geq -6.1429$

2. Resolver: $\max \pi^0(x - x^0)$ s.a $x \in D_0$
   La solución al problema de programación lineal es: $w^0 = (0,0)^t$ y su valor en la función objetivo es $\frac{50}{7}$. Como $\frac{50}{7} > 1$ ir a 3.

3. Sea $D_1 = D_0 \cap \{x : \pi^0(x - x^0) \geq 1\} = D_0 \cap \{x : -\frac{55}{28}x - \frac{5}{4}y + \frac{50}{7} \geq 1\}$.
   Empezando en $w^0 = (0,0)^t$ buscamos en vértice que sea un mínimo local en D_1

$$w^0 = (0,0),\ f(0,0) = -1.8 \quad (\tfrac{1}{2}, 2)^t,\ f(\tfrac{1}{2}, 2)^t = -2.45$$
$$(1,0),\ f(1,0) = -1.6 \quad (2,0),\ f(2,0) = -1 \quad (2.85, .43)$$
$$f(2.85, .43) = -2.75 \quad (1.8545, 2)\ f(1.8545, 2) = -2.3885$$

   Sea $x^1 = (2.85, .43)$ como $f(2.85, .43) = -2.75 > -3.4 = \gamma$ ir a la iteración 1.

Iteración 1.

1. En $x^1 = (2.85, .43)$ construir un corte $\gamma-$válido π^1 para (f, D_1).
   dos vértices adyacentes a x^1 son $y^1 = (2,0)$, $y^2 = (1.86, 2)$.
   $\min\{f(y^1) = -1, f(y^2) = -1.5244\} = -1.5244 > \gamma$
   $z^1 = x^1 + \theta_1(y^1 - x^1)$ donde $\theta = \sup\{t : t \geq 0, f(x^1 + t(x - x^1)) \geq \gamma\}$.
   De aquí que $\theta_1 = \sup\{0, 0.206875\}$ y $z^1 = (3.0258, 0.3410)^t$
   $z^2 = x^1 + \theta_2(y^2 - x^1)$ Donde, $\theta_2 = \sup\{0, 0.985\}$ y $z^2 = (1.8749, -1.1165)^t$.
   Para generar el corte hay que obtener una solución para el siguiente sistema:

$$\theta_1\pi(y^1 - x^1) \geq 1$$
$$\theta_2\pi(y^2 - x^1) \geq 1$$

   sustituyendo tenemos:

$$0.206875(x\ y)(.85, .43)^t \geq 1$$
$$0.985(x\ y)(-0.99, 1.57)^t \geq 1$$

   De aquí que $\pi^1 = (4.0641, 3.2091)$.

2. Resolver: $\max \pi^0(x - x^1)$ s.a $x \in D_1$
   La solución al problema de programación lineal es: $w^0 = (1.85, 2)^t$ y su valor en la función objetivo es 0.9926. Como $0.9926 < 1$ entonces $x^0 = (3,1)^t$ y $f(3,1) = -3.4$ es el mínimo global.

En la figura 2.3 se da un panorama general de los valores de γ usados y los cortes que se usaron durante la aplicación del algoritmo.

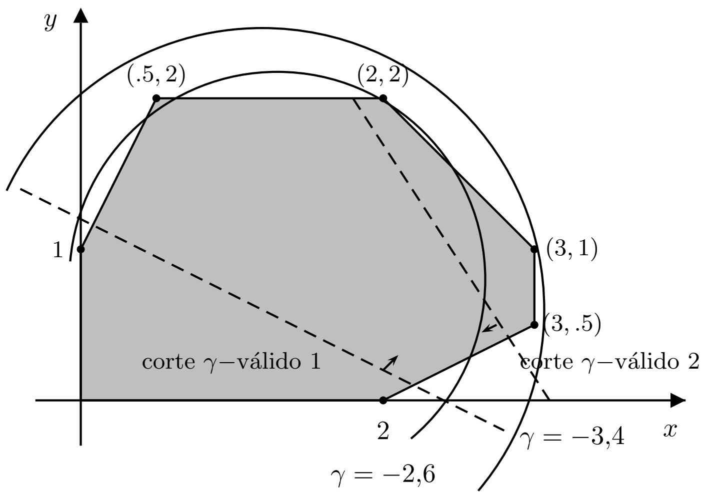


Figura 2.3: Cortes y valores de γ usados durante el algoritmo

2.3. Algoritmos de corte de cara

El algoritmo que se presenta en esta sección recibe dicho nombre gracias a su mecanismo para encontrar el mínimo global en un politopo; lo que veremos es que cada corte de cara que generamos, quita al menos una cara de D que ya no es necesaria, esto sin quitar ningún punto extremo de D que no haya sido ya eliminado por algún corte anterior. Así, este algoritmo se basa en encontrar caras de D que llamaremos caras extremas y eliminarlas a través de cortes de cara. El procedimiento para encontrar las caras extremas está basado en el trabajo de Majthay y Whinston (1974).

Definición 5. *Una cara F de un poliedro D es llamada una **cara extrema** de D relativa al poliedro M si*

$$\varnothing \neq F \cap M \subset riF \tag{2.16}$$

2.3.1. El problema de caras extremas

Dados dos poliedros D y M, encontrar una cara extrema de D relativa a M o bien demostrar que tal cara no existe.

Supongamos que las restricciones que definen a un politopo D son de la siguiente forma:

$$x_i = p_{i0} - \Sigma_{j \in J} p_{ij} x_j \quad (i \in B) \tag{2.17}$$

$$x_k \geq 0 \quad (k = 1, ..., n) \tag{2.18}$$

donde B es el conjunto de índices de las soluciones básicas ($|B| = m$) y J es el conjunto de índice de las variables no básicas ($|J| = n - m$).

Para cualquier punto $x \in \mathbb{R}^n$ definamos $Z(x) = \{j|x_j = 0\}$. De este modo es posible dar otra caracterización de las caras extremas, lo cual nos ayudará más adelante para construir un algoritmo exclusivamente para encontrarlas.

Proposición 2. *Sea $x^0 \in D \cap M$, $F_0 = \{x \in D : x_j = 0 \ \forall j \in Z(x^0)\}$, entonces F^0 es una cara extrema de D relativa a M si y sólo si para cualquier $i \in \{1, 2, ..., n\} \diagdown Z(x^0)$:*

$$0 < \text{mín}\{x_i : x \in D \cap M, x_j = 0 \ \forall j \in Z(x^0)\}.$$

Sea M el sistema de desigualdades lineales

$$\Sigma_{j=1}^{n} c_{ij} x_j \geq d_i \quad (i = n+1, ..., \bar{n})$$

Introduciendo las variables de holgura x_i $(i = n+1, ..., \bar{n})$y usando 2.17 y 2.18 podemos describir al politopo $D \cap M$ a través del siguiente sistema

$$x_i = \bar{p_{i0}} - \Sigma_{j \in \bar{J}} \bar{p_{ij}} x_j \quad (i \in \bar{B}) \tag{2.19}$$

$$x_i \geq 0 \quad (i = 1, ..., \bar{n}) \tag{2.20}$$

donde $\bar{B}$ es el conjunto de índices de las variables básicas ($|\bar{B}| = m + (\bar{n} - n)$), $\bar{J}$ es el conjunto de índices de las variables no básicas ($|\bar{J}| = n - m$).

En lo sucesivo, las variables x_i $(i = 1, ..., n)$ serán las variables 'originales', mientras que las otras $(x_i, i > n)$ como las 'no originales'.

En general, sea $\{1, 2, ..., n\} \diagdown Z(x^0) = \{i_1, i_2, ..., i_s\}$.

Procedimiento I.

Empezando desde $k = 1$ resolver:

$$(P_k) \quad \text{mín}\, x_{i_k} \quad s.a \tag{2.21}$$

$$\begin{aligned} &x_i = \bar{p_{i0}} - \Sigma_{j \in \bar{J}} \bar{p_{ij}} x_j && (i \in \bar{B}) \\ &x_i \geq 0 && (i = 1, ..., \bar{n}) \\ &x_j = 0 && \forall j \in Z(x^{k-1}) \end{aligned}$$

Sea ψ_k el valor óptimo y x^k la solución óptima básica de (P_k).

Sea $k \leftarrow k + 1$ y repetir el procedimiento hasta que $k = s$.

Proposición 3. *Si $Z(x^s) \neq \varnothing$, entonces $F = \{x \in D : x_j \ \forall j \in Z(x^s)\}$ es una cara extrema de D relativa a M, de otro modo no hay cara extrema de D relativa a M más que D misma.*

Para resolver el problema (P_k) para $k = 1$ recordemos que x^0 está dado por 2.19 y 2.20. Si $Z(x^0) = J \cap \{1, 2, ..., n\}$ es decir, todas las variables x_j, $j \in Z(x^0)$ son no básicas, entonces para resolver P_1 aplicamos el algoritmo simplex al programa lineal:

$$(P_1*) \quad \text{mín}\, x_{i_1} \text{ s.a} \quad 2.19 \text{ y } 2.20^*$$

donde * significa que todas las variables básicas originales en 2.19 y 2.20 deben omitirse.

Sin embargo puede suceder que $Z(x^0) \cap \bar{B} \neq \varnothing$, es decir, que algunas variables básicas originales sean iguales a cero. En este caso, debemos quitar las variables x_i, $i \in Z(x^0) \cap \bar{B}$, de la base siempre que sea posible, para hacerlo, como $x_i^0 = 0$ se cumple que

$$0 = \text{mín}\{x_i : 2.19,\ 2.20 \text{ y } x_j = 0\ \forall j \in \bar{J} \cap \{1, 2, ..., n\}\},$$

así $\bar{p_{ij}} \geq 0 \forall j \in \bar{J} \diagdown \{1, 2, ..., n\}$. Si $\bar{p_{ij}} > 0$ para al menos $j \in \bar{J} \diagdown \{1, 2, ..., n\}$, entonces pivoteando sobre el elemento (i, j) estaremos forzando a x_i a salir de la base. Por otro lado, si $\bar{p_{ij}} = 0 \forall j \in \bar{J} \cap \{1, 2, ..., n\}\}$ esto significa que x_i depende sólo de las variables x_j, con $j \in \bar{J} \cap \{1, 2, ..., n\}\}$. Siguiendo este procedimiento para toda $i \in Z(x^0) \cap \bar{B}$ transformaremos la tabla 2.19 en una donde las variables x_i, $i \in Z(x^0)$ que son básicas sólo depende de las variables no básicas originales. Hasta entonces es posible empezar el algoritmo simplex para minimizar x_{i_1}. Este mismo procedimiento se puede hacer para cada (P_k).

Ahora bien, el conjunto $Z(x^k)$ es igual al conjunto de índices Z_k de las variables no básicas originales en la tabla óptima de (P_k^*) más los índices de las variables básicas originales, las cuales se encuentran en un nivel cero. En particular, esto quiere decir que $F = \{x \in D : x_j = 0 \quad \forall j \in Z_s\}$ así, F es un vértice de D si y sólo si $|Z_s| = n - m$. Esto nos da un criterio para saber cuando hemos encontrado una cara extrema; cuando $k = s$ ó bien cuando $|Z_k| = n - m$.

2.3.2. Cortes de cara válidos

Sea $F = \{x \in D : X_j = 0 \quad \forall j \in Z\}$ una cara extrema de D relativa a M que se obtuvo a través del método anterior. Si $F = D$ entonces por la definición de cara extrema, $D \cap M \subset riD$ en este caso M no contiene ningún vértice de D. Ahora si F es un vértice de D entonces éste puede ser eliminado por algún corte sin eliminar ningún punto que sea una mejor solución. Entonces sólo nos queda considerar el caso en el que F es una cara propia de D pero no un vértice $(0 < |Z| < n - m)$.

Definición 6. *Sea F una cara propia pero no un vértice de D. Una desigualdad lineas $l(x) \geq 0$ es un* ***corte de cara*** *de éste elimina F sin eliminar cualquier vértice de D que este en M.*

La construcción de un corte facial se hace a través del siguiente procedimiento:

Sea $\alpha_j (j \in Z)$ un número escogido y para cada $h \in \{1, 2, ..., n\} \diagdown Z$ consideremos el programa de programación paramétrica:

$$(P_h(q))\, \text{mín}\{x_h : x \in D \cap M,\ \Sigma_{j \in Z} \alpha_j x_j \leq q\} \tag{2.22}$$

donde q es un parámetro no negativo. Por la proposición 1 se sigue que

$$0 < \text{mín}\{x_h : x \in D \cap M, x_j = 0\ \forall j \in Z\}. \tag{2.23}$$

entonces el valor óptimo en 2.22 para $q = 0$ es positivo. Así si,

$q_h = sup\{q : 0 <$valor óptimo de $(P_h(q))\}$ entonces $q_h > 0$ y así,

$$p := \text{mín}\{q_h : h \in \{1, 2, ..., n\} \diagdown Z\} > 0. \tag{2.24}$$

Proposición 4. *Sea $0 < |Z| < n - m$. Si $p < +\infty$ entonces la desigualdad*

$$\Sigma_{j \in Z} \alpha_j x_j \geq p \tag{2.25}$$

define un corte de cara válido.

2.3.3. Un algoritmo finito de corte

Algoritmo

Inicialización:

Buscar un vértice x^0 que sea un mínimo local.
Sea $\gamma = f(x^0)$, $D_0 = D$.
Elegir dos números $\delta_0 \geq 0$, $N > 1$ y sea $d_0 = +\infty$.

Iteración k=0,1,...:

0) Si $k = 0$ o $d_{k-1} \geq \delta_{k-1}$ ir a $1a$

   Sino ir a $1b$

1a) Construir un corte γ-válido $l_k(x) := \pi^k(x - x^k) - 1 \geq 0$ para (f, D) en x^k.
   Sea $d_k = \frac{1}{\|\pi^k\|}$, $\delta_k = \delta_{k-1}$ e ir a 2.

1b) Comenzar desde x^k, identificar una cara extrema F_k de D relativa a D_k (la intersección de D con los cortes generados previamente).

Si $F_k = D$ FIN x^0 es una solución óptima del problema de programación cóncava.

Si F_k es un vértice de D (i.e, $F_k = x^k$), entonces construye en x^k un corte γ-válido $l_k(x) := \pi^k(x-x^k)-1 \geq 0$ para (f, D). Sea $d_k = \frac{1}{\|\pi^k\|}$, $\delta_k = \frac{1}{N}\delta_{k-1}$ e ir a 2.

Si F_k es una cara propia pero no es un vértice de D, construir un corte de cara válido $l_k(x) \geq 0$.

Si este corte es infinito FIN x^0 es una solución óptima global.

Sino sea $d_k = +\infty$, $\delta_k = \frac{1}{N}\delta_{k-1}$ e ir a 3.

2) Resolver el programa lineal

$$\text{máx}\, l_k(x) \quad s.a \quad x \in D_k$$

obtener una solución óptima básica w^k.

Si $l_k(w^k) \leq 0$ FIN x^0 es una solución óptima global.

Sino ir a 3

3) Sea $D_{k+1} = D_k \cap \{x : l_k(x) \geq 0\}$ Encontrar un vértice x^{k+1} de D_{k+1} que sea una mínimo local de $f(x)$ sobre D_{k+1}.

Si $f(x^{k+1}) \geq \gamma$ ir a la iteración $k+1$

Sino ir a 4

4) Sea $\gamma \leftarrow f(x^{k+1})$, $x^0 \leftarrow x^{k+1}$, $D_0 \leftarrow D_{k+1}$ y regresar a la iteración 0.

2.4. Algoritmo de corte y división

Este método como su nombre lo indica se basa en hacer cortes para obtener una mejor solución además de generar una partición de la región de soluciones factibles, a través de la construcción de conos.

2.4.1. Partición de un cono

Definición 7. *Un cono K es un cono poliédrico convexo cuyo vértice es el origen y está generado por n vectores linealmente independientes $z^1, z^2, ..., z^n$.*

El cono K generado por tales vectores constituye una matriz no singular $Q = (z^1, z^2, ..., z^n)$ la cual se denotará como $K = con(Q)$.

Un punto x pertenece a K si y sólo si $x = \Sigma\lambda_i z^i = Q\lambda^t$ con $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n) \geq 0$ tenemos

$$K = \{x : Q^{-1}x \geq 0\} \tag{2.26}$$

Para poder realizar la partición de un cono es necesario tomar un punto dentro de éste, digamos u, como $u \in K$ entonces debe cumplir $Q^{-1}u = (\lambda_1, \lambda_2, ..., \lambda_n) \geq 0$, generamos un conjunto de índice $I = \{i : \lambda_i > 0\}$, de este modo podemos expresar a z^i como una combinación lineal de z_j con $(i \neq j)$ y u:

$$z^i = \frac{u - \Sigma_{j\neq i}\lambda_j z^j}{\lambda_i}$$

Así, la matriz $Q_i = (z^1, ..., z^{i-1}, u, z^{i+1}, ..., z^n)$ obtenida de Q a través de sustituir u por z^i sigue siendo no singular.

Observemos que $(intK_i) \cap (intK_j) = \varnothing \quad (j \neq i)$; $K = \cup\{K_i : i \in I\}$

Con esto lo que hemos construido es la partición de K con respecto a u.

2.4.2. Algoritmo

Con base en lo anterior daremos una breve descripción del algoritmo. Este método comienza dándole un valor a ϵ y al igual que los algoritmos anteriores se busca un vértice que sea un mínimo local de f sobre D, después se establece el valor para α, enseguida reescribimos el problema con respecto a x^0; con esto aseguramos que el cono sobre el cual trabajamos tenga como vértice el origen de la región de soluciones factibles, además con esto se satisface $\theta_i = \text{máx}\{t : f(te^i) \geq \gamma\} > 0$ para $i = 1, ..., n$. Así pues, hemos construido a la matriz Q_0 cuyas columnas son los vectores linealmente independientes generadores de K llamados z^i que son elementos de $G_\alpha := \{z : f(z) = \alpha, f(\lambda z) < \alpha\ \forall\lambda > 1\}$, es decir, los puntos z^i en $\mathbb{R}^n$ son $\alpha-$ extensiones de algún $y \neq 0$. Entonces, de la siguiente expresión $eQ^{-1}x \geq 1$, $\pi = eQ^{-1}$ con $Q = \{z^1, z^2, ..., z^n\}$ define un corte $\alpha-$válido para (f, D) en x^0.

Como subrutina es necesario resolver un programa de programación lineal cuya solución óptima se denotará $w(Q)$ cuyo valor óptimo en éste programa se denotará por $\mu(Q)$.

Si $\mu(Q) < 1$ entonces x^0 es un óptimo $\epsilon-$global. Pero si $f(w(Q)) < \alpha$ esto quiere decir que hemos encontrado un vértice x^1 de D con un valor más pequeño que el de $w(Q)$, es decir, $f(x^1) < f(w(Q)) < \alpha$, con esto modificamos la región de soluciones factibles D por $D \cap \{x : eQ^{-1}x \geq 1\}$ y empezamos el algoritmo con x^1 en lugar de x^0.

Si $\mu(Q) > 1$ y $f(w(Q)) \geq \alpha$ quiere decir que tenemos que seguir buscando sobre la región factible que dejó el último corte $D \cap \{x : eQ^{-1}x \geq 1\}$, con la finalidad de deshacernos de la mayor parte de D que no queremos, construiremos

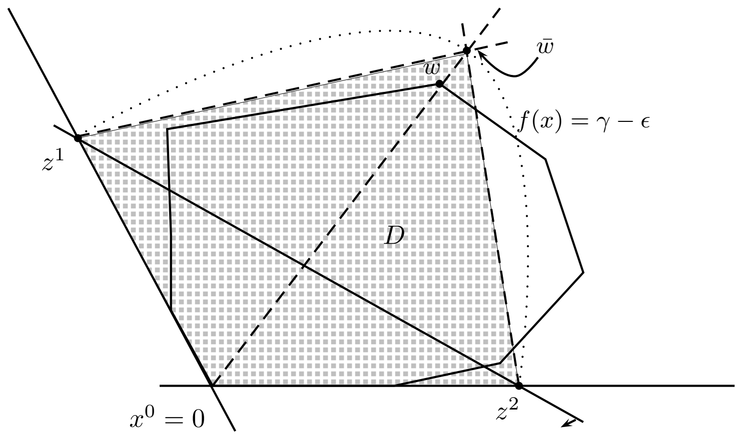


Figura 2.4: Gráfica general del proceso de corte y división

una $\alpha-$extensión $\bar{w}$ de w y dividiremos el cono $K = con(Q)$ con respecto a $\bar{w}$. La i-ésima partición de la matriz Q está dada por $Q_i = \{z^1, ..., z^{i-1}, \bar{w}, z^{i+1}, ..., z^n\}$. Como puede verse en la figura 2.4.

Ahora bien, para saber si hay puntos factibles que satisfacen $f(x) < \gamma - \epsilon$ en cualquiera de los subconos K_i para cada $i \in I$ es necesario resolver el programa lineal

$$\text{máx} \quad eQ^{-1}x \quad \text{s.a} \quad x \in D, \quad Q_i^{-1}x \geq 0$$

donde $Q_i^{-1}x \geq 0$ es una restricción que asegura que la región factible sobre la cual estamos trabajando es el cono K_i. En caso de que todos los programas lineales tengan un valor menor a 1, significa que no hay puntos $x \in D$ en ningún cono K_i que satisfaga $f(x) < \gamma - \epsilon$, por lo que podremos asegurar que x^0 es una solución $\epsilon-$global. En caso contrario, cada subcono K_i cuya solución en el programa lineal sea mayor que 1 tiene que ser explorado a través del mismo procedimiento de división.

Algoritmo

Elegir $\epsilon \geq 0$

Inicialización
Calcular un punto $z \in D$. Sea $M = D$

Fase I

Empezando desde z buscar un vértice x^0 de M que sea un mínimo local de $f(x)$ sobre M.

Fase II

0) Sea $\gamma = f(x^0)$, $\alpha = \gamma - \epsilon$.

   Reescribe el problema en su forma estándar con respecto a x^0.

   Construye $Q_0 = (z^{01}, z^{02}, ..., z^{0n})$ donde z^{0i} es la intersección de G_α con la $i-$ésima arista de K.

   Sea $\mathcal{M} = \{Q_0\}$

1) Para cada $Q \in \mathcal{M}$ resuelve el siguiente programa lineal

$$PL(Q, M) \quad \text{máx}\, eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

   para resolver una solución óptima básica $w(Q)$ y el valor óptimo $\mu(Q) = eQ^{-1}w(Q)$

   Si $f(w(Q)) < \alpha$ para alguna Q entonces, sean

$$M \leftarrow M \cap \{x : eQ^{-1}x \geq 1\} \quad z \leftarrow w(Q)$$

   y regresar a la Fase I.

   Sino ir a 2)

2) Sea $R = \{Q \in \mathcal{M} : \mu(Q) > 1\}$

   Si $\mathcal{R} = \varnothing$ Fin: x^0 es una solución óptima ϵ-global

   Sino ir a 3

3) Para cada $Q \in \mathcal{R}$ construye una α-extensión $\bar{w}(Q)$ de $w(Q)$ y divide Q con respecto a $\bar{w}(Q)$. Reemplaza Q por la partición resultante y que $\mathcal{M}'$ la colección de matrices resultantes.

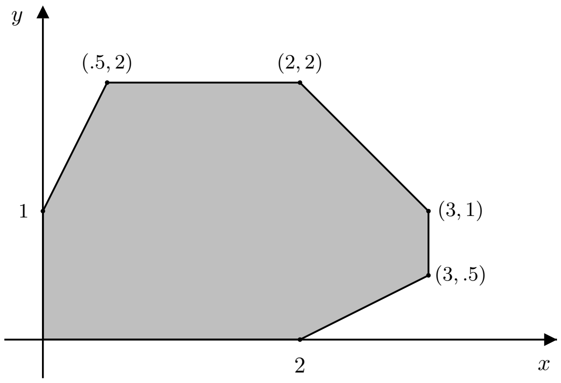
   Hacer $\mathcal{M} \leftarrow \mathcal{M}'$ y regresar a 1.

Ejemplo. Consideremos el ejemplo anterior.

$$\begin{array}{llcl}
\text{mín} & -(x-1.2)^2 & - & (y-0.6)^2 \\
\text{s.a} & & &
\end{array}$$

$$\begin{array}{rcrl}
-2x & + & y & \leq 1 \\
 & & y & \leq 2 \\
x & + & y & \leq 4 \\
x & & & \leq 3 \\
\frac{1}{2}x & - & y & \leq 1 \\
x & , & y & \geq 0
\end{array}$$

La región de soluciones factibles que representa el sistema de desigualdades definido en el ejemplo es la que enseguida se muestra:



Aplicando el algoritmo de corte y división:

Sea $\epsilon = \frac{1}{2}$

Inicialización.

Sea $z = (0,0) \in D$. Sea $M = D$

Fase I.

Empezando desde z buscar un vértice x^0 de M que sea un mínimo local de $f(x)$ sobre M. Sea $x^0 = (0,0)$ y $f(x^0) = -1.8$.

Fase II.

0) Sea $\gamma = -1.8$, sea $\alpha = -1.8 - 0.5 = -2.3$

   Reescribir el problema con respecto a x^0; como en este caso $x^0 = (0,0)$ el problema se queda igual.

   Construir $Q_0 = (z^{01}, z^{02})$ donde z^{0i} es la intersección de G_α con la i-ésima arista de K_0. En este caso K_0 tiene como origen el punto $(0,0)$ y está formado por el primer cuadrante de $\mathbb{R}^2$. De este modo los z^{0i} serán las $\alpha-$ extensiones, así que necesitamos de dos vértices $y^1 = (0,1)^t$ y $y^2 = (2,0)^t$

   $z^{01} = x^0 + \theta_1(y^1 - x^0)$ donde $\theta =$sup$\{t : t \geq 0, f(x^1 + t(x - x^1)) \geq \gamma\}$.

   De aquí que $\theta_1 =$sup$\{0, 1, 1.5275\}$ y $z^{01} = (0, 1.5275)^t$

   $z^2 = x^1 + \theta_2(y^2 - x^1)$ Donde, $\theta_2 =$sup$\{0, 1, 1.2965\}$ y $z^2 = (2.593, 0)^t$.

Así, la matriz que buscamos es la siguiente:

$$\mathbf{Q_0} = \begin{pmatrix} 0 & 2.593 \\ 1.5275 & 0 \end{pmatrix}$$

Sea $\mathcal{M} = \{Q_0\}$

1) Para cada $Q \in \mathcal{M}$ resolver el programa lineal correspondiente:

$$\text{máx} \quad eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

sustituyendo tenemos:

$$\text{máx} \quad 0.3857x + 0.6547y \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

de donde: $w(Q) = (2,2)$, $\mu(Q) = 2{,}0808$, $f(w(Q)) = -2.6$.

Como $-2.6 < -2.3$ entonces la región de soluciones factibles se modifica del siguiente modo:

$$\begin{aligned} M \leftarrow & M \cap \{x : eQ_0^{-1}x \geq 1\} \\ & M \cap \{x : 0.3857x + 0.6547y \geq 1\} \\ z \leftarrow & w(Q) = (2,2) \end{aligned}$$

y regresar a la Fase I.

Fase I.

Empezando desde $z = (2,2)$ buscar un vértice x^0 de M que sea un mínimo local de $f(x)$ sobre M. Sea $x^0 = (3,1)$ y $f(x^0) = -3.4$.

Fase II.

0) Sea $\gamma = -3.4$, sea $\alpha = -3.4 - 0.5 = -3.9$

Reescribir el problema con respecto a x^0; como en este caso $x^0 = (3,1)$ el problema es el siguiente:

$$\begin{array}{ll} \text{mín} & -(x-4.2)^2 \quad - \quad (y-1.6)^2 \\ \text{s.a} & \end{array}$$

$$\begin{array}{cccl} -2x & + & y & \leq -4 \\ & & y & \leq 3 \\ x & + & y & \leq 8 \\ x & & & \leq 3 \\ 0.5x & - & y & \leq 1.5 \\ 0.3857x & + & 0.6547y & \leq 2.8118 \\ x & & & \geq 3 \\ & & y & \geq 1 \end{array}$$

Construir $Q_0 = (z^{01}, z^{02})$ donde z^{0i} es la intersección de G_α con la i-ésima arista de K_0. En este caso K_0 tiene como origen el punto $(3,1)$ y los z^{0i} serán las $\alpha-$ extensiones, así que necesitamos de dos vértices $y^1 = (3,2)^t$ y $y^2 = (5,1)^t$

$z^{01} = x^0 + \theta_1(y^1 - x^0)$ donde θ =sup$\{t : t \geq 0, f(x^1 + t(x - x^1)) \geq \gamma\}$.

De aquí que θ_1 =sup$\{0, 1, 2, 2.17\}$ y $z^{01} = (3, 3.17)^t$

$z^2 = x^1 + \theta_2(y^2 - x^1)$ Donde, θ_2 =sup$\{0, 1, 1.541\}$ y $z^2 = (6.08, 1)^t$.

Así, la matriz que buscamos es la siguiente:

$$\mathbf{Q_0} = \begin{pmatrix} 3 & 6.08 \\ 3.17 & 1 \end{pmatrix}$$

Sea $\mathcal{M} = \{Q_0\}$

1) Para cada $Q \in \mathcal{M}$ resolver el programa lineal correspondiente:

$$\text{máx} \quad eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

sustituyendo tenemos:

$$\text{máx} \quad 0.1334x + 0.1893y \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

de donde: $w(Q) = (5,3)$, $\mu(Q) = 1.2349$, $f(w(Q)) = -2.6$.

Como $-2.6 > -3.9$ ir a 2)

2) Sea $\mathcal{R} = \{Q \in \mathcal{M} : \mu(\mathcal{Q}) > \infty\}$, entonces tenemos que $\mathcal{R} = \{Q_0\}$, como $\mathcal{R} \neq \varnothing$ ir a 3).

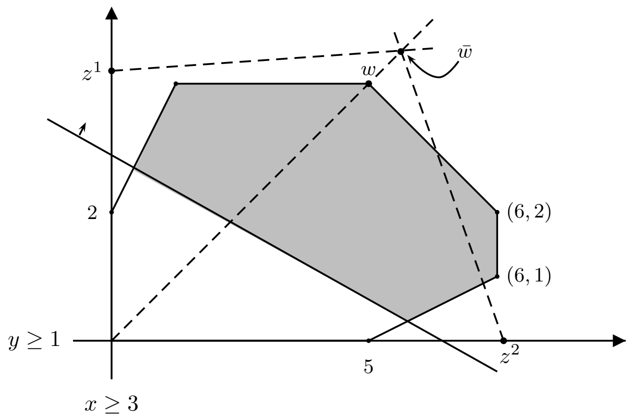
3) Para cada $Q \in \mathcal{R}$ construir una $\alpha-$extensión $\bar{w}(Q)$ de $w(Q)$ y dividir Q con respecto a $\bar{w}(Q)$.

$\bar{w}(Q) = x^0 + \theta(w(Q) - x^0)$ $\theta = sup\{0, 1, 1.132\}$ entonces

$\bar{w}(Q) = (5.264, 3.264)^t$. Y las dos nuevas matrices son:

$$\mathbf{Q_1} = \begin{pmatrix} 5.264 & 6.08 \\ 3.264 & 1 \end{pmatrix} \mathbf{Q_2} = \begin{pmatrix} 3 & 5.264 \\ 3.17 & 3.264 \end{pmatrix}$$

Ahora trabajamos sobre dos subconos que se muestran en la siguiente figura:



Nota: La escala de la gráfica se redujo a la mitad en relación con la anterior. Sea $\mathcal{M}' = \{Q_1, Q_2\}$ $\mathcal{M} \leftarrow \mathcal{M}'$ e ir a 1)

1) Para cada $Q \in \mathcal{M}$ resolver el programa lineal correspondiente: Para Q_1 tenemos:

$$\text{máx} \quad 0.1553x + 0.056y \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

donde $Q^{-1}x$ es equivalente a las siguientes dos desigualdades que forman parte de las restricciones de este programa.

$$-0.0686x + 0.417y \geq 0$$
$$0.2239x - 0.3610y \geq 0$$

de donde: $w(Q) = (6, 2)$, $\mu(Q) = 1.0438$, $f(w(Q)) = -3.4$.

Como $-3.4 > -3.9$ ir a 2.

Para Q_2 tenemos:

$$\text{máx} \quad -0.0136x + 0.3284y \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

donde $Q^{-1}x$ es equivalente a las siguientes dos desigualdades que forman parte de las restricciones de este programa.

$$-0.4734x + 0.7635y \geq 0$$
$$0.4598x - 0.4351y \geq 0$$

de donde: $w(Q) = (3.49, 3)$, $\mu(Q) = 0.9376$, $f(w(Q)) = -2.4641$.

Como $-2.4641 > -3.9$ ir a 2.

2) Sea $\mathcal{R} = \{Q \in \mathcal{M} : \mu(Q) > 1\} = \{Q_1\}$ ir a 3.

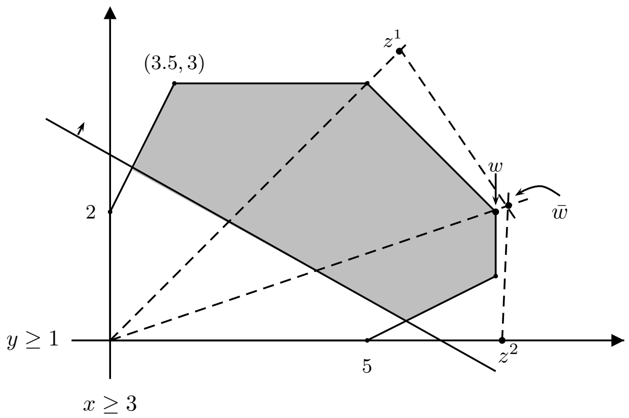
3) Para Q_1 construir una $\alpha-$extensión $\bar{w}$ de $w(Q)$.

   $\bar{w}(Q) = x^0 + \theta(w(Q) - x^0)$ $\theta = sup\{0, 1, 1.04165\}$ entonces

   $\bar{w}(Q) = (6.125, 2.0416)^t$. Y las dos nuevas matrices son:

$$\mathbf{Q_1} = \begin{pmatrix} 6.125 & 6.08 \\ 2.0416 & 1 \end{pmatrix} \mathbf{Q_2} = \begin{pmatrix} 5.264 & 6.125 \\ 3.264 & 2.0416 \end{pmatrix}$$

   Los subconos sobre los que tenemos que trabajar son:



   Sea $\mathcal{M}' = \{Q_1, Q_2\}$ $\mathcal{M} \leftarrow \mathcal{M}'$ e ir a 1)

1) Para cada $Q \in \mathcal{M}$ resolver el programa lineal correspondiente: Para Q_1 tenemos:

$$\text{máx} \quad 0.1657x - 0.0072y \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

   donde $Q^{-1}x$ es equivalente a las siguientes dos desigualdades que forman parte de las restricciones de este programa.

$$-0.159x + 0.9669y \geq 0$$
$$0.3247x - 0.9741y \geq 0$$

   de donde: $w(Q) = (6, 1{,}5)$, $\mu(Q) = 0.9834$, $f(w(Q)) = -3.25$.

   Como $-3.25 > -3.9$ ir a 2.

   Para Q_2 tenemos:

$$\text{máx} \quad 0.1323x + 0.0931y \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

donde $Q^{-1}x$ es equivalente a las siguientes dos desigualdades que forman parte de las restricciones de este programa.

$$\begin{array}{c} -0.2208x + 0.6625y \geq 0 \\ 0.3531x - 0.5694y \geq 0 \end{array}$$

de donde: $w(Q) = (6, 2)$, $\mu(Q) = 0.98$, $f(w(Q)) = -3.4$.

Como $-3.4 > -3.9$ ir a 2.

2) Sea $\mathcal{R} = \{Q \in \mathcal{M} : \mu(Q) > 1\} = \varnothing$ pues $\mu(Q_1) = 0.9834$ y $\mu(Q_2) = 0.98$. Por lo tanto hemos terminado $x^0 = (3, 1)$ es una solución $\epsilon-$óptima global.

Región de soluciones factibles respecto a $x^0 = (3, 1)$.

Observemos que debido a que el valor de ϵ es elegido por el usuario, el algoritmo podría tener una solución exacta en caso de que el valor de ϵ fuera muy pequeño, aunque esto ocasionaría que se necesitasen varias iteraciones. Sin embargo, si el valor de ϵ es grande, pocas iteraciones se harán para llegar a la solución óptima, aunque se perderá la exactitud en la solución.

3

Métodos de aproximación sucesiva

Los métodos de aproximación sucesiva se dividen básicamente en dos tipos: aproximación exterior y aproximación interior, como su nombre lo indica, en cada iteración nos acercan a la solución exacta o aproximada del problema original, ya sea por fuera o por dentro de la región factible. Estos algoritmos se basan en el principio de construir subproblemas de optimización (en este trabajo serán usados programas de programación lineal) que son más fáciles de resolver; los cuales representan una subrutina del algoritmo principal.

3.1. Aproximación exterior

Los métodos de aproximación exterior también conocidos como métodos de relajación, se basan en quitar algunas restricciones del conjunto de soluciones factibles, las cuales se van añadiendo poco a poco conforme el algoritmo y las soluciones obtenidas hasta el momento lo exigen, con esto se logra tener problemas más sencillos de resolver, si además la solución obtenida también es solución del problema original, entonces el problema queda resuelto. En caso contrario, habrá que continuar con el algoritmo hasta agotar todas las restricciones del problema.

Consideremos el problema de optimización global:

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{s.a} & \\ & x \in D \end{array} \qquad \text{(P)}$$

donde $f : \mathbb{R}^n \to \mathbb{R}$ es una función continua cóncava y $D \subset \mathbb{R}^n$ es un conjunto cerrado y supondremos que $\text{mín}\, f(D)$ existe.

Resolver (P) con métodos de aproximación exterior es resolver una sucesión de problemas relajados:

$$\begin{array}{lll} \text{mín} & f(x) & \\ \text{s.a} & & (Q_k) \\ & x \in D_k & \end{array}$$

donde $\mathbb{R}^n \supset D_1 \supset D_2 \supset \cdots \supset D$ y $\text{mín}\, f(D_k) \underset{k\to\infty}{\longrightarrow} \text{mín}\, f(D)$.

Usualmente los subconjuntos D_k pertenecen a una familia F con las siguientes propiedades:

1. Los conjuntos $D_k \subset \mathbb{R}^n$ son cerrados y cualquier problema (Q_k) con $D_k \in F$ tiene una solución y puede ser resuelto a través de algoritmos conocidos.
2. Para cualquier $D_k \in F$ que contiene a D y cualquier punto $x^k \in D_k \setminus D$ se puede definir una función de restricción $l_k : \mathbb{R}^n \to \mathbb{R}$ que satisface

$$l_k(x) \leq 0 \quad \forall x \in D \tag{3.1}$$

$$l_k(x^k) > 0 \tag{3.2}$$

$$\{x \in D_k : l_k(x) \leq 0\} \in F \tag{3.3}$$

Bajo estas suposiciones podemos establecer el siguiente método de solución:

Algoritmo de aproximación exterior:

Escoger $D_1 \in F$ tal que $D \subset D_1$

$k \leftarrow 1$

Iteración k (k=1,2,...)

Resuelve el problema relajado (Q_k) obteniendo una solución $x^k \in argminf(D_k)$.

k.a) Si $x^k \in D$ Fin: x^k resuelve (P).

k.b) Sino, construye una función de restricción $l_k : \mathbb{R}^n \to \mathbb{R}$ que satisface 3.1, 3.2, 3.3 y el conjunto $D_{k+1} = D_k \cap \{x : l_k(x) \leq 0\}$

$k \leftarrow k+1$.

Las condiciones 3.1 y 3.2 indican que el conjunto $\{x \in \mathbb{R}^n : l_k(x) = 0\}$ separa estrictamente $x^k \in D_k \setminus D$ de D. La restricción agregada $l_k(x) \leq 0$ corta un subconjunto de D_k. Sin embargo, como $D \subset D_k$ para todo k cada D_k constituye una aproximación exterior de D. En la figura 3.1 se muestra de manera general una idea de cómo funciona el proceso de aproximación exterior.

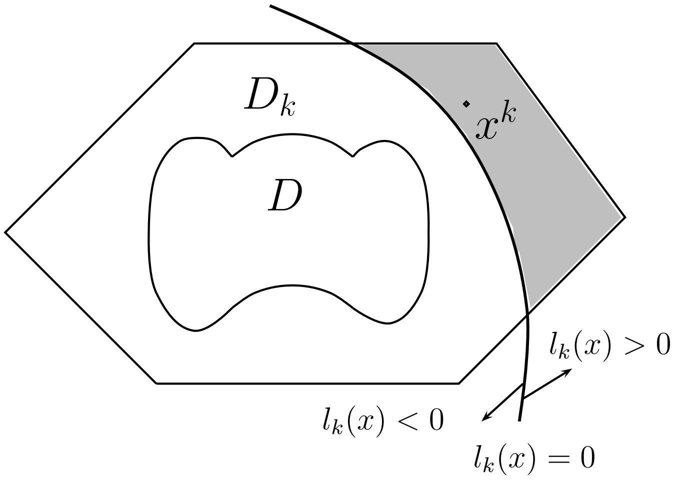


Figura 3.1: Proceso general de aproximación exterior

Como $D_k \supset D_{k+1} \supset D$ tenemos $\min f(D_k) \leq \min f(D_{k+1}) \leq \min f(D)$ y $x^k \in D$ entonces $x^k \in argminf(D)$.

Para asegurar que D_{k+1} es cerrado si D_k es cerrado, pedimos que la función de restricción sea semicontinua inferiormente.

Enseguida se muestra un teorema que nos garantiza la convergencia del algoritmo.

Teorema 4. *En el contexto del método de aproximación exterior presentado supongamos:*

1. *l_k es semicontinua inferiormente para cada $k = 1, 2, ...$;*

2. *cada subsucesión convergente $\{x^q\} \subset \{x^k\}$ que satisface $\{x^q\}_{q\to\infty} \to \bar{x}$ contiene una subsucesión $\{x^r\} \subset \{x^q\}$ tal que $\lim_{r\to\infty} l_r(x^r) = \lim_{r\to\infty} l_r(\bar{x})$ y*

3. *$\lim_{r\to\infty} l_r(\bar{x}) = 0$ implica que $\bar{x} \in D$*

Entonces todo punto de acumulación de la sucesión $\{x^k\}$ pertenece a D y así resolvemos (P).

3.1.1. Subproblemas (Q_k)

En optimización global los métodos de aproximación exterior que usan cortes afines, frecuentemente han sido aplicados a problemas que satisfacen

$$\min f(D_k) = \min f(V(D_k))$$

donde D_k es un politopo y $V_k = V(D_k)$ denota al conjunto de vértices de D_k. Aquí también supondremos que el mínimo se alcanza en alguno de los vértices del conjunto. Por lo que en cada iteración del algoritmo es importante saber quien es D_k y V_k.

Encontrando un politopo inicial D_1 y su conjunto de vértices V_1

Sea D un conjunto no vacío, compacto y convexo de dimensión n. D_1 debería ser un simple politopo muy ajustado que contiene a D y además con el menor número de vértices.

Denotemos:

$$\alpha_j := \text{mín}\{x_j : x \in D\} \qquad j = 1, ..., n \tag{3.4}$$
$$\alpha := \text{máx}\{\Sigma_{j=1}^{n} x_j : x \in D\} \tag{3.5}$$

Entonces:

$$D_1 := \{x \in \mathbb{R}^n : \alpha_j - x_j \leq 0 \ \ (j = 1, ..., n), \ \Sigma_{j=1}^{n} x_j - \alpha \leq 0\} \tag{3.6}$$

es un simplex que contiene a D. Entonces $n+1$ caras de D_1 se definen por $n+1$ hiperplanos $\{x \in \mathbb{R}^n : x_j = \alpha_j\}$ $(j = 1, ..., n)$ y $\{x \in \mathbb{R}^n : \Sigma_{j=1}^{n} x_j = \alpha\}$.

Cada uno de estos hiperplanos es un hiperplano soporte de D. El conjunto de vértices de D_1 es

$$V_1 = \{v^0, v^1, ..., v^n\}$$

donde

$$v^0 = (\alpha_1, ..., \alpha_n)^T \tag{3.7}$$

y

$$v^j = (\alpha_1, ..., \alpha_{j-1}, \beta_j, \alpha_{j+1}, ..., \alpha_n) \ (j = 1, ..., n) \tag{3.8}$$

con

$$\beta_j = \alpha - \Sigma_{i \neq j} \alpha_i$$

3.1.2. Nuevos vértices y direcciones extremas

Consideremos como conjunto convexo poliédrico actual al conjunto factible del problema relajado (Q_k) en la iteración k definido como:

$$D_k = \{x \in \mathbb{R}^n : l_i(x) \leq 0, i \in K\} \tag{3.9}$$

donde $K \subset \mathbb{N}$ es conjunto finito de índices y $k \notin K$.

Sea $l_k(x) \leq 0$ la nueva restricción que define al siguiente conjunto:

$$D_{k+1} = \{x \in \mathbb{R}^n : l_i(x) \leq 0, i \in K \cup \{k\}\}. \tag{3.10}$$

Denotemos por V_k, V_{k+1} los conjuntos de vértices de D_k y D_{k+1}, respectivamente y sea U_k y U_{k+1} el conjunto de direcciones extremas de D_k y D_{k+1}, respectivamente.

El siguiente lema caracteriza los nuevos vértices y direcciones extremas de D_{k+1}.

Lema Sea $P \subset \mathbb{R}^n$ un conjunto convexo poliédrico de dimensión n con un conjunto de vértices V y conjunto de direcciones extremas U. Sea $l(x) = ax+\beta \leq 0$ ($a \in \mathbb{R}^n, \beta \in \mathbb{R}$) sea una restricción lineal adicional y sean V' el conjunto de vértices y U' el conjunto de direcciones extremas de $P' = P \cap \{x : l(x) \leq 0\}$ entonces tenemos:

1. $w \in V' \diagdown V$ si y sólo si w es la intersección del hiperplano $\{x : l(x) = 0\}$ con una arista $[v^-, v^+]$ de P que satisface $l(v^-) < l(v^+) > 0$ o con una arista no acotada que sale de un vértice $v \in V$ en una dirección $u \in U$ que satisface $l(v) < 0$ y $au > 0$ o bien $l(v) > 0$ y $au < 0$.

2. $u \in U' \diagdown U$ si y sólo si u satisface $au = 0$ y es de la forma $u = \lambda u^- + \mu u^+$ con λ, $\mu > 0$, $au^- < 0$, $au^+ > 0$ lo que define una cara de dos dimensiones del cono de recesión de P.

Notemos que en el primer inciso del lema anterior, requiere que P tenga aristas y en el segundo inciso se necesita de la existencia de una cara de dos dimensiones del cono de recesión de P. De otro modo no podremos tener nuevos vértices y nuevas direcciones extremas respectivamente.

Una aplicación directa a este resultado es el siguiente procedimiento para calcular V_{k+1} y U_{k+1} de V_k y U_k

Sea $D_k = \{x \in \mathbb{R}^n : l_i(x) = a^i x + \beta_i \leq 0 (i \in K)\}$, donde K es un conjunto finito de índices que satisface $k \notin K$.

Método para encontrar vértices y direcciones extremas

Sea $l_k(x) = a^k x + \beta_k$ y definimos:

$$V_k^+ := \{v \in V_k : l_k(v) > 0\} \qquad V_k^- := \{v \in V_k : l_k(v) < 0\} \tag{3.11}$$

$$U_k^+ := \{u \in U_k : a^k u > 0\} \qquad U_k^- := \{u \in U_k : a^k u < 0\} \tag{3.12}$$

1. **Encontrando los vértices de $V_{k+1} \diagdown V_k$ que son puntos donde $\{x \in \mathbb{R}^n : l_k(x) = 0\}$ encuentra una arista acotada de D_k:**

   Para cualquier par $(v^-, v^+) \in V_k^- \times V_k^+$ sea $w = \alpha v^- + (1-\alpha)v^+$, donde

$$\alpha = \frac{l_k(v^+)}{l_k(v^+) - l_k(v^-)} \qquad (i.e., l_k(w) = 0) \tag{3.13}$$

Calcular

$$I(w) = \{i \in K : l_i(w) = 0\} \tag{3.14}$$

Si el rango de la matriz $A(w)$ que tiene renglones a^i, donde $l_i(x) = a^i x + \beta_i$, $a^i \in \mathbb{R}^n$, $\beta_i \in \mathbb{R}$, $i \in I(w)$, es menor que $n-1$, entonces w no puede estar en $V_{k+1} \diagdown V_k$. De otro modo, w esta en $V_{k+1} \diagdown V_k$.

Si V_k es acotado, entonces $V_{k+1} \diagdown V_k$ se determina de este modo.

2. **Encontrando los vértices de $V_{k+1} \diagdown V_k$ que son puntos donde $\{x \in \mathbb{R}^n : l_k(x) = 0\}$ encuentra una arista no acotada de D_k:**

   Para cualquier par $(u, v) \in \{U_k^- \times V_k^+\} \cup \{U_k^+ \times V_k^-\}$

   Determinar $w = v + \alpha u$, donde $\alpha = \dfrac{-l_k(v)}{a^k u}$ $(i.e., l_k(w) = 0)$.

   Calcular $I(w) = \{i : l_i(w) = 0,\ i \in K\}$ y como en el caso acotado decidir del rango de $A(w)$ si $w \in V_{k+1} \diagdown V_k$. Observemos que en este caso tenemos $I(w) = \{i : l_i(v) = 0,\ a^i u = 0,\ i \in K\}$.

3. **Encontrando las nuevas direcciones extremas:**

   Para cualquier par $(u^-, u^+) \in U_k^- \times U_k^+$

   determinar $u = (a^k u^+)u^- - (a^k u^-)u^+$. Veamos que $u \in$ cono U_k y $a^k u = 0$. Sea $J(u) = \{j : a^j u^- = a^j u^+ = 0,\ j \in K\}$.

   Si el rango del sistema de ecuaciones

   $$a^j x = 0 \quad j \in J(u)$$

   es menor que $n-2$, entonces u no puede ser una dirección extrema de D_{k+1}, de otro modo $u \in U_{k+1} \diagdown U_k$.

3.1.3. Aproximación exterior para problema con restricciones lineales

Consideremos primero el problema de minimizar una función cóncava $f : \mathbb{R}^n \to \mathbb{R}$ sobre un poliedro D generado por un sistema de desigualdades lineales de la siguiente forma

$$A_i x \leq b_i \quad (i = 1, 2, ..., m), \tag{3.15}$$
$$x_j \geq 0 \quad (j = 1, 2, ..., n), \tag{3.16}$$

donde $b_i \in \mathbb{R}$ y A_i es el i-ésimo renglón de una matriz A de $m \times n$.

El método de aproximación para resolver este problema cuando la región de soluciones factibles es no acotada se sustenta en las siguientes propiedades de una función cóncava $f : \mathbb{R}^n \to \mathbb{R}$.

Proposición 5. *Si la función cóncava $f(x)$ es acotada inferiormente sobre alguna semirecta, entonces es acotada inferiormente sobre cualquier semirecta paralela.*

Proposición 6. *Sea M cualquier conjunto convexo cerrado en $\mathbb{R}^n$. Si la función cóncava $f(x)$ es acotada inferiormente sobre todo rayo extremo de M, entonces es acotada inferiormente sobre una semirecta contenida en M.*

Corolario 2. *Sea $M \neq \varnothing$ un poliedro en $\mathbb{R}^n$ que no contiene rectas. Entonces $f(x)$ es acotada inferiormente sobre una arista no acotada de M o bien el mínimo de $f(x)$ sobre M es alcanzado en algún vértice de M.*

Si la región de soluciones factibles del problema es acotada, podemos usar el algoritmo que se describió en la sección anterior, simplemente haciendo algunas ligeras modificaciones en el inciso k.b) del siguiente modo; cuando llegamos a este inciso es porque la solución obtenida x^k no satisface alguna restricción, entonces en lugar de construir a $l_k(x)$, ésta será aquella que cumpla $i_k \in argmax\{A_i x^k - b_i : i = 1, ..., m\}$ y el conjunto $D_{k+1} = D_k \cap \{x : A_{i_k} x - b_{i_k} \leq 0\}$ y con esto vamos a la siguiente iteración.

Si la región de soluciones factibles del problema original es no acotada, entonces la región de soluciones factibles de los subproblemas también lo será, en este caso, es posible que el problema original tenga solución y algún subproblema no lo tenga, pero de acuerdo al corolario 2 es posible encontrar una arista no acotada de D_k para la cual $f(x) \to -\infty$, dicha arista tiene una dirección u^k, si esta dirección satisface $A_i u^k \leq 0$ para todo $i = 1, ..., m$ entonces de acuerdo a la proposición 6 $D = \varnothing$ o bien $f(x)$ es no acotada inferiormente sobre cualquier semirecta que va de un punto factible en la dirección u^k. Sino, u^k no satisface una de las desigualdades $A_i u \leq 0$ $i = 1, ..., m$ entonces definimos como $i_k \in argmax\{A_i u^k : i = 1, ..., m\}$ y $D_{k+1} = D_k \cap \{x : A_{i_k} x \leq b_{i_k}\}$ de este modo como $A_{i_k} u^k > 0$ entonces u^k ya no es una dirección de recesión de D_{k+1}.

La solución de cada subproblema Q_k depende de la solución y construcción de Q_{k-1}, notemos que Q_k difiere de Q_{k-1} en una sola restricción, además recordemos que el conjunto D_1 depende enteramente de nuestra elección por lo que sabremos en la primera iteración cuales son los elementos que forman a los conjunto de vértices V_1 y al de direcciones U_1, tomando en cuenta que en la iteración $k-1$ sabemos quienes son V_{k-1} y U_{k-1} y habiendo añadido la respectiva restricción al conjunto D_{k-1} para generar a D_k construir los conjuntos V_k y U_k se hará a través del método descrito en la sección anterior.

Retomando el corolario 2, podemos decir que si una función cóncava es acotada inferiormente sobre una semirecta que sale del origen, entonces el mínimo se deberá alcanzar sobre la semirecta en el origen. En caso de que exista una dirección $u^k \in U_k$ que satisface $f(\lambda u^k) < f(0)$ para alguna $\lambda > 0$ entonces $f(x)$ es no acotada inferiormente sobre la semirecta en la dirección u^k. De otro modo, el mínimo de la función será alcanzado en algún vértice x^k de D_k, donde $x^k \in argmin\{f(x) : x \in V_k\}$.

Enseguida presentamos el algoritmo:

Algoritmo de aproximación exterior:

Inicialización:

Toma un n-simplex D_1 que contiene a D.

Sea V_1 el conjunto de vértices de D_1, U_1 el conjunto de direcciones extremas de D_1

Sea $I_1 = \{1, 2, ..., m\}$.

Iteración k=1,2,...:

1) Para cada $u \in U_k$ verificar si existe $\lambda > 0$ tal que $f(\lambda u) < f(0)$. Si esto ocurre para algún $u^k \in U_k$, entonces:

   Si $A_i u^k \leq 0 \ \ (i \in I)$ Fin. D es vacío o el problema no tiene una solución óptima global y $f(x)$ es no acotada inferiormente sobre cualquier semirecta paralela a u^k contenida en D.

   Sino calcular

$$i_k \in argmax\{A_i u^k : i \in I_k\} \tag{3.17}$$

   e ir a 3

2) Si no existe $u \in U_k$ tal que $f(\lambda u) < f(0)$ para alguna $\lambda > 0$, entonces encuentra $x^k \in argmin\{f(x) : x \in V_k\}$

   Si $x^k \in D$, es decir, si $A_i x^k \leq b_i (\forall i \in I_k)$ Fin. x^k es una solución óptima global.

   Sino calcular

$$i_k \in argmax\{A_i x^k - b_i : i \in I_k\} \tag{3.18}$$

   e ir a 3

3) Formar

$$D_{k+1} = D_k \cap \{x : A_{i_k} x \leq b_{i_k}\}. \tag{3.19}$$

   Calcular el conjunto de vértices V_{k+1} y el conjunto de direcciones extremas U_{k+1} de S_{k+1} a partir de V_k y U_k.

   Sea $I_{k+1} = I_k \diagdown \{i_k\}$ e ir a la iteración $k + 1$.

Ejemplo:

$$\begin{array}{llcll} \text{mín} & \dfrac{x_1x_2}{x_1+x_2} & - & 0.05(x_1+x_2) & \\ \text{s.a} & & & & \\ & -3x_1 & + & x_2 & \leq 1 \\ & -3x_1 & - & 5x_2 & \leq -23 \\ & x_1 & - & 4x_2 & \leq 2 \\ & -x_1 & + & x_2 & \leq 5 \\ & x_1 & & & \geq 0 \\ & & & x_2 & \geq 0 \end{array}$$

Inicialización.

Sea $D_1 = \mathbb{R}^2_+$, $V_1 = \{(0,0)\}$, $U_1 = \{(1,0),(0,1)\}$, $I_1 = \{1,2,3,4\}$

Iteración 1.

1) Para cada $u \in U_4$ revisar si existe un $\lambda > 0$ tal que $f(\lambda u) < f(0)$. Sea $u^1 = (1,0) \in U_1$ entonces $f(\lambda u^1) = -0.05\lambda$ así cuando $\lambda \to +\infty$ $f(\lambda u^1) \to -\infty$, además $f(0) = 0$ y $f(\lambda u^1) < f(0)$ entonces:

   a) Para $(i \in I_1)$ $A_i u^1 = -3, -3, 1, -1$ como no para todo $i \in I$ $A_i u^1 \leq 0$ entonces ir al inciso b)

   b) Calcular $i_1 \in argmax\{A_i u^1 : i \in I\}$ entonces $i_1 = 3$ e ir a 3)

3) Construir $D_2 = D_1 \cap \{x : A_{i_1}x \leq b_{i_1}\}$ es decir,

   $D_2 = \{x : x_1 \geq 0, x_2 \geq 0, x_1 - 4x_2 \leq 2\}$.

   Calcular V_2, U_2 de S_2. $I_2 = \{1,2,4\}$.

   Para poder construir los vértices de S_2 ocuparemos el método descrito en la sección anterior, para ello definimos lo siguiente:

   $l_1(x) = x_1 - 4x_2 - 2$.

   $V_1^+ = \varnothing \quad V_1^- = \{(0,0)\} \quad U_1^+ = \{(1,0)\} \quad U_1^+ = \{(0,1)\}$

   $\{U_1^- \times V_1^+\} = \{(0,1)\} \quad \{U_1^+ \times V_1^-\} = \{(1,0),(0,0)\}$.

   Sea $w = u + \alpha v$ entonces, $w = (0,0) + \alpha(1,0)$ donde: $\alpha = \dfrac{-l_1(v)}{a^1 u} = \dfrac{-(-2)}{1} = 2$ Así, el vértice generado es el $w = (2,0)$, para verificar que pertenezca a V_2 hay que construir el conjunto $I(w) = \{i : l_i(v) = 0$ $a^i u = 0,\ i \in K\}$, recordemos que $K \subset \mathbb{N}$ tal que $k \notin K$, esto quiere decir que debemos aplicar v el vector con el cual generamos a w a las restricciones de D_1 excepto a la que acabamos de introducir, en este caso $l_1(x) = x_1 - 4x_2 - 2$. En cuanto a la expresión $a^i u$ haremos el producto punto del vector de coeficientes de la correspondiente restricción con el vector u con el cual generamos a w.

v	u	$l(v)$	au
$(0,0)$	$(1,0)$	$1(0)+0(0)-0=0$	$1(1)+0(0)=1$
$(0,0)$	$(1,0)$	$0(0)+1(0)-0=0$	$0(1)+1(0)=0$

Con esto tenemos que $I(w)=\{2\}$ y $A(w)=(0\ 1)$ que tiene rango 1.

Ahora para determinar las direcciones necesitamos elementos del conjunto $U_1^- \times U_1^+$ y determinar $u=(a^1u^+)u^- - (a^1u^-)u^+$. Sea $(u^-,u^+) = ((0,1),(1,0))$

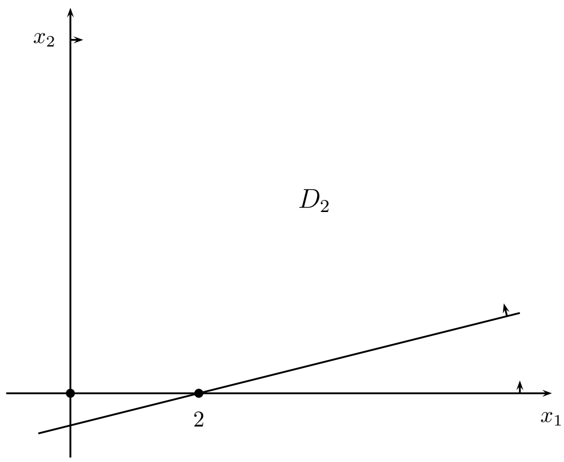
Así $u=((1\ {-4})(1\ 0)^t)(0\ 1)^t - ((1\ {-4})(0\ 1)^t)(1\ 0)^t = (4\ 1)^t$

Para saber si u pertenece a U_2 construimos $J(u)=\{j\in K : a^ju^- = a^ju^+ = 0\}$

u^-	u^+	au^-	au^+
$(0,1)$	$(1,0)$	$1(0)+0(1)=0$	$1(1)+0(0)=1$
$(0,1)$	$(1,0)$	$0(0)+1(1)=1$	$0(1)+1(0)=0$

en este caso $J(u)=\varnothing$ que tiene rango cero. Por lo tanto $V_2=\{(0,0),(2,0)\}$ y $U_2=\{(0,1),(4,1)\}$. Como $A_{i_1}u^1 = 1(1)-4(0)=1>0$ entonces $u^1=(1,0)$ no pertenece a U_2.

La región de soluciones factibles D_2 y V_2 se muestran en la siguiente gráfica:



Iteración 2.

1) Para cada $u \in U_4$ revisar si existe un $\lambda > 0$ tal que $f(\lambda u) < f(0)$. Sea $u^2 = (0,1) \in U_2$ entonces $f(\lambda u^2) = -0.05\lambda$ así cuando $\lambda \to +\infty$ $f(\lambda u^2) \to -\infty$, además $f(0) = 0$ y $f(\lambda u^2) < f(0)$ entonces:

   a) Para $(i \in I_2)$ $A_i u^2 = 1, -5, 1$ como no para todo $i \in I$ $A_i u^2 \leq 0$ entonces ir al inciso b)

   b) Calcular $i_1 \in argmax\{A_i u^1 : i \in I\}$ entonces $i_2 = 1$ e ir a 3)

3) Construir $D_3 = D_2 \cap \{x : A_{i_2} x \leq b_{i_2}\}$ es decir,

   $D_3 = \{x : x_1 \geq 0, x_2 \geq 0, x_1 - 4x_2 \leq 2, -3x_1 + x_2 \leq 1\}$.

   Calcular V_3, U_3 de S_3. $I_3 = \{2, 4\}$.

   Para poder construir los vértices de S_3 ocuparemos el método descrito en la sección anterior, para ello definimos lo siguiente:

   $l_2(x) = -3x_1 + x_2 - 1$.

$$V_2^+ = \varnothing \quad V_2^- = \{(0,0),(2,0)\} \quad U_2^+ = \{(0,1),(4,1)\} \quad U_2^+ = \varnothing$$

$$\{U_2^- \times V_2^+\} = \varnothing \quad \{U_2^+ \times V_2^-\} = \begin{array}{l} \{((0,1),(0,0)),\ ((0,1),(2,0)), \\ ((4,1),(0,0)),\ ((4,1),(2,0))\}. \end{array}$$

$(u,v) \in U_2^+ \times V_2^-$	construcción de w	α	w
$((0,1),(0,0))$	$(0,0) + 1(0,1)$	$\frac{-(-1)}{1}$	$(0,1)$
$((0,1),(2,0))$	$(2,0) + 7(0,1)$	$\frac{-(-7)}{1}$	$(2,7)$
$((4,1),(0,0))$	$(0,0) + \frac{-1}{11}(4,1)$	$\frac{-(-1)}{-11}$	$(\frac{-4}{11}, \frac{-1}{11})$
$((4,1),(2,0))$	$(2,0) - \frac{7}{11}(4,1)$	$\frac{-(-7)}{-11}$	$(\frac{-6}{11}, \frac{-7}{11})$

Enseguida veremos cual de los vértices generados pertenece a V_3.

w	v	u	$l(v)$	au
$(0,1)$	$(0,0)$	$(0,1)$	$1(0) + 0(0) - 0 = 0$	$1(0) + 0(1) = 0$
			$0(0) + 1(0) - 0 = 0$	$0(0) + 1(1) = 1$
			$1(0) - 4(0) - 2 = -2$	$1(0) - 4(1) = -4$

Para $w = (0,1)$, $I(w) = \{1\}$, $A(w) = (1\ 0)$ cuyo rango es 1. Por lo tanto $w = (0,1) \in V_3$.

w	v	u	$l(v)$	au
$(2,7)$	$(2,0)$	$(0,1)$	$1(2)+0(0)-0=2$	$1(0)+0(1)=0$
			$0(2)+1(0)-0=0$	$0(0)+1(1)=1$
			$1(2)-4(0)-2=0$	$1(0)-4(1)=-4$

Para $w=(2,7)$, $I(w)=\varnothing$, $A(w)=\varnothing$ cuyo rango es 0. Por lo tanto $w=(2,7)\notin V_3$.

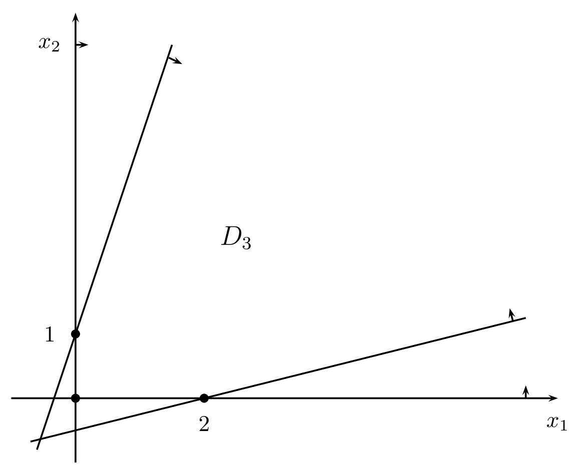
En cuanto a los vértices $(\frac{-4}{11},\frac{-1}{11})$ y $(\frac{-6}{11},\frac{-7}{11})$ aunque pudiesen satisfacer las condiciones para pertenecer a V_3 ambos no cumplen la restricción de no negatividad, por lo que podemos afirmar que no forman parte de V_3.

Ahora veamos cuales son las direcciones que forman parte de U_3.
$U_2^- \times U_2^+ = \{(0,1),(4,1)\}$ y determinar $u=(a^2u^+)u^- - (a^2u^-)u^+$. Sea $(u^-,u^+)=((4,1),(0,1))$
Así $u=((-3\ 1)(0\ 1)^t)(4\ 1)^t-((-3\ 1)(4\ 1)^t)(0\ 1)^t=(4\ 12)^t=(1\ 3)^t$
Para saber si u pertenece a U_3 construimos $J(u)=\{j\in K : a^ju^- = a^ju^+=0\}$

u^-	u^+	au^-	au^+
$(4,1)$	$(0,1)$	$1(4)+0(1)=4$	$1(0)+0(1)=0$
$(4,1)$	$(0,1)$	$0(4)+1(1)=1$	$0(0)+1(1)=1$
$(4,1)$	$(0,1)$	$1(4)-4(1)=0$	$1(0)-4(1)=-4$

en este caso $J(u)=\varnothing$ que tiene rango cero. Por lo tanto $V_3=\{(0,0),(2,0),(0,1)\}$ y $U_3=\{(4,1),(1,3)\}$. Como $A_{i_1}u^1=-3(0)+(1)=1>0$ entonces $u^2=(0,1)$ no pertenece a U_3.

En la siguiente gráfica se muestran D_3 y V_3.



Iteración 3.

1) Para cada $u \in U_3$ revisar si existe un $\lambda > 0$ tal que $f(\lambda u) < f(0)$.

   Para $f((4,1)) = \frac{4}{5} - 0.05(5) = 0.55$.

   Para $f((1,3)) = \frac{3}{4} - 0.05(5) = 0.55$.

   como no existe $\lambda > 0$ tal que $f(\lambda u) < f(0)$ entonces ir a 2.

2) Encontrar $x^3 \in argmin\{f(x) : x \in V_3\}$; $f(0,0) = 0$, $f(2,0) = -0.1$, $f(0,1) = -0.05$ entonces $x^3 = (2,0)$ e ir al inciso a.

   a) Ahora debemos sustituir x^3 en las $i \in I_3$ restricciones del problema:

   $$-3(2) - 5(0) = -6 \text{ lo cual no es menor o igual a } -23$$
   $$-(2) + (0) = -2 \text{ aquí } -2 \leq 5$$

   como x^3 no cumplió las dos restricciones anteriores, significa que no pertenece a la región de soluciones factibles construida hasta ahora, entonces ir a inciso b.

   b) Calcular $i_3 \in argmax\{A_{i_3} - b_i : i \in I_3\}$

   $$-3(2) - 5(0) + 23 = 17 \qquad -(2) + (0) - 5 = -7$$

   entonces $i_3 = 2$ e ir a 3.

3) Construir $D_4 = D_3 \cap \{x : A_{i_3}x \leq b_{i_3}\}$ es decir,

   $D_4 = \{x : x_1 \geq 0, x_2 \geq 0, x_1 - 4x_2 \leq 2, -3x_1 + x_2 \leq 1, -3x_1 - 5x_2 + 23 \leq 0\}$.

   Calcular V_4, U_4 de S_4. $I_4 = \{4\}$.

   Para poder construir los vértices de S_4 ocuparemos el método descrito en la sección anterior, para ello definimos lo siguiente:

   $l_4(x) = -3x_1 - 5x_2 + 23$.

$$V_3^+ = \{(0,0),(2,0),(0,1)\} \quad V_3^- = \varnothing \quad U_3^+ = \varnothing \quad U_3^+ = \{(4,1),(1,3)\}$$

$$\{U_3^+ \times V_3^-\} = \varnothing \quad \{U_3^- \times V_3^+\} = \begin{array}{l} \{((0,1),(0,0)),\ ((0,1),(2,0)), \\ ((0,1),(0,1)),\ ((4,1),(2,0))\} \\ ((4,1),(0,0)),\ ((4,1),(0,1))\}. \end{array}$$

$(u,v) \in U_3^- \times V_3^+$	construcción de w	α	w
$((4,1),(0,0))$	$(0,0)+\frac{23}{17}(4,1)$	$\frac{-(23)}{-17}$	$(\frac{92}{17},\frac{23}{17})$
$((4,1),(2,0))$	$(2,0)+1(4,1)$	$\frac{-(17)}{-17}$	$(6,1)$
$((4,1),(0,1))$	$(0,0)+\frac{18}{17}(4,1)$	$\frac{-(18)}{-17}$	$(\frac{72}{17},\frac{35}{17})$
$((1,3),(0,0))$	$(0,0)+\frac{23}{18}(1,3)$	$\frac{-(23)}{-18}$	$(\frac{23}{18},\frac{69}{18})$
$((1,3),(2,0))$	$(2,0)+\frac{17}{18}(1,3)$	$\frac{-(17)}{18}$	$(\frac{53}{18},\frac{51}{18})$
$((1,3),(0,1))$	$(0,1)+1(1,3)$	$\frac{-(18)}{-18}$	$(1,4)$

Enseguida veremos cual de los vértices generados pertenece a V_4.

w	v	u	$l(v)$	au
$(6,1)$	$(2,0)$	$(4,1)$	$1(2)+0(0)-0=2$	$1(4)+0(1)=4$
			$0(2)+1(0)-0=0$	$0(4)+1(1)=1$
			$1(2)-4(0)-2=0$	$1(4)-4(1)=0$
			$-3(2)+1(0)-1=-7$	$-3(4)+1(1)=-11$

Para $w=(6,1)$, $I(w)=\{1\}$, $A(w)=(1\ -4)$ cuyo rango es 1. Por lo tanto $w=(6,1)\in V_4$.

w	v	u	$l(v)$	au
$(1,4)$	$(0,1)$	$(1,3)$	$1(0)+0(1)-0=0$	$1(1)+0(3)=1$
			$0(0)+1(1)-0=1$	$0(1)+1(3)=3$
			$1(0)-4(1)-2=-6$	$1(1)-4(3)=-11$
			$-3(0)+1(1)-1=0$	$-3(1)+1(3)=0$

Para $w=(1,4)$, $I(w)=\{4\}$, $A(w)=(-3\ 1)$ cuyo rango es 1. Por lo tanto $w=(1,4)\in V_4$.

w	v	u	$l(v)$	au
$(\frac{92}{17},\frac{23}{17})$	$(0,0)$	$(4,1)$	$1(0)+0(0)-0=0$	$1(4)+0(1)=4$
			$0(0)+1(0)-0=0$	$0(4)+1(1)=1$
			$1(0)-4(0)-2=-2$	$1(4)-4(1)=0$
			$-3(0)+1(0)-1=-1$	$-3(4)+1(1)=-11$

Para $w=(\frac{92}{17},\frac{23}{17})$, $I(w)=\varnothing$, $A(w)=\varnothing$ cuyo rango es 0. Por lo tanto $w=(\frac{92}{17},\frac{23}{17})\notin V_4$.

w	v	u	$l(v)$	au
$(\frac{72}{17},\frac{35}{17})$	$(0,1)$	$(4,1)$	$1(0)+0(1)-0=0$	$1(4)+0(1)=4$
			$0(0)+1(1)-0=1$	$0(4)+1(1)=1$
			$1(0)-4(1)-2=-6$	$1(4)-4(1)=0$
			$-3(0)+1(1)-1=0$	$-3(4)+1(1)=-11$

Para $w=(\frac{72}{17},\frac{35}{17})$, $I(w)=\varnothing$, $A(w)=\varnothing$ cuyo rango es 0. Por lo tanto $w=(\frac{72}{17},\frac{35}{17})\notin V_4$.

w	v	u	$l(v)$	au
$(\frac{23}{18},\frac{69}{18})$	$(0,0)$	$(1,3)$	$1(0)+0(0)-0=0$	$1(1)+0(3)=1$
			$0(0)+1(0)-0=0$	$0(1)+1(3)=3$
			$1(0)-4(0)-2=-2$	$1(1)-4(3)=-11$
			$-3(0)+1(0)-1=-1$	$-3(1)+1(3)=0$

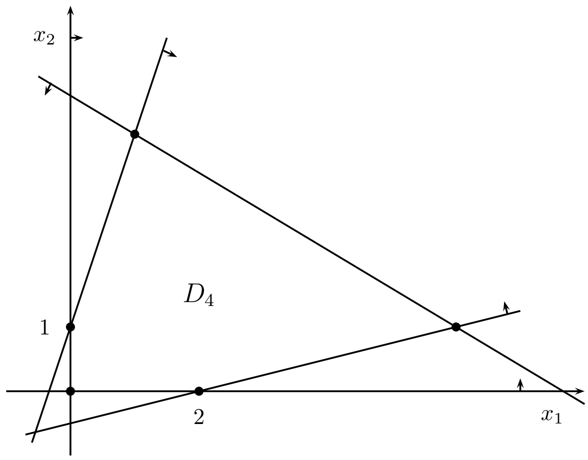
Para $w=(\frac{23}{18},\frac{69}{18})$, $I(w)=\varnothing$, $A(w)=\varnothing$ cuyo rango es 0. Por lo tanto $w=(\frac{23}{18},\frac{69}{18})\notin V_4$.

w	v	u	$l(v)$	au
$(\frac{53}{18},\frac{51}{18})$	$(2,0)$	$(1,3)$	$1(2)+0(0)-0=2$	$1(1)+0(3)=1$
			$0(2)+1(0)-0=0$	$0(1)+1(3)=3$
			$1(2)-4(0)-2=0$	$1(1)-4(3)=-11$
			$-3(2)+1(0)-1=-7$	$-3(1)+1(3)=0$

Para $w=(\frac{53}{18},\frac{51}{18})$, $I(w)=\varnothing$, $A(w)=\varnothing$ cuyo rango es 0. Por lo tanto $w=(\frac{53}{18},\frac{51}{18})\notin V_4$.

Como ninguna dirección satisface que $A_{i_3}u^3>0$ entonces $U_4=U_3$. De este modo $V_4=\{(6,1),(1,4)\}$, $I_4=\{4\}$.

En la siguiente gráfica se muestran D_4 y V_4.



Iteración 4.

1) Para cada $u \in U_4$ revisar si existe un $\lambda > 0$ tal que $f(\lambda u) < f(0)$.

   Para $f((4,1)) = \frac{4}{5} - 0.05(5) = 0.55$.

   Para $f((1,3)) = \frac{3}{4} - 0.05(5) = 0.55$.

   como no existe $\lambda > 0$ tal que $f(\lambda u) < f(0)$ entonces ir a 2.

2) Encontrar $x^4 \in argmin\{f(x) : x \in V_4\}$; $f(1,4) = \frac{4}{5} - 0.05(5) = 0.55$, $f(6,1) = \frac{6}{7} - 0.05(7) = 0.55$ entonces $x^4 = (6,1)$ e ir al inciso a.

   a) Ahora debemos sustituir x^4 en las $i \in I_4$ restricciones del problema:

   $$-1(6) + 1(1) = -5 < 5 \text{ aquí } -2 \leq 5$$

   como x^4 satisface la restricción anterior, significa que pertenece a la región de soluciones factibles construida hasta ahora. Por lo tanto $x^4 = (6,1)$ es el óptimo global.

3.2. Aproximación Interior

El método de aproximación interior es llamado así por construir una sucesión de politopos D_k con $k = 1, 2, ...$ que se aproximan desde adentro a D y satisfacen $D_1 \subset D_2 \subset \cdots \subset D$, de este modo los valores mínimos de f sobre estos conjuntos son no decrecientes, así se aproximan al valor óptimo desde arriba.

3.2.1. El problema (DG)

La anexión poliedral es una técnica originalmente concebida para resolver problemas que pueden ser formulados de la siguiente manera:

(DG) *Dado un poliedro D contenido en un cono $K_0 \subset \mathbb{R}^n$ en un conjunto convexo compacto G con $0 \in intG$, encontrar un punto $y \in D \setminus G$, o establecer que $D \subset G$*

En el capítulo anterior, para resolver un problema de programación cóncava, con $\gamma \in f(D)$ y $\epsilon > 0$ la tarea crucial era encontrar una solución factible y que satisficiera $f(y) < \gamma - \epsilon$ o bien establecer que tal punto no existía. Ahora la tarea consiste en resolver el problema (DG). Sea un vértice $x^0 \in D$ que satisface que $f(x^0) \geq \gamma$, un cono poliédrico $K_0 \supset D$, tomemos a $G = \{x : f(x) \geq \gamma - \epsilon\}$ el cual es compacto y convexo. Ahora bien, sabiendo resolver este tipo de problemas, el problema básico de programación cóncava se resuelve de la siguiente manera:

Empezando con una solución $z \in D$.

Fase I.

Buscar un mínimo local x^0 que sea un vértice de D tal que $f(x^0) \leq f(z)$.

Fase II.

Sea $\alpha = f(x^0) - \epsilon$.

Traslada el origen a x^0.

Construye un cono poliédrico $K_0 \supset D$.

Resuelve el problema (DG) para $G = \{x : f(x) \geq f(x^0) - \epsilon\}$

Si $D \subset G$ terminar; x^0 es una solución óptima global.

Sino, sea $y \in D \setminus G$ entonces $f(y) < f(x^0) - \epsilon$. Sea $z \leftarrow y$ y regresar a la Fase I.

Observación: En el algoritmo se pide trasladar el origen a x^0, esto es análogo a reescribir el problema en forma estándar con respecto a x^0.

3.2.2. Anexión Poliedral

Ahora nos ocuparemos de resolver el problema (DG) y por lo que acabamos de ver estaremos resolviendo el problema básico de programación cóncava.

Sea $D \subset K_0$, entonces $G = \{x : f(x) \geq \gamma - \epsilon\} \cap K_0$, sea P_1 un $n-$simplex.

Resolver el problema

$$(DP_1) \quad \text{Encontrar un punto } y^1 \in D\diagdown P_1$$

Si el punto y^1 no existe, entonces esto quiere decir que $D \subset G$, o bien si $y^1 \notin G$ significa que $y^1 \in D\diagdown G$ en cualquier caso terminamos.

Sino, construimos el punto z^1, el cual se genera desde el punto cero y la semirecta que pasa a través de y^1 hasta tocar la frontera de G (∂G)

Agrandar P_1 a $P_2 = conv(P_1 \cup \{z^1\})$

y repetir el método con P_2 en lugar de P_1.

La construcción de P_1 es a través del origen con los n puntos de la frontera de G cuando se intersectan con las n aristas del cono K_0.

Con este procedimiento estamos generando una sucesión de politopos cada vez más grandes que se aproximan desde adentro a $G \cap K_0$. Estos politopos $P_1 \subset P_2 \subset \cdots \subset G \cap K_0$ satisfacen

$$\begin{array}{lll} con(P_k) = K_0 & V(P_k)\diagdown\{0\} \subset \partial G & 0 \in P_k \\ y^k \in D\diagdown P_k & P_{k+1} = conv(P_k \cup \{z^k\}) & \end{array}$$

Observaciones:

1. y^k se obtiene de resolver el problema (DP_k).
2. z^k es la $G-$extensión de y^k.
3. $V(P_k)$ representa al conjunto de vértices de P_k.

Examinemos cómo es que podría terminar el algoritmo; si consideramos que $D\diagdown P_k \neq \varnothing$ entonces estaríamos diciendo que el conjunto $D\diagdown P_k$ tiene al menos un vértice de D, en caso contrario, es decir, en caso de que no haya ningún vértice en el conjunto entonces estaríamos diciendo que el conjunto de vértices de D es un subconjunto de P_k, pero si esto sucede, entonces el mismo D sería un conjunto de P_k.

Con esto en mente, para poder resolver el problema DG es necesario que el punto y^k que buscamos sea un vértice de D. De este modo cada conjunto P_{k+1} generado durante el algoritmo tiene al menos un vértice que no pertenece a ningún conjunto P_i con $i = 1, ..., k$ construido anteriormente. Ahora bien, recordemos que el número de vértices de D es finito, entonces el método antes descrito puede terminar de dos formas; la primera es habiendo generado un politopo P_h que contenga a D, concluyendo así que $D \subset G$ o bien con un vértice $y^h \in D\diagdown G$.

Para resolver el problema DP_k, el cual como ya vimos, consiste en encontrar un vértice $y^k \in D\diagdown P_k$, necesitaremos lo que se conoce como *cara de un politopo* P de dimensión $n-1$, el cual se forma de la intersección de P con un hiperplano soporte. La siguiente definición nos ayudará a encontrar dicha cara.

Definición 8. *Una cara de un politopo se dice que es transversal si su correspondiente hiperplano soporte no contiene al origen 0.*

Una cara transversal se determina a través de la expresión $vx = 1$. Ahora, usando las propiedades de los politopos sabemos que el cono de P_k es K_0, si denotamos V_k como el conjunto de todas las caras transversales de P_k entonces podremos determinar a P_k a través de las siguientes desigualdades:

$$\begin{array}{ll} vx \leq 1 & \quad (v \in V_k) \\ x \in K_0 & \end{array}$$

Con esta información generamos tanto programas lineales como elementos tenga V_k;

$$\text{máx} \quad vx \quad \text{s.a} \quad x \in D$$

denotaremos como $\mu(v)$ al valor óptimo de cada programa lineal y la siguiente proposición establece un criterio de optimalidad global.

Proposición 7. *Si $\mu(v) \leq 1$ para todo $v \in V_k$, entonces $D \subset P_k$. Si $\mu(v) > 1$ para algún $v \in V_k$, entonces cualquier solución óptima básica y^k de $LP(v, D)$ satisface $y^k \in D\diagdown P_k$.*

Con lo que hemos visto hasta el momento, el algoritmo para resolver el problema está casi hecho, sólo nos falta saber como encontrar V_k en cada iteración.

3.2.3. Calculando las caras de un politopo

El conjunto V_1 consiste de un solo elemento, el cual se genera a partir de los hiperplanos de la cara que pasan por las n intersecciones de la frontera de G con las aristas de K_0. Recordemos que el conjunto P_{k+1} se genera a partir de la envolvente convexa de $(P_k \cup \{z^k\})$, esto quiere decir que si sabemos cuales son los elementos de V_k, a partir de éste podremos obtener al conjunto V_{k+1}.

Encontrando una nueva cara
Dado un politopo P de dimensión completa en $\mathbb{R}^n$ que contiene al cero, cuyo conjunto de facetas transversales es conocida y dado un punto $z \notin P$ calcular el conjunto de facetas transversales del politopo $P' = conv(P \cup \{z\})$.

Una manera más fácil de resolver este problema es utilizando lo que vimos en la sección 3.1.2

Encontrando nuevos vértices
Dado un poliedro S de dimensión completa en $\mathbb{R}^n$ que contiene al cero en su interior, cuyo conjunto de vértices es conocido y dado un hiperplano $zx = 1$ donde z es el vector normal, calcular el conjunto de vértices del poliedro $S' = S \cap \{x : zx \leq 1\}$.

La siguiente proposición establece que cada cara transversal generada corresponde a un vértice del conjunto S.

Proposición 8. *Sea P un politopo de dimensión completa que contiene al cero, y sea $S = \{x : zx \leq 1 \quad \forall z \in P\}$ sea el polar de P. Entonces $0 \in intS$ y cada cara transversal $vx = 1$ de P corresponde a un vértice v de S y viceversa; cada cara no transversal $vx = 0$ de P corresponde a una dirección extrema v de S y viceversa.*

Corolario 3. *Sea P un politopo de dimensión completa que contiene al 0, y sea $z \notin P$. Si S denota al polar de P entonces cada cara transversal $vx = 1$ del politopo $P' = conv(P \cup \{z\})$ corresponde a un vértice v del poliedro $S' = S \cap \{x : zx \leq 1\}$ y viceversa; cada cara no transversal de P' corresponde a una dirección extrema de S' y viceversa.*

3.2.4. Un algoritmo de anexión poliedral

Algoritmo

Elegir $\epsilon \geq 0$
Inicialización:

Calcular un punto $z \in D$. Sea $M = D$.

Fase I.

Empezar con z, buscar un vértice x^0 de M que es un mínimo local de $f(x)$ sobre M.

Fase II.

0) Sea $\alpha = f(x^0) - \epsilon$. Trasladar el origen a x^0 y construir un cono K_0 que contenga a M (reescribir el problema en su forma estándar con respecto a x^0; entonces $K_0 = \mathbb{R}^n_+$).

   Para cada $i = 1, 2, ..., n$ construir la intersección u^i de la i-ésima arista de K_0 con la superficie $f(x) = \alpha$. Sea

$$S_1 = \{x : u^i x \leq 1, \quad i = 1, 2, ..., n\}$$

   Calcular $v^1 = eQ_1^{-1}$, donde $Q_1 = (u^1, u^2, ..., u^n)$. Sea $V_1 = \{v^1\}$, $V_1^* = V_1$. Sea $k = 1$ (Para $k > 1$, V_k^* es el conjunto de nuevos vértices de S_k).

1) Para cada $v \in V_k^*$ resolver el programa lineal

$$LP(v; M) \quad \text{maximizar } vx \quad \text{s.a} \quad x \in M$$

para obtener el valor óptimo $\mu(v)$ y una solución óptima básica $w(v)$.

Si tenemos $f(w(v)) < \alpha$ para algún $v \in V_k^*$, entonces sea

$$M \leftarrow M \cap \{x : v^1 x \geq 1\},$$

donde v^1 fue definido en el paso 0) y regresar a la Fase I.

Sino ir a 2)

2) Calcular $v^k \in$arg máx$\{\mu(v) : v \in V_k\}$.

Si $\mu(v^k) \leq 1$ entonces Fin x^0 es una solución óptima ϵ global de (BCP).

Sino ir a 3)

3) Sea z^k la α-extensión de $w(v^k)$. Generar

$$S_{k+1} = S_k \cap \{x : z^k x \leq 1\}$$

Calcular el conjunto de vértices V_{k+1} de S_{k+1} y sea $V_{k+1}^* = V_{k+1} \setminus V_k$. Sea $k \leftarrow k+1$ y regresar a 1).

Ejemplo:

$$\begin{array}{cccccl} \text{mín} & -(x_1 - 4.2)^2 & - & (x_2 - 1.9)^2 & \\ \text{s.a} & & & & \\ & -x_1 & + & x_2 & \leq 3 \\ & x_1 & + & x_2 & \leq 11 \\ & 2x_1 & - & x_2 & \leq 16 \\ & -x_1 & - & x_2 & \leq -1 \\ & & & x_2 & \leq 5 \\ & x_1 & & & \geq 0 \\ & & & x_2 & \geq 0 \end{array}$$

Sea $\epsilon = 0$

Inicialización.

Calcular un punto $z = (8, 0) \in D$. Sea $M = D$.

Fase I.

Empezando en $z = (8, 0)$ buscamos un vértice x^0 de M que sea un mínimo local de $f(x)$ sobre M. Para obtener a x^0 comparamos los valores en la función objetivo de los vértices adyacentes a z y x^0 será aquel que dicho valor sea el más pequeño.

v	$f(v)$
$(8,0)$	-18.05
$(1,0)$	-13.85
$(9,2)$	-23.05

así, $x^0 = (9,2)$

Fase II.

0) Sea $\alpha = -23.05$. Traslada el origen a x^0 y construye un cono K_0 que contenga a M. Al trasladar el origen a x^0 es análogo a reescribir el problema en forma estándar con respecto de x^0 por lo que ahora el problema será el siguiente:

$$\begin{array}{rrcrcl}
\text{mín} & -(x_1-13.2)^2 & - & (x_2-3.9)^2 & & \\
\text{s.a} & & & & & \\
 & -x_1 & + & x_2 & \leq & -4 \\
 & x_1 & + & x_2 & \leq & 22 \\
 & 2x_1 & - & x_2 & \leq & 32 \\
 & -x_1 & - & x_2 & \leq & -12 \\
 & & & x_2 & \leq & 7 \\
 & x_1 & & & \geq & 9 \\
 & & & x_2 & \geq & 2
\end{array}$$

y el cono K_0 sobre el cual trabajaremos está dado por: $(9,2) + \lambda_1(1,0) + \lambda_2(0,1)$.

Ahora construiremos los vectores u^i que son la intersección de la i-ésima arista de K_0 con la superficie $f(x) = \alpha$. $u^1 = (9,2) + \theta_1((9,3) - (9,2))$ como $\theta_1 = 4.226$ entonces $u^1 = (9, 6.226)$
$u^2 = (9,2) + \theta_1((10,2) - (9,2))$ como $\theta_2 = 8.61$ entonces $u^2 = (17.61, 2)$
Sea

$$S_1 = \{x : u^i x \leq 1, i = 1,2.\}$$

$$S_1 = \{x : 9x_1 + 6.226x_2 \leq 1,\ 17.61x_1 + 2x_2 \leq 1\}$$

Calculamos $v^1 = eQ_1^{-1}$ donde $Q = (u^1, u^2)$.

Como

$$\mathbf{Q_1} = \begin{pmatrix} 9 & 17.61 \\ 6.226 & 2 \end{pmatrix} \qquad \mathbf{Q_1^{-1}} = \begin{pmatrix} -0.0218 & 0.1922 \\ 0.0679 & -0.0982 \end{pmatrix}$$

entonces

$$\mathbf{v^1} = \mathbf{eQ_1^{-1}} = \begin{pmatrix} 0.0461 & 0.0940 \end{pmatrix}$$

Sea $V_1 = \{v^1\}$, $V_1^* = V_1$. Sea $k = 1$. (Para $k > 1$, V_k^* es el conjunto de nuevos vértices de S_k).

1) Para $v^1 \in V_1^*$ resolver el siguiente programa:

$$\begin{array}{rccc} \text{máx} & 0.0461x_1 & + & 0.0940x_2 \\ \text{s.a} & x \in M & & \end{array}$$

La solución a este problema es $w(v^1) = (15, 7)$ y su valor óptimo es $\mu(w(v^1)) = 1.3495$, además el valor de la solución en la función objetivo del problema es $f(w(v^1)) = -12.85 > \alpha$ entonces ir a 2.

2) Calcular $v^1 \in argmax\{\mu(v) : v \in V_1\}$, como en V_1 sólo existe $v^1 = \begin{pmatrix} 0.0461 & 0.0940 \end{pmatrix}$ y $\mu(w(v^1)) = 1.3495$ entonces ir a 3.

3) Sea z^1 la $\alpha-$extensión de $w(v^1)$. Esta $\alpha-$extensión la generamos del siguiente modo:

$$z^1 = \begin{pmatrix} 9 \\ 2 \end{pmatrix} + \theta \begin{pmatrix} 15 \\ 7 \end{pmatrix} - \begin{pmatrix} 9 \\ 2 \end{pmatrix}$$

de aquí que

$$z^1 = \begin{pmatrix} 15.9792 \\ 7.816 \end{pmatrix}$$

Construir

$$S_2 = S_1 \cap \{x : z^1 x \leq 1\}$$

$$S_2 = \{x : 9x_1 + 6.226x_2 \leq 1,\ 17.61x_1 + 2x_2 \leq 1,\ 15.9792x_1 + 7.816x_2 \leq 1\}$$

Calcular V_2 de S_2, para ello, utilizaremos el método descrito en la primer sección del presente capítulo, donde $U_1 = \{u^1 = (-0.1461, 0.2112),\ u^2 = (0.0539, -0.4745)\}$

$l_2(x) = 15.9792 * x_1 + 7.816x_2 - 1$

$$V_1^- = \varnothing \quad V_1^+ = \{(0.0461, 0.0940)\} \quad U_1^- = \{u^1, u^2)\} \quad U_1^+ = \varnothing$$

$$\begin{array}{ll} \{U_1^+ \times V_1^-\} = \varnothing \quad \{U_1^- \times V_1^+\} = & \{((-0.1461, 0.2112), (0.0461, 0.0940)), \\ & ((0.0539, -0.4745), (0.0461, 0.0940)). \end{array}$$

$$\begin{array}{rl} (u, v) \in U_1^+ \times V_1^- & ((-0.1461, 0.2112), (0.0461, 0.0940)) \\ \text{construcción de } w & (0.0461, 0.0940) + 0.6892(-0.1461, 0.2112) \\ \alpha & \frac{-(.4713)}{-.6838} \\ w & (-0.0546, 0.2396) \end{array}$$

$(u,v)\in U_1^+\times V_1^-$	$((0.0539,-0.4745),(0.0461,0.0940))$
construcción de w	$(0.0461,0.0940))+0.1655(0.0539,-0.4745)$
α	$\frac{-(.4713)}{-2.8474}$
w	$(0.0550,0.0155)$

Enseguida veremos cual de los vértices generados pertenece a V_2.

w	$(-0.0546,0.2396)$
v	$(0.0461,0.0940))$
u	$(-0.1461,0.2112)$
$l(v)$	$9(0.0461)+6.226(0.0940)-1=0$
	$17.61(0.0461)+2(0.0940)-1=0$
au	$9(-0.1461)+6.226(0.2112)=0$
	$17.61(-0.1461)+2(0.2112)=-2.15$

Para $w=(-0.0546,0.2396)$, $I(w)=\{1\}$, $A(w)=(9\ \ 6.226)$ cuyo rango es 1. Por lo tanto $w\in V_2$.

w	$(0.0550,0.0155)$
v	$(0.0461,0.0940)$
u	$(0.0539,-0.4745)$
$l(v)$	$9(0.0461)+6.226(0.0940)-1=0$
	$17.61(0.0461)+2(0.0940)-1=0$
au	$9(0.0539)+6.226(-0.4745)=-3.47$
	$17.61(0.0539)+2(-0.4745)=0$

Para $w=(0.0550,0.0155)$, $I(w)=\{2\}$, $A(w)=\{17.61\ \ 2\}$ cuyo rango es 1. Por lo tanto $w\in V_2$.

entonces $V_2^*=V_2\diagdown V_1=$ entonces $V_2^*=\{v^{21}=(-0.0546,0.2396),v^{22}=(0.0550,0.0155)\}$.

Sea $k=2$ e ir a 1.

1) Para cada $v\in V_2^*$ resolver máx vx s.a $x\in M$.

$$\text{máx } -0.0546x_1+0.2396x_2 \text{ s.a } x\in M$$

La solución a este problema es: $w(v^{21})=(11,7)$ cuyo valor óptimo es $\mu(w(v^{21}))=1.079$ y su valor en la función objetivo es $f(w(v^{21}))=-14.45>-23.05$

$$\text{máx } 0.0550x_1 + 0.0155x_2 \text{ s.a } x \in M$$

La solución a este problema es: $w(v^{22}) = (18, 4)$ cuyo óptimo es $\mu(w(v^{22})) = 1.05$ y su valor en la función objetivo es $f(w(v^{22})) = -23.05$

Como ninguna solución $w(v)$ de los respectivos problemas al sustituirlos en la función objetivo $f(w(v))$ resulto ser menor estrictamente que α entonces nos seguimos al paso 2.

2) Calcular $v^2 \in argmax\{\mu(v) : v \in V_2\}$ entonces $v^2 = v^{21} = (-0.0546, 0.2396)$ pues $\mu(w(v^{21})) = 1.079$ y $\mu(w(v^{22})) = 1.05$ entonces ir a 3.

3) Sea z^2 la $\alpha-$extensión de $w(v^2) = (11, 7)$

$$z^2 = \begin{pmatrix} 9 \\ 2 \end{pmatrix} + \theta \left[\begin{pmatrix} 11 \\ 7 \end{pmatrix} - \begin{pmatrix} 9 \\ 2 \end{pmatrix} \right]$$

cómo $\theta = 1.283$ entonces

$$z^2 = \begin{pmatrix} 11.566 \\ 8.415 \end{pmatrix}$$

Construir

$$S_3 = S_2 \cap \{x : z^1 x \leq 1\}$$

$$S_3 = \begin{cases} x : & 9x_1 + 6.226x_2 \leq 1 \\ & 17.61x_1 + 2x_2 \leq 1 \\ & 15.9792x_1 + 7.816x_2 \leq 1 \\ & 11.566x_1 + 8.415x_2 \leq 1 \end{cases}$$

Calcular V_3 de S_3

$l_3(x) = 11.566 * x_1 + 8.415x_2 - 1$

Dado que la restricción l_3 intersecta a una arista acotada de S_2, la construcción del vértice será del siguiente modo:

$$V_2^- = \{(0.055, 0.0155)\} \quad V_2^+ = \{(-0.0546, 0.2396)\}$$

$$\{V_2^- \times V_2^+\} = \quad \{((0.055, 0.0155), (-0.0546, 0.2396))$$

$$\begin{array}{rl} (v^-, v^+) \in V_2^- \times V_2^+ & ((0.055, 0.0155), (-0.0546, 0.2396)) \\ \text{construcción de } w & \alpha v^- + (1-\alpha)v^+ \\ \alpha & \dfrac{l_3(v^+)}{l_3(v^+) - l_3(v^-)} = \dfrac{0.3847}{0.6181} = 0.6224 \\ w & (0.0136, 0.1001) \end{array}$$

$V_3 = \{(0.0136, 0.1001), (0.0550, 0.0155)\}$ entonces $V_3^* = \{(0.0136, 0.1001)\}$ y la construcción de U_3 es la siguiente: Consideremos u^1 y u^2 de la iteración anterior y

$$\begin{array}{rl} (u^-, u^+) \in U_2^- \times U_2^+ & ((-0.1461, 0.2112), (0.0539, -0.4745)) \\ \text{construcción de } u & (a^3u^+)u^- - (a^3u^-)u^+ \\ u & (-0.4876, 0.6701) \end{array}$$

$U_3 = \{(-0.4876, 0.6701), (0.0539, -0.4745)\}$

Sea $k = 3$ e ir a 1.

1) Para cada $v \in V_3^*$ resolver máx vx s.a $x \in M$.

$$\text{máx } 0.0136x_1 + 0.1001x_2 \text{ s.a } x \in M$$

La solución a este problema es: $w(v^{31}) = (11, 7)$ cuyo valor óptimo es $\mu(w(v^{31})) = .5504$ y su valor en la función objetivo es $f(w(v^{31})) = -14.45 > -23.05$ ir a 2

2) Calcular $v^3 \in argmax\{\mu(v) : v \in V_3\}$ entonces $v^3 = v^{22} = (-0.0550, 0.0155)$ pues $\mu(w(v^{22})) = 1.05$ y $\mu(w(v^{31})) = .55$, el valor más grande no es menor que 1, así que ir a 3.

3) Sea z^3 la $\alpha-$extensión de $w(v^3) = (18, 4)$

$$z^3 = \begin{pmatrix} 9 \\ 2 \end{pmatrix} + \theta \left[\begin{pmatrix} 18 \\ 4 \end{pmatrix} - \begin{pmatrix} 9 \\ 2 \end{pmatrix} \right]$$

cómo $\theta = 1$ entonces

$$z^3 = \begin{pmatrix} 18 \\ 4 \end{pmatrix}$$

Construir

$$S_4 = S_3 \cap \{x : z^1x \leq 1\}$$

$$S_4 = \begin{cases} x : & 9x_1 + 6.226x_2 \leq 1 \\ & 17.61x_1 + 2x_2 \leq 1 \\ & 15.9792x_1 + 7.816x_2 \leq 1 \\ & 11.566x_1 + 8.415x_2 \leq 1 \\ & 18x_1 + 4x_2 \leq 1 \end{cases}$$

Calcular V_4 de S_4

$l_4(x) = 18 * x_1 + 4x_2 - 1$

Al haber añadido la restricción l_4 para formar S_4 nuevos vértices se generan al intersectar l_4 con aristas de S_3, en este caso, intersecto a dos; una acotada y otra no acotada, para encontrar dichos vértices necesitaremos de U_3^-, U_3^+, V_3^- y V_3^+.

$$
\begin{aligned}
U_3^- &= \{(-0.4876, 0.6701), (0.0539, -.,4745)\} \\
U_3^+ &= \varnothing \\
V_3^- &= \{(0.0136, 0.1001)\} \\
V_3^+ &= \{(0.055, 0.0155)\}
\end{aligned}
$$

Intersección con la arista acotada:

$$
\begin{aligned}
(v^-, v^+) \in V_3^- \times V_3^+ = &\quad ((0.0136, 0.1001), (0.055, 0.0155)) \\
\alpha = &\quad \frac{l_4(v^+)}{l_4(v^+) - l_4(v^-)} = \frac{-1.0520}{-1.0520 - 0.6452} = 0.6198 \\
w = &\quad \alpha v^- + (1 - \alpha)v^+ = (0.0293, 0.0679)
\end{aligned}
$$

Para la arista no acotada:

$$
\begin{aligned}
(u^-, v^+) \in U_3^- \times V_3^+ = &\quad ((-0.4876, 0.6701), (0.055, 0.0155)) \\
&\quad ((0.0539, -0.4745), (0.055, 0.0155)) \\
\alpha = &\quad \frac{-l_4(v)}{a_4(u)} = \frac{-.052}{-6.0964} = 0.0085 \\
w = &\quad v + \alpha u = (0.0509, 0.0212)
\end{aligned}
$$

$$
\begin{aligned}
\alpha = &\quad \frac{-l_4(v)}{a_4(u)} = \frac{-.052}{-.9278} = 0.056 \\
w = &\quad v + \alpha u = (0.058, -0.0111)
\end{aligned}
$$

De este modo $V_4 = \{(0.0293, 0.0679), (0.058, -0.0111), (0.0136, 0.1001)\}$ y $V_4^* = \{(0.0293, 0.0679), (0.058, -0.0111)\}$

Sea $k = 4$ e ir a 1

1) Para cada $v \in V_4^*$ resolver máx vx s.a $x \in M$.

$$\text{máx } 0.0293x_1 + 0.0679x_2 \text{ s.a } x \in M$$

La solución a este problema es: $w(v^{41}) = (15, 7)$ cuyo valor óptimo es $\mu(w(v^{41})) = .9148$ y su valor en la función objetivo es $f(w(v^{31})) = -12.85 > -23.05$

$$\text{máx } 0.058x_1 - 0.011x_2 \text{ s.a } x \in M$$

La solución a este problema es: $w(v^{42}) = (18, 4)$ cuyo valor óptimo es $\mu(w(v^{42})) = 1$ y su valor en la función objetivo es $f(w(v^{42})) = - - 23.05$ ir a 2

2) Calcular $v^4 \in argmax\{\mu(v) : v \in V_4\}$ entonces $v^4 = v^{42} = (0.0580, 0.011)$ pues $\mu(w(v^{42})) = 1$, $\mu(w(v^{41})) = .9148$ y $\mu(w(v^31)) = 0{,}55$, el valor más grande es 1, por lo tanto $x^0 = (9, 2)$ es el óptimo global.

Ejemplo:

$$\begin{array}{lccccccc}
\text{mín} & -|x_1 + \frac{1}{2}x_2 + \frac{2}{3}x_3|^{\frac{3}{2}} & & & - & x_1^2 & \\
\text{s.a} & & & & & & \\
& x_1 & + & x_2 & + & x_3 & \leq 2 \\
& x_1 & + & x_2 & - & \dfrac{1}{4}x_3 & \leq 1 \\
& -2x_1 & - & 2x_2 & + & x_3 & \leq 1 \\
& & & & & x_3 & \leq 3 \\
& x_1 & & & & & \geq 0 \\
& & & x_2 & & & \geq 0 \\
& & & & & x_3 & \geq 0
\end{array}$$

Sea $\epsilon = 0.5$

Inicialización.

Calcular un punto $z = (0, 0, 0) \in D$. Sea $M = D$.

Fase I.

Sea $x^0 = (0, 0, 0)$ un vértice de M.

Fase II.

0) Sea $\alpha = 0 - 0.5 = -0.5$. Enseguida hay que trasladar el origen a x^0, pero $x^0 = (0, 0, 0)$ el problema se queda igual. El cono que contiene a M es $K_0 = \mathbb{R}^3_+$.

Construir u^i con $i = 1, 2, 3$, la i-ésima arista de K_0 que intersecta con la superficie $f(x) = \alpha$.

$$u^1 = \theta_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad u^2 = \theta_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad u^3 = \theta_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\theta_1 = 0.448 \quad \theta_2 = 1.26 \quad \theta_3 = 0.945$$

los vectores u^i $i = 1, 2, 3.$ serán las columnas de la matriz Q.

$$u^1 = \begin{pmatrix} 0.448 \\ 0 \\ 0 \end{pmatrix} \qquad u^2 = \begin{pmatrix} 0 \\ 1.26 \\ 0 \end{pmatrix} \qquad u^3 = \begin{pmatrix} 0 \\ 0 \\ 0.945 \end{pmatrix}$$

Ahora construimos al conjunto $S_1 = \{x : v^i x \leq 1\}$

$$S_1 = \{x : 0.448x_1 \leq 1,\ 1.26x_2 \leq 1,\ 0.945x_3 \leq 1\}$$

Calcular $v^1 = eQ_1^{-1}$

Sean:

$$Q_1 = \begin{pmatrix} 0.448 & 0 & 0 \\ 0 & 1.26 & 0 \\ 0 & 0 & 0.945 \end{pmatrix} \qquad Q_1^{-1} = \begin{pmatrix} 2.2321 & 0 & 0 \\ 0 & 0.7937 & 0 \\ 0 & 0 & 1.0582 \end{pmatrix}$$

así $v^1 = eQ_1^{-1} = \begin{pmatrix} 2.2321 & 0.7937 & 1.0582 \end{pmatrix}$ Sea $V_1 = \{v^1\}$, $V_1^* = V_1$

1) Para cada $v \in V_1^*$ resolver $\max v^1 x$ s.a $x \in M$

$$\max\ 2.2321x_1 + 0.7937x_2 + 1.0582 \text{ s.a } x \in M$$

La solución a este problema es: $w(v^1) = (1.2, 0, 0.8)$ cuyo valor óptimo es $\mu(w(v^1)) = 3.525$ y su valor en la función objetivo es $f(w(v^1)) = -3.722 < -0.5$ entonces la región de soluciones factibles es:

$$M \leftarrow M \cap \{x : v^1 x \geq 1\}$$
$$M \leftarrow M \cap \{x : 2.2321x_1 + 0.7937x_2 + 1.0582x_3 \geq 1\}$$

e ir a la fase I.

Fase I.

Sea $x^0 = (1.2, 0, 0.8)$ con $f(x^0) = -3.722$

0) Sea $\alpha = -3.722 - 0.5 = -4.222$ y trasladar el origen a x^0, lo cual es análogo a reescribirlo en forma estándar con respecto a x^0, por lo que ahora el problema es:

$$\begin{array}{ll} \min & -|x_1 + \frac{1}{2}x_2 + \frac{2}{3}x_3 - 1.73|^{\frac{3}{2}} \quad - \quad (x_1 - 1.2)^2 \\ \text{s.a} & \end{array}$$

$$\begin{array}{rcrcrl}
x_1 & + & x_2 & + & x_3 & \leq 4 \\
x_1 & + & x_2 & - & \frac{1}{4}x_3 & \leq 2 \\
-2x_1 & - & 2x_2 & + & x_3 & \leq -.6 \\
 & & & & x_3 & \leq 3.8 \\
2.2321x_1 & + & 0.7937x_2 & + & 1.0582x_3 & \geq 4.52508 \\
x_1 & & & & & \geq 1.2 \\
 & & x_2 & & & \geq 0 \\
 & & & & x_3 & \geq 0.8
\end{array}$$

El cono $K_0 = x^0 + \lambda_i e^i$ con $i = 1, 2, 3$.

$$u^1 = \begin{pmatrix} 1.2 \\ 0 \\ 0.8 \end{pmatrix} + \theta_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad u^2 = \begin{pmatrix} 1.2 \\ 0 \\ 0.8 \end{pmatrix} + \theta_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad u^3 = \begin{pmatrix} 1.2 \\ 0 \\ 0.8 \end{pmatrix} + \theta_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\theta_1 = 1.5265 \quad \theta_2 = 5.22 \quad \theta_3 = 3.915$$

los vectores u^i $i = 1, 2, 3.$ serán las columnas de la matriz Q.

$$u^1 = \begin{pmatrix} 2.7265 \\ 0 \\ 0.8 \end{pmatrix} \qquad u^2 = \begin{pmatrix} 1.2 \\ 5.22 \\ 0.8 \end{pmatrix} \qquad u^3 = \begin{pmatrix} 1.2 \\ 0 \\ 4.715 \end{pmatrix}$$

Ahora construimos al conjunto $S_1 = \{x : u^i x \leq 1\}$

$$S_1 = \{x : 2.7265x_1 + 0.8x_3 \leq 1,\ 1.2x_1 + 5.22x_2 + 0.8x_3 \leq 1,\\ 1.2x_1 + 4.715x_3 \leq 1\}$$

Calcular $v^1 = eQ_1^{-1}$

Sean:

$$Q_1 = \begin{pmatrix} 2.7265 & 1.2 & 1.2 \\ 0 & 5.22 & 0 \\ 0.8 & 0.8 & 4.715 \end{pmatrix} \qquad Q_1^{-1} = \begin{pmatrix} 0.3964 & -0.0757 & -0.1009 \\ 0 & 0.1916 & 0 \\ -0.0673 & -0.0197 & .2292 \end{pmatrix}$$

así $v^1 = eQ_1^{-1} = \begin{pmatrix} .3291 & 0.0962 & .1283 \end{pmatrix}$ Sea $V_1 = \{v^1\}$, $V_1^* = V_1$

1) Para cada $v \in V_1^*$ resolver $\max v^1 x$ s.a $x \in M$

$$\max\ 0.3291x_1 + 0.0962x_2 + 0.1283 \ \text{ s.a } \ x \in M$$

La solución a este problema es: $w(v^1) = (2.4, 0, 1.6)$ cuyo valor óptimo es $\mu(w(v^1)) = 0.9951$ y su valor en la función objetivo es $f(w(v^1)) = -3.786 > -4.22$ entonces ir a 2

2) Calcular $v^1 \in argmax\{\mu(v) : v \in V_1\}$, el único vértice en V_1 es v^1 cuyo valor óptimo en el programa lineal es $\mu(w(v^1)) = 0.9951 < 1$ por lo tanto $x^0 = (1.2, 0, 0.8)$ con $f(x^0) = -3.72$ es el óptimo global.

Como pudimos observar, en los algoritmos de aproximación sucesiva presentados, se hace uso de la construcción de los nuevos conjuntos de vértices y direcciones extremas que definen a la región factible de los subproblemas en cuestión, tarea que resulta un tanto laboriosa. Para mas detalles sobre la teoría y las demostraciones de los teoremas ver [8].

4

Métodos de partición sucesiva

El presente capítulo está dedicado al estudio de los métodos de partición sucesiva, también conocidos como métodos de ramificación y acotamiento. Dicho capítulo está dividido en cuatro secciones; en la primera se dará la noción general del proceso principal de ramificación y acotamiento y en las tres secciones siguientes se desarrollarán tres métodos; el algoritmo cónico, el simplicial y el rectangular, con sus respectivos ejemplos.

4.1. Ramificación y Acotamiento

El problema de minimización cóncava que deseamos resolver es el siguiente:

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{s.a} & \\ & x \in D \end{array} \qquad \text{(P)}$$

donde $f : A \to \mathbb{R}$. Supondremos que $D \subset A \subset \mathbb{R}^n$, también supondremos que $\text{mín}\, f(D)$ existe.

En general el método de ramificación y acotamiento procede del siguiente modo:

Se empieza con un conjunto factible relajado $M_0 \supset D$, el cual conforme avancen las iteraciones será dividido en subconjuntos finitos M_i $i \in I$, este proceso de división es lo que se conoce como proceso de ramificación y consistirá de una partición, la cuál definiremos más adelante. Ahora bien, cada subconjunto M_i tendrá asignada su cota inferior denotada por $\beta(M_i)$ y su cota superior $\alpha(M_i)$, siempre que esto sea posible. Dichas cotas deben satisfacer:

$$\beta(M_i) \leq \text{ínf}\, f(M_i \cap D) \leq \alpha(M_i)$$

Considerando que $\beta := \text{mín}_{i\in I}\,\beta(M_i)$, $\alpha := \text{mín}_{i\in I}\,\alpha(M_i)$ sobre todas las cotas, entonces tenemos: $\beta \leq \text{mín}\, f(D) \leq \alpha$.

Estas cotas nos ayudarán a determinar un criterio de optimalidad. Si $\alpha = \beta$ (o si $\alpha - \beta \leq \epsilon$ para una $\epsilon > 0$ dada), esto indica que hemos encontrado la solución óptima global. Sino, entonces debemos elegir alguno de los subconjuntos M_i, dividirlo para así obtener una partición más refinada de M_0. Con esto se determinaran nuevas y mejores cotas sobre los nuevos elementos de la partición y el proceso se repite hasta hallar la solución óptima.

Este proceso tiene la ventaja de que es posible deshacernos de partes de la región factible que sabemos de acuerdo a los criterios establecidos que no contienen a la solución óptima. Desafortunadamente es posible que durante el proceso se encuentre de manera pronta a un punto factible candidato a ser el óptimo, el cual podría ser considerado como tal después de varias iteraciones, ésto debido a la dependencia de la cotas actuales, ya que la exactitud de la solución aproximada es medida a través de la diferencia $\alpha - \beta$.

Definición 9. *Sea B un subconjunto de $\mathbb{R}^n$ e I un conjunto finito de índices. Un conjunto $\{M_i : i \in I\}$ de subconjuntos de B se dice que es una partición de B si*

$$B = \cup_{i\in I} M_i \quad y \quad M_i \cap M_j = \partial M_i \cap \partial M_j \quad \forall i, j \in I,\; i \neq j$$

donde ∂M_i denota la frontera (relativa) de M_i

Conjuntos partición

Los conjuntos más comunes que son sometidos a una partición dentro del proceso son conos poliédricos, simplices y rectángulos. En las próximas tres secciones se desarrollarán los algoritmos para cada tipo de conjunto partición respectivamente, mostrando como se lleva a cabo la partición.

Definición 10. *Un conjunto partición M que satisface $M \cap D = \varnothing$ es llamado no factible; un conjunto partición M que satisface $M \cap D \neq \varnothing$ es llamado factible. Un conjunto partición M se dice que es incierto si no sabemos cuando M es factible o no factible.*

Aquí es importante destacar que durante el proceso se considera que el ínfimo y el mínimo sobre el conjunto vacío es igual a $+\infty$.

Algoritmo general de Ramificación y Acotamiento

Paso 0 (Inicialización)

Escoger $M_0 \supseteq D$, $S_{M_0} \subset D$, $-\infty < \beta_0 \leq \text{mín}\, f(D)$.

Sea $\mathcal{M}_0 = \{M_0\}$, $\alpha_0 = \text{mín}\, f(S_{M_0})$, $\beta(M_0) = \beta_0$.

Si $\alpha_0 < \infty$ entonces escoge $x^0 \in$ argmin $f(S_{M_0})$ (es decir, $f(x^0) = \alpha_0$).

Si $\alpha_0 - \beta_0 = 0$ entonces FIN: $\alpha_0 = \beta_0 = \text{mín}\, f(D)$, x^0 es una solución óptima.

Sino, ir al Paso 1.

Paso k (k=1,2,...)

Al inicio del Paso k tenemos la partición $\mathcal{M}_k$ de un subconjunto de M_0 que es la parte de la región factible que contiene a la solución óptima. Más aún, para todo $M \in \mathcal{M}_{k-1}$ tenemos $S_M \subseteq M \cap D$ y cotas $\beta(M)$, $\alpha(M)$ que satisfacen:

$$\beta(M) \leq \text{ínf}\, f(M \cap D) \leq \alpha(M) \quad \textit{Si se sabe que M es factible} \tag{4.1}$$
$$\beta(M) \leq \text{ínf}\, f(M) \quad \textit{Si M es incierta} \tag{4.2}$$

Además, tenemos la actual cota inferior y superior β_{k-1}, α_{k-1} que satisfacen

$$\beta_{k-1} \leq \text{mín}\, f(D) \leq \alpha_{k-1}$$

Finalmente, si $\alpha_{k-1} < \infty$ entonces tenemos un punto $x^{k-1} \in D$ que satisface $f(x^{k-1}) = \alpha_{k-1}$ (el mejor punto factible obtenido hasta ahora).

k.1 Eliminar todo $M \in \mathcal{M}_{k-1}$ que satisface $\beta(M) \geq \alpha_{k-1}$. Sea $\mathcal{R}_k$ la colección de los conjuntos restantes en la partición $\mathcal{M}_{k-1}$.

k.2 Elegir una colección no vacía de conjuntos $\mathcal{P}_k \subset \mathcal{R}_k$ y construir una partición de todos los elementos de $\mathcal{P}_k$. Sea $\mathcal{P}'_k$ la colección de los nuevos elementos de la partición.

k.3 Eliminar todo $M \in \mathcal{P}'_k$ para los cuales se sabe que $M \cap D = \varnothing$ o si se sabe que el mín $f(D)$ no se alcanza ahí. Sea $\mathcal{M}'_k$ la colección de los elementos restantes de $\mathcal{P}'_k$.

k.4 Asignar a cada $M \in \mathcal{M}'_k$ un conjunto S_M y una cantidad $\beta(M)$ que satisface $S_M \subseteq M \cap D$

$$\beta(M) \leq \text{ínf}\, f(M \cap D) \leq \alpha(M) \quad \textit{Si se sabe que M es factible}$$
$$\beta(M) \leq \text{ínf}\, f(M) \quad \textit{Si M es incierta}$$

Más aún, se requiere

$$S_M \supseteq M \cap S_M, \quad \beta(M) \geq \beta(M') \quad \textit{cuando } M \subset M' \in \mathcal{M}_{k-1}$$

Sea $\alpha(M) = \text{mín}\, f(S_M)$

k.5 Sea $\mathcal{M}_k$=($\mathcal{R}_k \setminus \mathcal{P}_k$)$\cup$ $\mathcal{M}'_k$.

Calcular: $\alpha_k = \text{ínf}\ \{\alpha(M) : M \in \mathcal{M}_k\}$ y $\beta_k = \text{mín}\{\beta(M) : M \in \mathcal{M}_k\}$

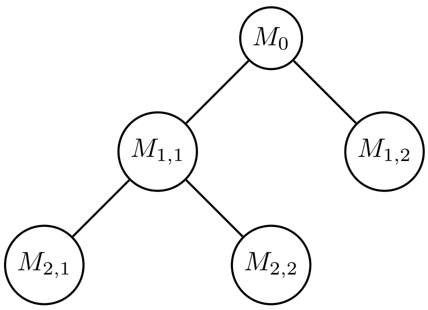
Si $\alpha_k < \infty$, entonces sea $x^k \in D$ tal que $f(x^k) = \alpha_k$.

k.6 Si $\alpha_k - \beta_k = 0$ FIN: $\alpha_k = \beta_k = \min f(D)$, x^k es una solución óptima. Sino ir al paso $k+1$.

Observaciones sobre el algoritmo

La condición que satisfacen los conjuntos partición M para que sean eliminados en el paso k es que $\beta(M) \geq \alpha_{k-1}$, de este modo cuando llegamos a que $\alpha_k = \beta_k$ es por que todos los elementos partición han sido eliminados. Ahora bien, cada vez que el algoritmo nos lleva al paso k es por que aun hay elementos de la región factible por dividir.

Una manera gráfica de visualizar el procedimiento es a través de un árbol binario como se muestra a continuación.



donde el nodo M_0 representa la región factible original y cada rama que sale de algún nodo, representa la partición de la región que parece prometedora.

Finitud y Condiciones de Convergencia

Como se pudo apreciar en las observaciones anteriores, que el algoritmo sea finito o convergente depende del límite de la diferencia $\alpha_k - \beta_k$, para garantizar tales condiciones aquí sólo enunciaremos los resultados que lo sustentan pero las pruebas pueden ser consultadas en [Hosrt, Tuy 1995] pág 126.

Consideremos las sucesiones anidadas decrecientes de elementos partición refinados sucesivamente, es decir, sucesiones $\{M_{k_q}\}$ tales que

$$M_{k_q} \in P_{k_q} \quad M_{k_{q+1}} \subset M_{k_q}.$$

Definición 11. *Una operación acotamiento es llamada* finitamente consistente *si, en cada paso, cualquier elemento partición no eliminado puede ser refinado y si cualquier sucesión decreciente M_{k_q} de elementos partición refinados sucesivamente es finita.*

Teorema 5. *En el método de ramificación y acotamiento supongamos que la operación acotamiento es finitamente consistente. Entonces el procedimiento termina en un número finito de pasos.*

Corolario 4. *Si el procedimiento de ramificación y acotamiento es finito entonces se generan al menos una infinidad de sucesiones decrecientes* $\{M_{k_q}\}$ *de elementos partición refinados sucesivamente.*

Definición 12. *Una operación acotamiento es llamada* consistente *si en cada paso cualquier elemento partición no eliminado puede ser refinado y si cualquier sucesión infinita decreciente* $\{M_{k_q}\}$ *de elementos partición sucesivamente refinados satisface*

$$\lim_{q\to\infty}(\alpha_{k_q} - \beta(M_{k_q})) = 0 \tag{4.3}$$

Definición 13. *Una operación selección es llamada* completa *si para todo* $M \in \cup_{p=1}^{\infty} \cap_{k=p}^{\infty} R_k$ *tenemos* $\inf f(M \cap D) \geq \alpha := \lim_{k\to\infty} \alpha_k$

es decir, que cualquier porción del conjunto factible que se ha dejado sin explorar para siempre, al finalizar el algoritmo no debe ser mejor que las partes eliminadas.

Teorema 6. *En un procedimiento de ramificación y acotamiento infinito supongamos que la operación acotamiento es consistente y la operación selección es completa. Entonces*

$$\alpha := \lim_{k\to\infty} \alpha_k = \lim_{k\to\infty} f(x^k) = \min f(D)$$

Teorema 7. *Una operación de selección se dice que es de acotamiento mejorado si al menos cada vez después de un número finito de pasos,* P_k *satisface la relación*

$$P_k \cap argmin\{\beta(M) : M \in R_k\} \neq \varnothing \tag{4.4}$$

es decir, al menos un elemento partición donde la actual cota inferior es alcanzada es seleccionado para más particiones en el paso k *del algoritmo.*

Teorema 8. *En el procedimiento de ramificación y acotamiento infinito supongamos que la operación acotamiento es consistente y la operación selección es de acotamiento mejorado. Entonces el procedimiento es convergente:*

$$\alpha := \lim_{k\to\infty} \alpha_k = \lim_{k\to\infty} f(x^k) = \min f(D) = \lim_{k\to\infty} \beta_k = \beta$$

Corolario 5. *Sea* $f : \mathbb{R}^n \to \mathbb{R}$ *una función continua,* D *un conjunto cerrado y* $C(x^0) := \{x \in D : f(x) \leq f(x^0)\}$ *es acotado. En un procedimiento infinito de ramificación y acotamiento supongamos que la operación acotamiento es consistente y la operación selección es completa. Entonces todo punto de acumulación de* $\{x^0\}$ *resuelve el problema* (P).

Definición 14. *Una operación de acotamiento inferior es llamada fuertemente consistente si en cada paso cualquier elemento partición no eliminado puede ser refinado y si cualquier sucesión infinita decreciente* $\{M_{k_q}\}$ *de elemento partición sucesivamente refinados posee una subsucesión* $\{M_{k_q}\}$ *satisface:*

$$\bar{M} \cap D \neq \varnothing \qquad \beta(M_{k_{q'}})_{q'\to\infty} \to \min f(\bar{M} \cap D)$$
$$\textit{donde} \qquad \bar{M} = \cap_q M_{k_q}.$$

4.1.1. Cotas Inferiores

Hay varias formas de encontrar cotas inferiores para el proceso de partición sucesiva, algunas mediante el uso de subfuncionales convexos, mediante dualidad o bien a través de vértices mínimos [Horst, Tuy 1995] pág 145. Siendo éste último método el que usaremos durante todo el capítulo. Básicamente para hallar las cotas inferiores será necesario construir un programa lineal, el cual al ser resuelto, su valor óptimo nos proporcionará dicha cota.

Vértices mínimos

Dado que M es un politopo y la función con la que trabajamos es cóncava, podremos construir cotas inferiores para $\text{ínf } f(M \cap D)$ o para $\text{ínf } f(M)$ si minimizamos a otra función que esté que coincida con f sobre el conjunto de vértices $V(M)$ de M, de este modo

$$\text{ínf } f(M \cap D) \geq \text{mín } f(M) = \text{mín } f(V(M)) \tag{4.5}$$

escogiendo así a $\beta(M) = \text{mín } f(V(M))$.

Si lo que deseamos es encontrar cotas más ajustadas podríamos usar el método de relajación de aproximación exterior a través de corte, de este modo el politopo que se construye en cada iteración denotado por P satisfará que $M \cap D \subseteq P \subset M$ de aquí que

$$\text{ínf } f(D \cap M) \geq \text{mín } f(V(M)) \geq \text{mín } f(M) \tag{4.6}$$

Así $\beta(M) = \text{mín } f(V(P))$ es una mejor cota que $\text{mín } f(V(M))$.

4.2. Algoritmos cónicos

Como una herramienta para la estructura del algoritmo es necesario que consideremos el siguiente problema:

(DG) *Dado un politopo D contenido en un cono $K_0 \subset \mathbb{R}^n$ y un conjunto convexo compacto G con $0 \in intG$, encontrar un punto $y \in D \setminus G$, o establecer que $D \subset G$.*

Si $p(x)$ es una función convexa positivamente homogénea tal que $G = \{x : p(x) \leq 1\}$, entonces $D \subset G$ es equivalente á maximizar $p(D) \leq 1$, en este caso, resolver el problema (DG) se reduce a resolver el problema de maximización convexa:

$$maximizar \ \ p(x) \ \ s.a \ \ x \in D \tag{4.7}$$

Para construir un procedimiento basado en conos para resolver el problema 4.7, a través del esquema de ramificación y acotamiento es necesario determinar tres operaciones básicas: ramificar, acotar, y elegir.

1. **Ramificar (subdivisión cónica).** Suponemos que un cono es de la forma $K = con(Q)$ con $Q = (z^1, z^2, ..., z^n)$, $z^i \in \partial G$ (la frontera de G), $i = 1, 2, ..., n$. Dado tal cono, una subdivisión de K es determinar el punto $u \in K$ tal que $u \neq \lambda z^i \ \forall \lambda \geq 0$, $i = 1, 2, ..., n$. Si $u = \Sigma_{i \in I} \lambda_i z^i$ $(\lambda_i > 0)$ y $\bar{u}$ es el punto donde la semirecta desde 0 a través de u encuentra ∂G, entonces la partición de K con respecto a u consiste de los subconos $K_i = con(Q_i)$, $i \in I$ con $Q_i = (z^1, ..., z^{i-1}, \bar{u}, z^{i+1}, ..., z^n)$. Así, para determinar la operación de ramificación, se debe especificar una regla para asignar a cada cono $K = con(Q)$, $Q = (z^1, z^2, ...z^n)$, un punto $u(Q) \in K$ el cual no caiga en ninguna arista de K.

2. **Acotamiento.** Para cualquier cono $K = con(Q)$, $Q = (z^1, z^2, ...z^n)$, el hiperplano $eQ^{-1}x = 1$ pasa a través de $(z^1, z^2, ...z^n)$, es decir la función lineal $h(x) = eQ^{-1}x$ concuerda con $p(x)$ en $(z^1, z^2, ...z^n)$. Así, $h(x) \geq p(x)$ para todo $x \in K$ y el valor $\mu(Q) = \max\{eQ^{-1}x : x \in K \cap D\}$ satisfará $\mu(Q) \geq \max p(K \cap D)$, es decir, $\mu(Q)$ es una cota superior para $p(K \cap D)$.

3. **Elección.** La regla más simple es elegir el cono $K = con(Q)$ con la $\mu(Q)$ más larga de entre todos los conos de interés actual.

Una condición suficiente para asegurar la convergencia del algoritmo, es a través de la consistencia de la operación de acotamiento, a la cual es posible recurrir debido a que estamos frente a una selección de mejoramiento de cota.

Consideremos un cono $K = con(Q)$ y sea $\omega(Q)$ la solución óptima del programa lineal

$$\max \quad eQ^{-1}x \quad \text{s.a} \quad x \in D, \quad Q^{-1}x \geq 0$$

Observemos que por ser un programa lineal, su solución es un vértice de $D \cap K$ que satisface $eQ^{-1}\omega(Q) = \mu(Q)$ el valor óptimo del programa.

Si ahora consideramos una sucesión infinita anidada de cono $K_s = con(Q_s)$, $Q_s = (z^{s1}, z^{s2}, ..., z^{sn})$ con $s = 1, 2, ...$ generados a través de la división de conos descrita. Para cada s sea $\omega^s = \omega(Q_s)$, $u^s = u(Q_s)$ donde $u^s \neq \omega^s$ y denotemos por q^s y $\bar{\omega}^s$ los puntos donde la recta que va del cero a través de ω^s intersecta al simplex $[z^{s1}, z^{s2}, ..., z^{sn}]$ y a la frontera de G ∂G respectivamente. La siguiente definición nos ayudará a establecer la condición de normalidad en el proceso de partición de un cono.

Definición 15. *Una sucesión* $K_s = con(Q_s)$ *con* $s = 1, 2, ...$ *se dice que es normal para una conjunto dado* $D \diagdown G$ *si* $\liminf_{s\to\infty||q^s-\bar{\omega}^s||=0}$.

Un proceso de división de un cono se dice que es normal si cualquier sucesión anidada de conos que genera es normal.

Por lo mencionado anteriormente, la condición de normalidad para el proceso de división de un cono es suficiente para la consistencia de la operación acotamiento, de este modo se asegura la convergencia del proceso.

4.2.1. Subrutina principal

Procedimiento (DG):

Elegir una regla $u : Q \to u(Q)$ para la operación de división de los conos que genere un proceso de subdivisión normal cónico.

1. Calcular las intersecciones $z^{01}, z^{02}, ..., z^{0n}$ de las aristas de K_0 con ∂G. Sea $Q_0 = (z^{01}, z^{02}, ..., z^{0n})$, $M = \{Q_0\}$, $P = M$.

2. Para cada matriz $Q \in P$ resolver el programa lineal $LP(Q, D)$ para obtener el valor óptimo $\mu(Q)$ y una solución óptima básica $w(Q)$ de este programa.

   Si $w(Q) \notin G$ para alguna Q entonces terminar: $y = w(Q)$.

   Sino, $w(Q) \in G$ para todo $Q \in P$; entonces ir a 3.

3. En M borrar todo $Q \in P$ tal que $\mu(Q) \leq 1$. Sea R la colección de matrices restantes.

   Si $R = \varnothing$ terminar: $D \subset G$.

   Sino, $R \neq \varnothing$; ir a 4.

4. Elegir $Q^* \in argmax\{\mu(Q) : Q \in R\}$ y dividir $K^* = con(Q^*)$ con respecto a $u^* = u(Q^*)$. Sea P^*la colección de matrices correspondientes a la partición de Q^*.

5. Reemplazar Q^* por P^* en R y denotar por M^* la colección restante de matrices. Sea $P \leftarrow P^*$, $M \leftarrow M^*$ y regresar a 2.

M

4.2.2. Proceso básico de subdivisión normal cónico

Sea $K = con(Q)$ cualquier cono con $Q = (z^1, z^2, ..., z^n)$ donde $z^i \in \partial G$ es generado durante el proceso.

$\omega(Q)$ es la solución óptima del programa lineal asociado a Q.

$u(Q)$ es el punto con respecto del cual se divide K.

$Z(Q)$ $(n-1)-$simplex que es la sección de K a través del hiperplano $H_0 = \{x : eQ_0^{-1}x = 1\}$ a través de $z^{01}, z^{02}, ..., z^{0n}$

Cuando K es subdividido con respecto a $\omega(Q)$ $(u(Q) = \omega(Q))$ lo que se está realizando es una $\omega-$subdivisión.

Para la realización del proceso [8] establecen que es posible realizar dos métodos para la subdivisión de los conos; la $\omega-$subdivisión o bien la bisección, pero hacerlo de manera separada provoca que el primero sea un procedimiento normal siempre y cuando no sea degenerado. Por otro lado, si sólo se usara la bisección, éste por no depender de los datos del problema converge de manera muy lenta. Así que para poder encontrar un procedimiento que sea convergente y eficiente mezclaron ambos del siguiente modo:

1 Elegir una sucesión infinita creciente $\Delta \subset \{0, 1, ...\}$.

2 Sea $\tau(K_0) = 0$ para el cono inicial $K_0 = con(Q_0)$. En cada iteración, hay un índice $\tau(K)$ para el cono $K = con(Q)$ que será subdividido.

  a Si $\tau(K) \notin \Delta$ hacer una $\omega-$subdivisión de $K = con(Q)$, es decir, elegir $u(Q) = \omega(Q)$. Sea $\tau(K') = \tau(K)+1$ para todo subcono $K' = con(Q')$ en la partición.

  b) Sino si $\tau(K) \in \Delta$, hacer la bisección de $K = con(Q)$, es decir, $u(Q)$ será el punto medio de la arista más larga de $Z(Q)$. Sea $\tau(K') = \tau(K) + 1$ para todo subcono $K' = con(Q')$ en la partición.

4.2.3. Algoritmo cónico normal

Consideremos el problema de programación cóncava básico:

$$minimizar\ f(x) \tag{4.8}$$
$$s.a\ \ Ax \leq b \tag{4.9}$$
$$x \geq 0 \tag{4.10}$$

donde suponemos que las restricciones definen un politopo D y la función cóncava $f : \mathbb{R}^n \to \mathbb{R}$ tiene conjuntos de nivel acotados.

Algoritmo cónico normal para programas cóncavos básicos

Elegir $\epsilon \geq 0$.

Inicialización:
Calcular un punto $z \in D$. Elegir $M = D$.

Fase I.

Empezar con z encontrar un vértice x^0 de D tal que $f(x^0) \leq f(z)$. Sea $\bar{x}$ el mejor de entre x^0 y todos los vértices de D adyacentes a x^0; sea $\gamma = f(\bar{x})$.

Fase II.

Elegir una sucesión creciente infinita Δ de números naturales.

0) Sea $\alpha = \gamma - \epsilon$. Trasladar el origen a x^0 y construir un cono $K_0 \supset D$. Para cada $i = 1, 2, ..., n$ calcular el punto z^{01} donde la i-ésima arista de K_0 encuentra la superficie $f(x) = \alpha$. Sea $Q_0 = (z^{01}, z^{02}, ..., z^{0n})$, $\mathcal{M} = \{Q_0\}$, $P = \mathcal{M}$, $\tau(Q_0) = 0$.

1) Para cada $Q = (z^1, ..., z^n) \in P$ resolver

$$LP(Q, M) \quad \text{máx}\, eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

para obtener el valor óptimo $\mu(Q)$ y una solución óptima básica $w(Q)$.

Si para algún Q, $f(w(Q)) < \gamma$, sea $z \leftarrow w(Q)$,

$$M \leftarrow M \cap \{x : eQ_0^{-1}x \geq 1\}$$

y regresar a la Fase I.

Sino ir a 2).

2) En $\mathcal{M}$ borrar todos los $Q \in P$ que satisfacen que $\mu(Q) \leq 1$. Sea $\mathcal{R}$ la colección de matrices restantes.

Si $\mathcal{R} = \varnothing$, terminar: $\bar{x}$ es una solución óptima ϵ-global para el programa cóncavo básico.

Sino si $\mathcal{R} \neq \varnothing$ ir a 3).

3) Elegir $Q_* \in argmax\{\mu(Q) : Q \in R\}$

   a) Si $\tau(Q_*) \notin \Delta$ entonces dividir Q_* con respecto a $w(Q_*)$ y sea $\tau(Q) = \tau(Q_*) + 1$ para cada miembro Q de la partición.
   b) Sino, dividir Q^* con respecto al punto medio de la arista más larga de $Z(Q_*)$ y sea $\tau(Q) = \tau(Q_*) + 1$ para cada miembro Q de la partición.

4) Sea $\mathcal{P}_*$ la partición de Q_*, $\mathcal{M}_*$ la colección obtenida de $\mathcal{R}$ por haber reemplazado Q_* por $\mathcal{P}_*$. Sea $\mathcal{P} \leftarrow \mathcal{P}_*$, $\mathcal{M} \leftarrow \mathcal{M}_*$ y regresar a 1).

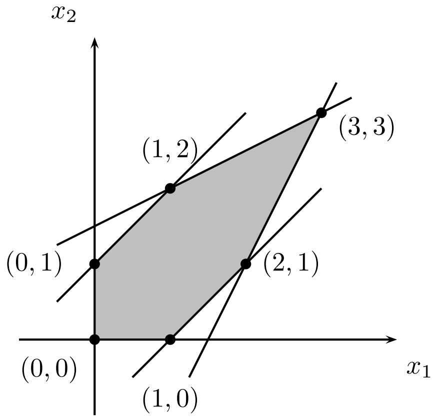
Ejemplo.

Consideremos el siguiente problema de minimización cóncava.

$$\text{mín} \quad -2(x_1 - 1.2)^2 \quad -2(x_2 - .2)^2$$

$$\begin{array}{llll}
\text{s.a} & & & \\
-x_1 & + & x_2 & \leq 1 \\
x_1 & - & x_2 & \leq 1 \\
-x_1 & + & 2x_2 & \leq 3 \\
2x_1 & - & x_2 & \leq 3 \\
x_1 & & & \geq 0 \\
 & & x_2 & \geq 0
\end{array}$$

La siguiente gráfica representa la región de soluciones factibles asociada al problema anterior.



Sea $\epsilon = 0.2$

Inicialización.

Sea $z = (0,0) \in D$. Sea $M = D$

Fase I.

Sea $x^0 = (0,0)$ tal que $f(x^0) \leq f(z) = -2.96$ y sea $\bar{x} = (0,1)$ el mínimo local, tal que $f(\bar{x}) = -4.16$ entonces sea $\gamma = -4.16$

Fase II.

Sea $\Delta = \{2, 4, 6, ...\}$

0) Sea $\alpha = \gamma - \epsilon = -4.16 - 0.2 = -4.36$. Ahora hay que trasladar el origen al punto x^0, como $x^0 = (0,0)$, el problema permanece igual.

   Construir un cono $K_0 = x^0 + \lambda_i e^i$ con $i = 1,2$ y $\lambda_i \geq 0$.

$$z^{01} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \theta_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad z^{02} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \theta_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\theta_1 = 2.663 \quad \theta_2 = 1.0603$$

   los vectores z^{0i} $i = 1,2$. serán las columnas de la matriz Q_0.

$$z^{01} = \begin{pmatrix} 2.663 \\ 0 \end{pmatrix} \qquad z^{02} = \begin{pmatrix} 0 \\ 1.0603 \end{pmatrix}$$

   Sea $Q_0 = (z^{01}, z^{02})$, $\mathcal{M} = \{Q_0\}$, $\mathcal{P} = \mathcal{M}$, $\tau(Q_0) = 0$

1) Para cada $Q \in \mathcal{P}$ resolver

$$\text{máx} \quad eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando eQ_1^{-1}

$$Q_0^{-1} = \begin{pmatrix} 0.3755 & 0 \\ 0 & 0.9431 \end{pmatrix}$$

entonces el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.3755x_1 & + & 0.9431x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.3755x_1 & & & \geq 0 \\ & & & 0.9431x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q) = (3,3)$ y su valor óptimo $\mu(Q) = 3.9558$ donde $f(\omega(Q)) = -22.16 < -4.16 = \gamma$. Así, sea $z \leftarrow \omega(Q) = (3,3)$

$$M \leftarrow M \cap \{x : eQ_0^{-1}x \geq 1\}$$
$$M \leftarrow M \cap \{x : 0.3755x_1 + 0.9431x_2 \geq 1\}$$

ir a la Fase I.

Fase I.

Empezando en $z \leftarrow (3,3)$, sea $x^0 = (3,3)$ tal que $f(x^0) \leq f(z)$ y sea $\bar{x} = (3,3)$ con $\gamma = -22.16$

Fase II.

Sea $\Delta = \{2, 4, 6, ...\}$

0) Sea $\alpha = \gamma - \epsilon = -22.16 - 0.2 = -22.36 = \alpha$
Trasladar el origen a x^0, como $x^0 = (3,3)$ el problema sobre el cual trabajaremos ahora es el siguiente:

$$\text{mín} \quad -2(x_1 - 4.2)^2 \quad -2(x_2 - 3.2)^2$$

$$\begin{array}{lcll} \text{s.a} & & & \\ -x_1 & + & x_2 & \leq 1 \\ x_1 & - & x_2 & \leq 1 \\ -x_1 & + & 2x_2 & \leq 6 \\ 2x_1 & - & x_2 & \leq 6 \\ 0.3755x_1 & + & 0.9431x_2 & \geq 4.9558 \\ x_1 & & & \geq 3 \\ & & x_2 & \geq 3 \end{array}$$

Construir el cono $K_0 = x^0 + \lambda_i e^i$ con $\lambda_i \geq 0$ para $i = 1, 2$ donde $x^0 = (3, 3)$.

$$z^{01} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \theta_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad z^{02} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \theta_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\theta_1 = 4.5377 \quad \theta_2 = 3.3209$$

los vectores z^{0i} $i = 1, 2$. serán las columnas de la matriz Q_0.

$$z^{01} = \begin{pmatrix} 7.5377 \\ 3 \end{pmatrix} \qquad z^{02} = \begin{pmatrix} 3 \\ 6.3209 \end{pmatrix}$$

Sea $Q_0 = (z^{01}, z^{02})$, $\mathcal{M} = \{Q_0\}$, $\mathcal{P} = \mathcal{M}$, $\tau(Q_0) = 0$

1) Para cada $Q \in \mathcal{P}$ resolver

$$\text{máx} \quad eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando eQ_1^{-1}

$$Q_1^{-1} = \begin{pmatrix} 0.1663 & -0.0789 \\ -0.0789 & 0.1957 \end{pmatrix}$$

entonces el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0874x_1 & + & 0.1167x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.1663x_1 & - & 0.0789x_2 & \geq 0 \\ & -0.0789x_1 & + & 0.1957x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q) = (6, 6)$ y su valor óptimo $\mu(Q) = 1.2246$ donde $f(\omega(Q)) = -22.16 = \gamma$. ir a 2.

2) Eliminar en $\mathcal{M}$ todo $Q \in \mathcal{P}$ tal que $\mu(Q) \leq 1$. Para este problema no satisface dicha condición Q_0 por lo que $\mathcal{R} = \{Q_0\}$ e ir a 3

3) Escoger $Q_* \in argmax\{\mu(Q) : Q \in \mathcal{R}\}$ así, $Q_* = Q_0$

   a) $\tau Q_* = \tau(Q^0) = 0$ y $0 \notin \Delta$ entonces hay que subdividir Q_* con respecto a $\omega(Q_*) = (6, 6)$. Para realizar dicho proceso necesitamos a $x^0 = (3, 3)$ a $\omega(Q_*) = (6, 6)$ y su $\alpha-$extensión.

$$\omega * (Q_*) = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \quad + \theta \begin{pmatrix} 6-3 \\ 6-3 \end{pmatrix} \quad = \begin{pmatrix} 3+3\theta \\ 3+3\theta \end{pmatrix}$$

   donde $\theta = 1.00362$

$$\begin{pmatrix} 6.01086 \\ 6.01086 \end{pmatrix}$$

de donde

$$Q_1 = \begin{pmatrix} 6.01086 & 3 \\ 6.01086 & 6.3209 \end{pmatrix} \qquad Q_2 = \begin{pmatrix} 7.5377 & 6.01086 \\ 3 & 6.01086 \end{pmatrix}$$

donde: $\tau(Q_1) = 1$ y $\tau(Q_2) = 1$, $\mathcal{P}_* = \{Q_1, Q_2\}$, $\mathcal{M}_* = \{Q_1, Q_2\}$, $\mathcal{P} = \mathcal{P}_*$, $\mathcal{M} = \mathcal{M}_*$ e ir a 1.

1) Resolver para cada $Q \in \mathcal{P}$

$$\text{máx} \quad eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando Q_1^{-1}

$$Q_1^{-1} = \begin{pmatrix} 0.3167 & -0.1503 \\ -0.3011 & 0.3011 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{lllll} \text{máx} & 0.0155x_1 & + & 0.1508x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.3167x_1 & - & 0.1503x_2 & \geq 0 \\ & -0.3011x_1 & + & 0.3011x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_1) = (6, 6)$ su valor óptimo $\mu(Q_1) = 0.9978$ y $f(\omega(Q_1)) = -22.16$.

Ahora veamos que sucede con la otra partición: Calculando Q_2^{-1}

$$Q_2^{-1} = \begin{pmatrix} 0.2204 & -0.2204 \\ -0.1100 & 0.2764 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{lllll} \text{máx} & 0.1164x_1 & + & 0.0560x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.2204x_1 & - & 0.2204x_2 & \geq 0 \\ & -0.1100x_1 & + & 0.2764x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_2) = (6, 6)$ su valor óptimo $\mu(Q_1) = 0.9984$ y $f(\omega(Q_1)) = -22.16$ e ir a 2.

2) Eliminar en $\mathcal{M}$ todo $Q \in \mathcal{P}$ que satisfaga que $\mu(Q) \leq 1$. Así, $\mathcal{R} = \varnothing$ por lo tanto $\bar{x} = (3, 3)$ es un óptimo $\epsilon-$global.

Supongamos ahora que con el mismo ejemplo ahora elegimos en la Fase II a $\Delta = \{0, 2, 4, ...\}$, que a diferencia del conjunto Δ que acabamos de usar, éste contiene al elemento 0. ¿Cuáles son los cambios con el nuevo Δ? Pues bien, al inicio todo es igual, hasta que llegamos a la iteración donde trasladamos el problema al punto $x^0 = (3, 3)$ y nos encontramos en el inciso 3). A continuación mostramos como termina el problema.

3) Escoger $Q_* \in argmax\{\mu(Q) : Q \in \mathcal{R}\}$, entonces $Q_* = Q_0$.

   b) Como $\tau(Q_*) = 0$ y $0 \in \Delta$ entonces hay que particionar a Q_* con respecto al punto medio de la arista más grande de $Z(Q_*)$, donde $Z(Q_*)$ es el $n-1$ simplex generado con las columnas de Q_*.

$$Q_* = \begin{pmatrix} 7.5377 & 3 \\ 3 & 6.3209 \end{pmatrix}$$

de aquí, el 1−simplex $Z = [z^{01} \;\; z^{02}]$ es el segmento de recta cuyos puntos extremos son z^{01} y z^{02}, así que necesitamos obtener a su punto medio

$$\begin{pmatrix} \dfrac{7.5377+3}{2} & , & \dfrac{3+6.3209}{2} \end{pmatrix}$$

$$\omega(Q_*) = \begin{pmatrix} 5.26885 & , & 4.66045 \end{pmatrix}$$

Ahora generamos su $\alpha-$extensión

$$\omega * (Q_*) = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \theta \begin{pmatrix} 5.26885 - 3 \\ 4.66045 - 3 \end{pmatrix}$$

como $\theta = 1.5324$ entonces

$$\omega * (Q_*) = \begin{pmatrix} 6.4768 \\ 5.5445 \end{pmatrix}$$

entonces

$$Q_1 = \begin{pmatrix} 6.4768 & 3 \\ 5.5445 & 6.3209 \end{pmatrix} \qquad Q_2 = \begin{pmatrix} 7.5377 & 6.4768 \\ 3 & 5.5445 \end{pmatrix}$$

$\tau(Q_1) = 1$ y $\tau(Q_2) = 1$

4) $\mathcal{P}_* = \{Q_1, Q_2\}$, $\mathcal{M}_* = \{Q_1, Q_2\}$ e ir a 1.

1) Para cada $Q \in \mathcal{P}$ resolver

$$\text{máx} \;\; eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando Q_1^{-1}

$$Q_1^{-1} = \begin{pmatrix} 0.2601 & -0.1234 \\ -0.2281 & 0.2665 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0319x_1 & + & 0.143x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.2601x_1 & - & 0.1234x_2 & \geq 0 \\ & -0.2281x_1 & + & 0.2665x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_1) = (6, 6)$ su valor óptimo $\mu(Q_1) = 1.0518$ y $f(\omega(Q_1)) = -22.16$.

Ahora veamos que sucede con la otra partición:

Calculando Q_2^{-1}

$$Q_2^{-1} = \begin{pmatrix} 0.2479 & -0.2896 \\ -0.1342 & 0.3371 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.1138x_1 & + & 0.0474x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.2479x_1 & - & 0.2896x_2 & \geq 0 \\ & -0.1342x_1 & + & 0.3371x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_2) = (5.244793, 4.489586)$ su valor óptimo $\mu(Q_1) = 0.8096638$ y $f(\omega(Q_1)) = -5.4630$.

Como ningún $\omega(Q)$ satisface $f(\omega(Q)) < \gamma$ ir a 2.

2) En $\mathcal{M}$ eliminar toda Q tal que $\mu(Q) \leq 1$; eliminamos a Q_2. Sea $\mathcal{R} = \{Q_1\}$ e ir a 3

3) Escoger $Q_* \in argmax\{\mu(Q) : Q \in \mathcal{R}\}$, entonces $Q_* = Q_1$.

   a) Como $\tau(Q_*) = 1$ y $1 \notin \Delta$ entonces hay que particionar a Q_* con respecto a $\omega(Q_*)$; tenemos $x^0 = (3, 3)$, $\omega(Q) = (6, 6)$ y construiremos su $\alpha-$extensión.

$$\omega(Q_*) = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \theta \begin{pmatrix} 6-3 \\ 6-3 \end{pmatrix}$$

   como $\theta = 1.00362$

$$\omega(Q_*) = \begin{pmatrix} 6.01086 \\ 6.01086 \end{pmatrix}$$

   donde

$$Q_1 = \begin{pmatrix} 6.01086 & 3 \\ 6.01086 & 6.3209 \end{pmatrix} \qquad Q_2 = \begin{pmatrix} 6.4768 & 6.01086 \\ 5.5445 & 6.01086 \end{pmatrix}$$

   y $\tau(Q_1) = 2$, $\tau(Q_2) = 2$

4) $\mathcal{P}_* = \{Q_1, Q_2\}$, $\mathcal{M}_* = \{Q_1, Q_2\}$ e ir a 1.

1) Resolver para cada $Q \in \mathcal{P}$

$$\text{máx} \quad eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando Q_1^{-1}

$$Q_1^{-1} = \begin{pmatrix} 0.3167 & -0.1503 \\ -0.3011 & 0.3011 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0155x_1 & + & 0.1508x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.3167x_1 & - & 0.1503x_2 & \geq 0 \\ & -0.3011x_1 & + & 0.3011x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_1) = (6,6)$ su valor óptimo $\mu(Q_1) = 0.9978$ y $f(\omega(Q_1)) = -22.16$.

Ahora veamos que sucede con la otra partición: Calculando Q_2^{-1}

$$Q_2^{-1} = \begin{pmatrix} 1.0726 & -1.0726 \\ -0.9894 & 1.1558 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0832x_1 & + & 0.0831x_2 & \\ \text{s.a} & x \in M & & & \\ & 1.0726x_1 & - & 1.0726x_2 & \geq 0 \\ & -0.9894x_1 & + & 1.1558x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_2) = (6,6)$ su valor óptimo $\mu(Q_1) = 0.9978$ y $f(\omega(Q_1)) = -22.16$ e ir a 2.

2) Eliminamos en $\mathcal{M}$ todo Q tal que $\mu(Q) \leq 1$, por lo tanto $\mathcal{R} = \varnothing$. Por lo tanto $\bar{x} = (3,3)$ es una solución ϵ global.

Ahora bien, si para este ejemplo usamos $\Delta = \{0, 1, 2, ...\}$ el proceso de partición se hace a través de bisecciones, los cambios comienzan en la aplicación de la modificación anterior en el inciso 3) cuando en este caso tendremos que $\tau(Q_1) = 1$ donde $1 \in \Delta$ como enseguida mostraremos:

3) Escoge $Q_* \in argmax\{\mu(Q) : Q \in \nabla\}$, entonces $Q_* = Q_1$

b) Como $\tau(Q_1) = 1$, $1 \in Delta$ por lo tanto dividiremos con respecto al punto medio de la arista más grande.

$$Q_* = \begin{pmatrix} 6.4768 & 3 \\ 5.5445 & 6.3209 \end{pmatrix}$$

su punto medio es

$$\omega* = \left(\frac{6.4768+3}{2}, \frac{5.5445+6.3209}{2}\right) \qquad = \left(4.7384, 5.9327\right)$$

Su $\alpha-$extensión es

$$\omega* = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \theta \begin{pmatrix} 4.7384 - 3 \\ 5.9327 - 3 \end{pmatrix}$$

como $\theta = 1{,}172$

$$\omega* = \begin{pmatrix} 5.0374 \\ 6.4371 \end{pmatrix}$$

Así, Q_* se particiona del siguiente modo:

$$Q_1 = \begin{pmatrix} 5.0374 & 3 \\ 6.4371 & 6.3209 \end{pmatrix} \qquad Q_2 = \begin{pmatrix} 6.4768 & 5.0374 \\ 5.5445 & 6.4371 \end{pmatrix}$$

donde $\tau(Q_1) = 2$ y $\tau(Q_2) = 2$

4) $\mathcal{P} = \{Q_1, Q_2\}$, $\mathcal{M} = \{Q_1, Q_2\}$, $\mathcal{P} \leftarrow \mathcal{P}_*$, $\mathcal{M} \leftarrow \mathcal{M}$ e ir a 1

1) Resolver para cada $Q \in \mathcal{P}$

$$\text{máx} \;\; eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando Q_1^{-1}

$$Q_1^{-1} = \begin{pmatrix} 0.5045 & -0.2394 \\ -0.5138 & 0.4020 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{lllll} \text{máx} & -0.0093x_1 & + & 0.1626x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.5045x_1 & - & 0.2394x_2 & \geq 0 \\ & -0.5138x_1 & + & 0.4020x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_1) = (3.645569, 4.645569)$ su valor óptimo $\mu(Q_1) = 0.7214657$ y $f(\omega(Q_1)) = -4.7941$.

Ahora veamos que sucede con la otra partición: Calculando Q_2^{-1}

$$Q_2^{-1} = \begin{pmatrix} 0.4677 & -0.3660 \\ -0.4029 & .4706 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{lllll} \text{máx} & 0.0649x_1 & + & 0.1046x_2 & \\ \text{s.a} & x \in M & & & \\ & .4677x_1 & - & 0.3660x_2 & \geq 0 \\ & -0.4029x_1 & + & 0.4706x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_2) = (6, 6)$ su valor óptimo $\mu(Q_1) = 1.017$ y $f(\omega(Q_2)) = -22.16$ e ir a 2.

2) Eliminamos en $\mathcal{M}$ todo Q tal que $\mu(Q) \leq 1$; borramos a Q_1. Sea $\mathcal{R} = \{Q_2\}$ e ir a 3.

3) Escoger $Q_* \in argmax\{\mu(Q) : Q \in \mathcal{R}\}$, $Q_* = Q_2$.

   b) Como $\tau(Q_2) = 2$, $2 \in \Delta$ entonces haremos la partición con respecto al punto medio de la arista más grande.

$$Q_* = \begin{pmatrix} 6.4768 & 5.0374 \\ 5.5445 & 6.4371 \end{pmatrix}$$

   su punto medio es

$$\omega * = \left(\frac{6.4768 + 5.0374}{2}, \frac{5.5445 + 6.4371}{2} \right) \qquad = \big(5.7571, 5.9908\big)$$

   Su $\alpha-$extensión es

$$\omega * = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \theta \begin{pmatrix} 5.7571 - 3 \\ 5.9908 - 3 \end{pmatrix}$$

   como $\theta = 1.03734$

$$\omega * = \begin{pmatrix} 5.8601 \\ 6.1025 \end{pmatrix}$$

   Así, Q_* se particiona del siguiente modo:

$$Q_1 = \begin{pmatrix} 5.8601 & 5.0374 \\ 6.1025 & 6.4371 \end{pmatrix} \qquad Q_2 = \begin{pmatrix} 6.4768 & 5.8601 \\ 5.5445 & 6.1025 \end{pmatrix}$$

   donde $\tau(Q_1) = 3$ y $\tau(Q_2) = 3$

4) $\mathcal{P} = \{Q_1, Q_2\}$, $\mathcal{M} = \{Q_1, Q_2\}$, $\mathcal{P} \leftarrow \mathcal{P}_*$, $\mathcal{M} \leftarrow \mathcal{M}$ e ir a 1

1) Resolver para cada $Q \in \mathcal{P}$

$$\text{máx} \;\; eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando Q_1^{-1}

$$Q_1^{-1} = \begin{pmatrix} 0.9220 & -0.7216 \\ -0.8741 & 0.8394 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0479x_1 & + & 0.1178x_2 & \\ \text{s.a} & x \in M & & & \\ & 0.9220x_1 & - & 0.7216x_2 & \geq 0 \\ & -0.8741x_1 & + & 0.8394x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_1) = (5.541813, 5.770906)$ su valor óptimo $\mu(Q_1) = 0.9452656$ y $f(\omega(Q_1)) = -16.82$.

Ahora veamos que sucede con la otra partición: Calculando Q_2^{-1}

$$Q_2^{-1} = \begin{pmatrix} 0.8677 & -0.8332 \\ -0.7883 & .9209 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0793x_1 & + & 0.0877x_2 & \\ \text{s.a} & x \in M & & & \\ & .8677x_1 & - & 0.8332x_2 & \geq 0 \\ & -0.7883x_1 & + & 0.9209x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_2) = (6, 6)$ su valor óptimo $\mu(Q_1) = 1.002$ y $f(\omega(Q_2)) = -22.16$ e ir a 2.

2) Eliminamos en $\mathcal{M}$ todo Q tal que $\mu(Q) \leq 1$; borramos a Q_1. Sea $\mathcal{R} = \{Q_2\}$ e ir a 3.

3) Escoger $Q_* \in argmax\{\mu(Q) : Q \in \mathcal{R}\}$, $Q_* = Q_2$.

   b) Como $\tau(Q_*) = 3$, $3 \in \Delta$ entonces subdividimos con respecto al punto medio de la arista más grande.

$$Q_* = \begin{pmatrix} 6.4768 & 5.8601 \\ 5.5445 & 6.1025 \end{pmatrix}$$

   su punto medio es

$$\omega* = \left(\frac{6.4768 + 5.8601}{2}, \frac{5.5445 + 6.1025}{2} \right)$$

$$\omega* = \big(6.1685, 5.8235\big)$$

Su $\alpha-$extensión es

$$\omega* = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \theta \begin{pmatrix} 6.1685 - 3 \\ 5.8235 - 3 \end{pmatrix}$$

como $\theta = 1.01532$

$$\omega* = \begin{pmatrix} 6.2170 \\ 5.8668 \end{pmatrix}$$

Así, Q_* se particiona del siguiente modo:

$$Q_1 = \begin{pmatrix} 6.2170 & 5.8601 \\ 5.8668 & 6.1025 \end{pmatrix} \qquad Q_2 = \begin{pmatrix} 6.4768 & 6.2170 \\ 5.5445 & 5.8668 \end{pmatrix}$$

donde $\tau(Q_1) = 4$ y $\tau(Q_2) = 4$

4) $\mathcal{P} = \{Q_1, Q_2\}$, $\mathcal{M} = \{Q_1, Q_2\}$, $\mathcal{P} \leftarrow \mathcal{P}_*$, $\mathcal{M} \leftarrow \mathcal{M}$ e ir a 1

1) Resolver para cada $Q \in \mathcal{P}$

$$\text{máx} \;\; eQ^{-1}x \quad \text{s.a} \quad x \in M \quad Q^{-1}x \geq 0$$

Calculando Q_1^{-1}

$$Q_1^{-1} = \begin{pmatrix} 1.7146 & -1.6465 \\ -1.6483 & 1.7467 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0662x_1 & + & 0.1003x_2 & \\ \text{s.a} & x \in M & & & \\ & 1.7146x_1 & - & 1.6465x_2 & \geq 0 \\ & -1.6483x_1 & + & 1.7467x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_1) = (6, 6)$ su valor óptimo $\mu(Q_1) = 0.999$ y $f(\omega(Q_1)) = -22.16$.

Ahora veamos que sucede con la otra partición: Calculando Q_2^{-1}

$$Q_2^{-1} = \begin{pmatrix} 1.6630 & -1.7622 \\ -1.5716 & 1.8359 \end{pmatrix}$$

el problema de programación lineal a resolver es:

$$\begin{array}{llcll} \text{máx} & 0.0914x_1 & + & 0.0736x_2 & \\ \text{s.a} & x \in M & & & \\ & 1.6630x_1 & - & 1.7622x_2 & \geq 0 \\ & -1.5716x_1 & + & 1.8359x_2 & \geq 0 \end{array}$$

la solución es: $\omega(Q_2) = (5.680241, 5.360481)$ su valor óptimo $\mu(Q_1) = 0.9137054$ y $f(\omega(Q_2)) = -13.7176$ e ir a 2.

2) Eliminamos en $\mathcal{M}$ todo Q tal que $\mu(Q) \leq 1$; borramos a Q_1. Sea $\mathcal{R} = \varnothing$. Por lo tanto $\bar{x} = (3, 3)$ es una solución ϵ óptima global.

4.3. Algoritmos simpliciales

En esta sección se estudiaran los algoritmos simpliciales, dichos procesos reciben su nombre debido a que la región factible sobre la cual trabajan esta contenida en un inicio en un conjunto llamado simplex, el cuál a través de las iteraciones será dividido en subconjuntos que también son simplex. La forma en la cual son divididos puede ser de dos maneras; por medio del método llamado subdivisión radial o bien a través de bisección. Al proceso de dividir es lo que conocemos por ramificación. Recordemos de la primera sección que es necesaria la construcción de cotas. En este caso obtendremos a la cota inferior $\beta(M)$ a través de la envolvente convexa φ_M de f sobre M. Como f es una función cóncava, φ será una función afín, una vez obtenida ésta última será usada en el programa lineal mín φ_M sobre el conjunto $D \cap M$. Para obtener a φ será necesario resolver un sistema de ecuaciones lineales, el cual satisface que φ_M coincide con f en los vértices de M.

4.3.1. Proceso Normal Simplicial de Subdivisión

Para poder desarrollar adecuadamente el algoritmo es necesario construir un proceso de subdivisión, para ello consideremos cualquier $n-$simplex $M = [v^1, v^2, ..., v^n] \in \mathbb{R}^n$ y tomemos a φ como: $\varphi_M(x) = \pi(M)x + r(M)$ donde $\pi(M) \in \mathbb{R}^n$ y $r(M) \in \mathbb{R}$. Recordemos que $\varphi_M(x)$ es una función afín que coincide con $f(x)$ en los vértices de M.

Definición 16. *Una sucesión infinita anidada de simplex M se dice que es no degenerada si* $\lim_{s\to\infty} ||\pi(M_s)|| < \infty$, *es decir, si existe una subsucesión $\Delta \subset \{1, 2, ...\}$ y una constante η tal que $||\pi(M_s)|| \leq \eta \ \forall s \in \Delta$. Un proceso de subdivisión simplicial es no degenerado si cualquier sucesión anidada infinita de simplex que produce es no degenerada.*

Proposición 9. *Sea $M_s = [v^{s1}, v^{s2}, ..., v^{sn}]$ con $s = 1, 2, ...$, una sucesión infinita anidada de simplises tales que M_{s+1} es obtenido a partir de M_s a través de una subdivisión radial con respecto al punto ω^s. Si la sucesión es no degenerada entonces*

$$\liminf_{s\to\infty} |f(\omega^s) - \varphi_{M_s}(\omega^s)| = 0$$

Corolario 6. *Un proceso de subdivisión simplicial es normal si para cualquier sucesión infinita anidada de simplex M_s que genera satisface alguna de las siguientes condiciones:*

- *es exhaustivo*
- *es no degenerado y satisface $w^s = \omega^s$ para todo s finito.*

El proceso básico NSS

Con el fin de hacer un algoritmo convergente y eficiente es que se combinarán dos métodos de partición de un poliedro; la subdivisión radial y la bisección, usando al conjunto Δ, donde Δ es propuesto de tal manera que satisfaga con ser una sucesión infinita creciente de números naturales. Y la manera en la que este proceso de elección de partición funciona es el siguiente:

Sea $\tau(M_0) = 0$ para el simplex inicial M_0 en la iteración $k = 0, 1, 2, ...$, si M_k es el simplex a subdividir, entonces

a) Si $\tau(M_k) \notin \Delta$ entonces subdividir a M_k con respecto a $w^k = \omega^k$, es decir, estamos haciendo una $\omega-$subdivisión de M_k y hacer $\tau(M) = \tau(M_k) + 1$ para cada subsimplex M de la subdivisión.

b) Sino, si $\tau(M_k) \in \Delta$ entonces bisectar M_k, es decir, subdividir con respecto a w^k que es el punto medio de la arista más grande de M_k. Sea $\tau(M) = \tau(M_k) + 1$ para cada subsimplex M de la subdivisión.

4.3.2. Algoritmo Normal Simplicial

Elegir $\epsilon \geq 0$

Inicialización.

Elegir un $n-$simplex $M_0 \supset D$ tal que un vértice de M_0 también lo sea de D.

Sea x^0 la mejor solución hasta ahora.

Sea $\mathcal{M}_0 = \mathcal{N}_0 = \{M_0\}$

Iteración k=0,1,...

1) Para cada $M \in \mathcal{N}_k$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M y resolver el programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

obtener la solución óptima básica $\omega(M)$ y su valor óptimo $\beta(M)$.

2) Eliminar de $M \in \mathcal{M}_k$ tal que $\beta(M) \geq f(x^k) - \epsilon$.

Sea $\mathcal{R}_k$ la colección restante de simplices.

3) Si $\mathcal{R} = \varnothing$, Fin: x^k es una solución ϵ óptima global, sino ir a 4

4) Elegir $M_k \in argmin\{\beta(M) : M \in \mathcal{R}_k\}$ y subdivide de acuerdo al proceso elegido.

5) Actualiza los datos, x^{k+1} será la mejor solución conocida hasta este punto.

   Sea $\mathcal{N}_{k+1}$ el conjunto de subsimplices de M_k que se generaron en la subdivisión en el paso 4.

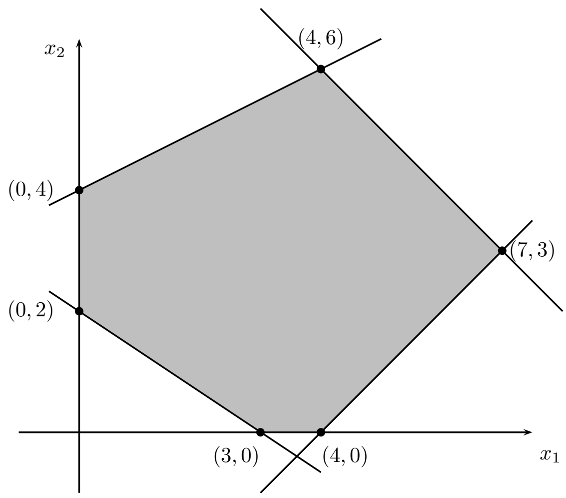
   Sea $\mathcal{M}_{k+1} = (\mathcal{R}_k \diagdown \{M_k\}) \cup \mathcal{N}_{k+1}$.

   Sea $k \leftarrow k+1$ y regresar al paso 1.

Ejemplo.

$$\begin{array}{lclc} & \text{mín} & -(x_1 - x_2 - 4)^2 & \\ \\ x_1 & + & x_2 & \leq 10 \\ -x_1 & + & 2x_2 & \leq 8 \\ -2x_1 & - & 3x_2 & \leq -6 \\ x_1 & - & x_2 & \leq 4 \\ x_1 & & & \geq 0 \\ & & x_2 & \geq 0 \end{array}$$

La siguiente gráfica representa la región de soluciones factibles del planteamiento anterior.



Sea $\epsilon = 0.2$.

Sea $M_0 = conv\{(0,0),(0,10),(10,0)\} \supset D$ el $n-$simplex.

Sea $x^0 = (0,2)$ la mejor solución disponible. Sea $\mathcal{M}_0 = \mathcal{N}_0 = \{M_0\}$

Iteración 0.

1) Para cada $M \in \mathcal{N}_0$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

   $\varphi(x_1, x_2) = a_1 x_1 + a_2 x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{rcrcl} & & & a_3 & = -16 \\ & & 10a_2 & + \; a_3 & = -196 \\ 10a_1 & & & + \; a_3 & = -36 \end{array}$$

   de aquí tenemos que $a_1 = -2$, $a_2 = -18$, $a_3 = -16$.

   Por lo tanto la función afín es $\varphi_M(x) = -2x_1 - 18x_2 - 16$

   Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

   es decir

$$\text{mín} \quad -2x_1 - 18x_2 - 16 \quad \text{s.a} \quad x \in D \cap M$$

   La solución al programa lineal es: $\omega(M) = (4, 6)$ y su valor óptimo es $\beta(M) = -132$

2) Eliminar toda $M \in \mathcal{M}_0$ tal que $\beta(M) \geq f(x^0) - \epsilon$, donde $f(x^0) - \epsilon = -36 - 0.2 = -36.2$, como $-132 < -36.2$ no se elimina nada.

   Sea $\mathcal{R}_0 = \{M_0\}$

3) Como $\mathcal{R}_0 \neq \varnothing$ ir a 4

4) Elegir $M_0 \in argmin\{\beta(M) : M \in \mathcal{R}_0\}$.

   $M_0 = conv\{(0,0), (0,10), (10,0)\}$ subdividir de acuerdo al proceso elegido.

   Sea $\Delta = \{2, 4, 6...\}$ Sea $\tau(M_0) = 0$

   a) Como $0 \notin \Delta$ entonces subdividimos con respecto a $w^0 = \omega^0$
      Ahora tenemos dos subsimplices cuyas envolventes convexas son:
      $SM_{01} = \{(0,0), (10,0), (4,6)\}$ $\tau(M_{01}) = 1$
      $SM_{02} = \{(0,0), (0,10), (4,6)\}$ $\tau(M_{02}) = 1$

5) Sea $x^1 = (4,6)$ $f(x^1) = -36$. Sea $\mathcal{N}_1 = \{SM_{01}, SM_{02}\}$. Sea $\mathcal{M}_1 = (\mathcal{R}_0 \diagdown \{M_0\}) \cup \mathcal{N}_\infty$, así, $\mathcal{M}_1 = \mathcal{N}_1$ pues $\mathcal{R}_0 = \{M_0\}$

   Sea $k \leftarrow 1$ ir a 1

Iteración 1.

1) Para cada $M \in \mathcal{N}_1$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

$\varphi_{M_{01}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3.$ resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & & & a_3 & = -16 \\ 10a_1 & & & + & a_3 & = -36 \\ 4a_1 & + & 6a_2 & + & a_3 & = -36 \end{array}$$

de aquí tenemos que $a_1 = -2$, $a_2 = -2$, $a_3 = -16$.

Por lo tanto la función afín es $\varphi_M(x) = -2x_1 - 2x_2 - 16$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad -2x_1 - 2x_2 - 16 \quad \text{s.a} \quad x \in D \cap M \cap 6x_1 - 4x_2 \geq 0$$

La solución al programa lineal es: $\omega(M_{01}) = (4, 6)$ y su valor óptimo es $\beta(M_{01}) = -36$

Veamos que sucede con el otro subsimplex: $\varphi_{M_{02}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3.$ resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & & & a_3 & = -16 \\ & & 10a_2 & + & a_3 & = -196 \\ 4a_1 & + & 6a_2 & + & a_3 & = -36 \end{array}$$

de aquí tenemos que $a_1 = 22$, $a_2 = -18$, $a_3 = -16$.

Por lo tanto la función afín es $\varphi_M(x) = 22x_1 - 18x_2 - 16$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad 22x_1 - 18x_2 - 16 \quad \text{s.a} \quad x \in D \cap M \cap 6x_1 - 4x_2 \leq 0$$

La solución al programa lineal es: $\omega(M_{02}) = (0, 4)$ y su valor óptimo es $\beta(M_{02}) = -88$

2) Eliminar toda $M \in \mathcal{M}_1$ tal que $\beta(M) \geq f(x^1) - \epsilon$, donde $f(x^1) - \epsilon = -36 - 0.2 = -36.2$, como $-36 \leq -36.2$ eliminamos SM_{01}.

   Sea $\mathcal{R}_1 = \{M_{02}\}$

3) Como $\mathcal{R}_1 \neq \varnothing$ ir a 4

4) Elegir $M_1 \in argmin\{\beta(M) : M \in \mathcal{R}_1\}$.

   $M_1 = conv\{(0,0),(0,10),(4,6)\}$ subdividir de acuerdo al proceso elegido. Sabemos que $\tau(M_1) = 1$

   a) Como $1 \notin \Delta$ entonces subdividimos con respecto a $w^1 = \omega^1 = (0,4)$
      Ahora tenemos dos subsimplices cuyas envolventes convexas son:
      $SM_{11} = \{(0,0),(0,4),(4,6)\}$ $\tau(M_{11}) = 2$
      $SM_{12} = \{(0,4),(0,10),(4,6)\}$ $\tau(M_{12}) = 2$

5) Sea $x^2 = (0,4)$ $f(x^2) = -64$. Sea $\mathcal{N}_2 = \{SM_{11}, SM_{12}\}$. Sea $\mathcal{M}_2 = (\mathcal{R}_1 \diagdown \{M_1\}) \cup \mathcal{N}_{\in}$, así, $\mathcal{M}_2 = \mathcal{N}_2$ pues $\mathcal{R}_1 = \{M_1\}$

   Sea $k \leftarrow 2$ ir a 1

Iteración 2.

1) Para cada $M \in \mathcal{N}_2$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

   $\varphi_{M_{11}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & & & a_3 & = & -16 \\ & & 4a_2 & + & a_3 & = & -64 \\ 4a_1 & + & 6a_2 & + & a_3 & = & -36 \end{array}$$

   de aquí tenemos que $a_1 = -13$, $a_2 = -12$, $a_3 = -16$.

   Por lo tanto la función afín es $\varphi_M(x) = -13x_1 - 12x_2 - 16$

   Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

   es decir

$$\text{mín} \quad -13x_1 - 12x_2 - 16 \quad \text{s.a} \quad x \in D \cap M \cap 6x_1 - 4x_2 \leq 0$$

   La solución al programa lineal es: $\omega(M_{11}) = (4,6)$ y su valor óptimo es $\beta(M_{11}) = -140$

   Veamos que sucede con el otro subsimplex: $\varphi_{M_{12}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & 4a_2 & + & a_3 & = -64 \\ & & 10a_2 & + & a_3 & = -196 \\ 4a_1 & + & 6a_2 & + & a_3 & = -36 \end{array}$$

de aquí tenemos que $a_1 = 18$, $a_2 = -22$, $a_3 = 24$.

Por lo tanto la función afín es $\varphi_M(x) = 18x_1 - 22x_2 + 24$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad 18x_1 - 22x_2 + 24 \quad \text{s.a} \quad x \in D \cap M \cap x_1 + 2x_2 \geq 8$$

La solución al programa lineal es: $\omega(M_{12}) = (0, 4)$ y su valor óptimo es $\beta(M_{12}) = -64$

2) Eliminar toda $M \in \mathcal{M}_2$ tal que $\beta(M) \geq f(x^2) - \epsilon$, donde $f(x^2) - \epsilon = -64 - 0.2 = -64.2$, como $-64 \leq -36.2$ eliminamos SM_{02}.

   Sea $\mathcal{R}_2 = \{M_{11}\}$

3) Como $\mathcal{R}_2 \neq \varnothing$ ir a 4

4) Elegir $M_2 \in argmin\{\beta(M) : M \in \mathcal{R}_2\}$.

   $M_2 = conv\{(0,0), (0,4), (4,6)\}$ subdividir de acuerdo al proceso elegido. Sabemos que $\tau(M_2) = 2$

   a) Como $2 \in \Delta$ entonces subdividimos con respecto a la arista más grande, la cual va desde el punto $(0, 0)$ hasta el punto $(4, 6)$ y su punto medio es $(2, 3)$.

      Ahora tenemos dos subsimplices cuyas envolventes convexas son:

      $SM_{21} = \{(0,0), (0,4), (2,3)\}$ $\tau(M_{11}) = 3$

      $SM_{22} = \{(0,4), (2,3), (4,6)\}$ $\tau(M_{12}) = 3$

5) Sea $x^3 = (0, 4)$ $f(x^2) = -64$. Sea $\mathcal{N}_3 = \{SM_{21}, SM_{22}\}$. Sea $\mathcal{M}_3 = (\mathcal{R}_2 \diagdown \{M_2\}) \cup \mathcal{N}_\ni$, así, $\mathcal{M}_3 = \mathcal{N}_3$ pues $\mathcal{R}_2 = \{M_2\}$

   Sea $k \leftarrow 3$ ir a 1

Iteración 3.

1) Para cada $M \in \mathcal{N}_3$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

   $\varphi_{M_{21}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3.$ resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & & & a_3 & = -16 \\ & & 4a_2 & + & a_3 & = -64 \\ 2a_1 & + & 3a_2 & + & a_3 & = -25 \end{array}$$

de aquí tenemos que $a_1 = 13.5$, $a_2 = -12$, $a_3 = -16$.

Por lo tanto la función afín es $\varphi_M(x) = 13.5x_1 - 12x_2 - 16$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad 13.5x_1 - 12x_2 - 16 \quad \text{s.a} \quad x \in D \cap M \cap \tfrac{1}{2}x_1 + x_2 \leq 4$$

La solución al programa lineal es: $\omega(M_{21}) = (0, 4)$ y su valor óptimo es $\beta(M_{21}) = -64$

Veamos que sucede con el otro subsimplex: $\varphi_{M_{22}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & 4a_2 & + & a_3 & = -64 \\ 2a_1 & + & 3a_2 & + & a_3 & = -25 \\ 4a_1 & + & 6a_2 & + & a_3 & = -36 \end{array}$$

de aquí tenemos que $a_1 = -\frac{5}{2}$, $a_2 = -2$, $a_3 = -14$.

Por lo tanto la función afín es $\varphi_M(x) = -\frac{5}{2}x_1 - 2x_2 - 14$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad -\tfrac{5}{2}x_1 - 2x_2 - 14 \quad \text{s.a} \quad x \in D \cap M \cap \tfrac{1}{2}x_1 + x_2 \geq 4$$

La solución al programa lineal es: $\omega(M_{22}) = (4, 6)$ y su valor óptimo es $\beta(M_{22}) = -36$

2) Eliminar toda $M \in \mathcal{M}_3$ tal que $\beta(M) \geq f(x^3) - \epsilon$, donde $f(x^3) - \epsilon = -64 - 0.2 = -64.2$, como $-64 > -64.2$ y $-36 > -64.2$ entonces $\mathcal{R}_3 = \varnothing$.

3) Como $\mathcal{R}_3 = \varnothing$ Fin; x^3 es una solución ϵ global.

4.4. Algoritmos rectangulares

Esta sección está dedicada al estudio de algoritmos rectangulares, los cuales toman dicho nombre por la forma en la cual se lleva a cabo la partición del poliedro. Este tipo de procedimiento por lo general se lleva a cabo cuando la función f es separable o bien puede ser escrita como tal.

Un rectángulo (o hiperrectángulo) $M \subseteq \mathbb{R}^n$ es un conjunto no vacío de la forma $M = \{x \in \mathbb{R}^n | r \leq x \leq s\}$, donde $r, s \in \mathbb{R}^n$. Cuando se implementa la partición sucesiva a través de rectángulos, el conjunto inicial M_0 es escogido como un rectángulo que contiene a D y cada partición de M_0 y los subconjuntos de M_0 son elementos que a su vez también son rectángulos.

Es importante mencionar que hay varias formas de hacer la partición de rectángulos en subrectángulos. Por mencionar algunos, existen el refinamiento fuerte, el débil y la bisección.

Cotas inferiores

Con respecto a la obtención de las cotas inferiores $\beta(M)$ de M, es necesario considerar que se puede hacer de varias formas [3], en este trabajo veremos el caso cuando D es un poliedro y f es separable o bien puede llevarse a dicha forma. Así, calcular $\beta(M)$ se hace a través de un programa lineal, donde la función objetivo $\varphi_M(x)$ es la envolvente convexa de f sobre M, la cual resulta ser una función afín. De este programa lineal obtenemos a $\omega(M)$ que representa a la solución óptima, a $\beta(M)$ que es el valor óptimo del problema lineal que a su vez es la cota buscada. De este modo, la sucesión infinita anidada de rectángulos M_h con $h = 0, 1, ...$ generada en el proceso, para cada h tiene $\omega^h = \omega(M_h)$ y $\varphi_h(x) = \varphi_{M_h}(x)$. La siguiente definición nos ayuda a establecer la normalidad del proceso de partición.

Definición 17. *Una sucesión anidada M_h, $h = 0, 1, ..$ se dice que es normal si* $\liminf_{h\to\infty} |f(\omega^h) - \beta(M_h)| = 0$.

Así, decimos que un proceso de división rectangular es normal si cualquier sucesión infinita anidada de rectángulos que genera es normal.

Construcción del proceso de subdivisión normal rectangular

La construcción del proceso de subdivisión normal rectangular se hará a través del refinamiento débil y es el que se desarrolla a continuación.

Consideremos al poliedro M_k el cual satisface que $\beta(M_k) < f(x^k) - \epsilon$, de este modo, tomando $\omega^k = \omega(M_k)$, $\varphi_k(x) = \varphi_{M_k}$, $\varphi_{k,j}(x_j) = \varphi_{M_{k,j}}(x_j)$ tenemos que $f(\omega^k) - \varphi_k(\omega^k) > 0$.

1. Sea $M_k = \{x : r^k \leq x \leq s^k\}$.
2. Escoger un índice $j_k \in \{1, 2, ..., n\}$ y un número $w^k \in (r^k_{j_k}, s^k_{j_k})$.
3. Usar el hiperplano $x_{j_k} = w^k$ y subdividir M_k en dos rectángulos

$$M_{k,1} = \{x \in M_k : x_{j_k} \leq w^k\}, \quad M_{k,2} = \{x \in M_k : x_{j_k} \geq w^k\}$$

Proposición 10. *La forma de elegir j_k y ω^k constituyen una regla de subdivisión y generan un proceso de subdivisión normal rectangular si satisfacen cualquiera de las siguientes condiciones:*

i) $j_k \in argmax_j |f_j(\omega_j^k) - \varphi_{kj}(\omega_j^k)|$ *y* $w^k = \omega^k$

ii) $j_k \in argmax_j \delta_{kj}$ *y* $\omega^k = (r_{j_k}^k + s_{j_k}^k)/2$, *donde* δ_{kj} *es tal que* $f_j(\omega_j^k) - \varphi_{kj}(\omega_j^k) \leq \delta_{kj}$, $\delta_{kj} \to 0$ *si* $s_j^k - r_j^k \to 0$

Observemos que en este refinamiento si escogemos $w^k = (r_{j_k}^k + s_{j_k}^k)/2$ entonces estaríamos realizando el proceso de partición a través de bisección.

4.4.1. Algoritmo normal rectangular

Elegir $\epsilon \geq 0$

Inicialización.

Elegir una rectángulo M_0 que contiene a D. Sea x^0 la mejor solución factible hasta ahora.

Sea $\mathcal{M}_0 = \mathcal{N}_0 = \{M_0\}$.

Iteración k=0,1,...:

1) Para cada $M \in \mathcal{N}_k$ calcula la función afín $\varphi_M(x)$ que coincida con $f(x)$ en los vértices de M y resolver el siguiente problema de programación lineal:

$$\min \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

obtener la solución óptima básica $\omega(M)$ y el valor óptimo $\beta(M)$.

2) Eliminar de $M \in \mathcal{M}_k$ tal que $\beta(M) \geq f(x^k) - \epsilon$.

   Sea $\mathcal{R}_k$ la colección restante de simplices.

3) Si $\mathcal{R} = \varnothing$, Fin: x^k es una solución ϵ óptima global, sino ir a 4

4) Elegir $M_k \in argmin\{\beta(M) : M \in \mathcal{R}_k\}$ y subdivide del siguiente modo:

   a) Sea $M_k = \{x : r^k \leq x \leq s^k\}$.
   b) Escoger un índice $j_k \in \{1, 2, ..., n\}$ y un número $w^k \in (r_{j_k}^k, s_{j_k}^k)$.
   c) Usar el hiperplano $x_{j_k} = w^k$ y subdividir M_k en dos rectángulos

$$M_{k,1} = \{x \in M_k : x_{j_k} \leq w^k\}, \quad M_{k,2} = \{x \in M_k : x_{j_k} \geq w^k\}$$

5) Actualiza los datos, x^{k+1} será la mejor solución conocida hasta este punto.

   Sea $\mathcal{N}_{k+1}$ el conjunto de subsimplices de M_k que se generaron en la subdivisión en el paso 4.

   Sea $\mathcal{M}_{k+1} = (\mathcal{R}_k \diagdown \{M_k\}) \cup \mathcal{N}_{k+1}$.

   Sea $k \leftarrow k+1$ y regresar al paso 1.

El siguiente teorema establece la convergencia del algoritmo.

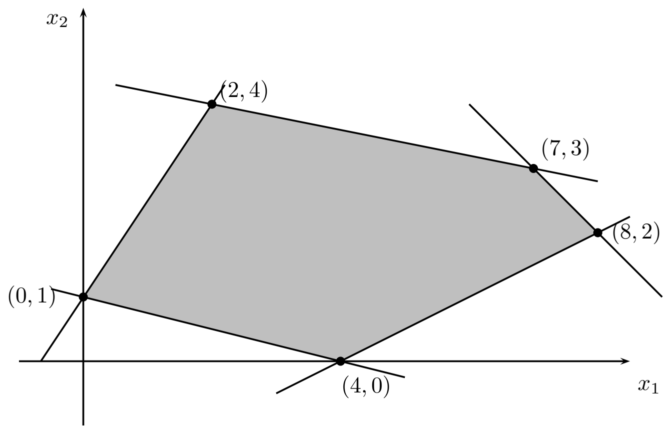
Teorema 9. *El algoritmo normal rectangular puede ser infinito si $\epsilon = 0$ y en este caso genera una sucesión $\{x^k\}$ donde todo punto de acumulación es una solución óptima global.*

Ejemplo.

$$\text{mín} \quad -x_1 - 4x_2^{3/2}$$

$$\begin{array}{llrl}
x_1 & + & x_2 & \leq 10 \\
x_1 & + & 5x_2 & \leq 22 \\
-3x_1 & - & 2x_2 & \leq 2 \\
-x_1 & - & 4x_2 & \leq -4 \\
x_1 & - & 2x_2 & \leq 4 \\
x_1 & & & \geq 0 \\
 & & x_2 & \geq 0
\end{array}$$

La siguiente gráfica representa la región de soluciones factibles del planteamiento anterior.



Sea $\epsilon = 0.2$.

Sea $M_0 = \{x : 0 \leq x_1 \leq 8, 0 \leq x_2 \leq 4\}$.

Sea $x^0 = (2,4)$ $f(x^0) = -36$ la mejor solución disponible. Sea $\mathcal{M}_0 = \mathcal{N}_0 = \{M_0\}$

Iteración 0.

1) Para cada $M \in \mathcal{N}_0$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

   $\varphi(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1,2,3$. resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & & & a_3 & = 0 \\ & & 4a_2 & + & a_3 & = -32 \\ 8a_1 & & & + & a_3 & = -64 \\ 8a_1 & + & 4a_2 & + & a_3 & = -96 \end{array}$$

   de aquí tenemos que $a_1 = -8$, $a_2 = -8$, $a_3 = 0$.

   Por lo tanto la función afín es $\varphi_M(x) = -8x_1 - 8x_2$

   Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

   es decir

$$\text{mín} \quad -8x_1 - 8x_2 \quad \text{s.a} \quad x \in D \cap M$$

   La solución al programa lineal es: $\omega(M) = (8,2)$ y su valor óptimo es $\beta(M) = -80$

2) Eliminar toda $M \in \mathcal{M}_0$ tal que $\beta(M) \geq f(x^0) - \epsilon$, donde $f(x^0) - \epsilon = -36 - 0.2 = -36.2$, como $-80 < -36.2$ no se elimina nada.

   Sea $\mathcal{R}_0 = \{M_0\}$

3) Como $\mathcal{R}_0 \neq \varnothing$ ir a 4

4) Elegir $M_0 \in argmin\{\beta(M) : M \in \mathcal{R}_0\}$.

   $M_0 = \{x : 0 \leq x_1 \leq 8, 0 \leq x_2 \leq 4\}$ subdividir de acuerdo al proceso elegido.

   Sea $j_0 \in \{1,2\}$ $w^0 \in (r^0_{j0}, s^0_{j0})$ y usando el hiperplano $x_{j0} = w^0$ subdividiremos a M_0 en dos rectángulos. Sea $j_0 = 1$ entonces $w^0 \in (0,8)$, sea $w^0 = 4$. Así el hiperplano $x_1 = 4$ genera la siguiente partición:

   $M_{01} = \{x \in M_0 : x_1 \leq 4\}$ y $M_{02} = \{x \in M_0 : x_1 \geq 4\}$

5) Sea $x^1 = (8,2)$ $f(x^1) = -75.3137$. Sea $\mathcal{N}_1 = \{M_{01}, M_{02}\}$. Sea $\mathcal{M}_1 = (\mathcal{R}_0 \diagdown \{M_0\}) \cup \mathcal{N}_\infty$, así, $\mathcal{M}_1 = \mathcal{N}_1$ pues $\mathcal{R}_0 = \{M_0\}$

   Sea $k \leftarrow 1$ ir a 1

Iteración 1.

1) Para cada $M \in \mathcal{N}_1$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

   $\varphi_{M_{01}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} & & & & a_3 & = 0 \\ & & 4a_2 & + & a_3 & = -32 \\ 4a_1 & & & + & a_3 & = -16 \\ 4a_1 & + & 4a_2 & + & a_3 & = -48 \end{array}$$

de aquí tenemos que $a_1 = -4$, $a_2 = -8$, $a_3 = 0$.

   Por lo tanto la función afín es $\varphi_M(x) = -4x_1 - 8x_2$

   Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

   es decir

$$\text{mín} \quad -4x_1 - 8x_2 \quad \text{s.a} \quad x \in D \cap M \cap x_1 \leq 4$$

   La solución al programa lineal es: $\omega(M_{01}) = (4, 3.6)$ y su valor óptimo es $\beta(M_{01}) = -44.8$

   Veamos que sucede con el otro subsimplex: $\varphi_{M_{02}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{ccccccc} 4a_1 & & & + & a_3 & = -16 \\ 4a_1 & + & 4a_2 & + & a_3 & = -48 \\ 8a_1 & & & + & a_3 & = -64 \\ 8a_1 & + & 4a_2 & + & a_3 & = -96 \end{array}$$

de aquí tenemos que $a_1 = -12$, $a_2 = -8$, $a_3 = 32$.

   Por lo tanto la función afín es $\varphi_M(x) = -12x_1 - 8x_2 + 32$

   Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad -12x_1 - 8x_2 + 32 \quad \text{s.a} \quad x \in D \cap M \cap x_1 \geq 4$$

La solución al programa lineal es: $\omega(M_{02}) = (8, 2)$ y su valor óptimo es $\beta(M_{02}) = -80$

2) Eliminar toda $M \in \mathcal{M}_1$ tal que $\beta(M) \geq f(x^1) - \epsilon$, donde $f(x^1) - \epsilon = -75.3137 - 0.2 = -75.5137$, como $-36 \leq -36.2$ eliminamos M_{01}.

   Sea $\mathcal{R}_1 = \{M_{02}\}$

3) Como $\mathcal{R}_1 \neq \varnothing$ ir a 4

4) Elegir $M_1 \in argmin\{\beta(M) : M \in \mathcal{R}_1\}$.

   $M_1 = \{x : 4 \leq x_1 \leq 8, 0 \leq x_2 \leq 4\}$ subdividir de acuerdo al proceso elegido.

   $j_i \in \{1, 2\}$, sea $j_1 = 2$ y $w^1 \in (r_2^1, s_2^1) = (0, 4)$.

   Así, usamos el hiperplano $x^2 = 2$ donde $w^1 = 2$. Ahora generamos la siguiente partición:

   $M_{11} = \{x \in M_1 : x_2 \leq 2\}$ y $M_{12} = \{x \in M_1 : x_2 \geq 2\}$

5) Sea $x^2 = (8, 2)$ $f(x^2) = -75.3137$. Sea $\mathcal{N}_2 = \{M_{11}, M_{12}\}$. Sea $\mathcal{M}_2 = (\mathcal{R}_1 \diagdown \{M_1\}) \cup \mathcal{N}_\in$, así, $\mathcal{M}_2 = \mathcal{N}_2$ pues $\mathcal{R}_1 = \{M_1\}$

   Sea $k \leftarrow 2$ ir a 1

Iteración 2.

1) Para cada $M \in \mathcal{N}_2$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

   $\varphi_{M_{11}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{ccccccl} 4a_1 & & & + & a_3 & = -16 \\ 4a_1 & + & 2a_2 & + & a_3 & = -27.3137 \\ 8a_1 & & & + & a_3 & = -64 \\ 8a_1 & + & 2a_2 & + & a_3 & = -75.3137 \end{array}$$

de aquí tenemos que $a_1 = -12$, $a_2 = -5.6568$, $a_3 = 32$.

Por lo tanto la función afín es $\varphi_M(x) = -12x_1 - 5.6568x_2 + 32$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad -12x_1 - 5.6568x_2 + 32 \quad \text{s.a} \quad x \in D \cap M \cap x_2 \leq 2$$

La solución al programa lineal es: $\omega(M_{11}) = (8,2)$ y su valor óptimo es $\beta(M_{11}) = -75.3136$

Veamos que sucede con el otro subsimplex: $\varphi_{M_{12}}(x_1, x_2) = a_1x_1+a_2x_2+a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{lllllll} 4a_1 & + & 2a_2 & + & a_3 & = -27.3137 \\ 8a_1 & + & 2a_2 & + & a_3 & = -75.3137 \\ 8a_1 & + & 4a_2 & + & a_3 & = -96 \\ 8a_1 & & & + & a_3 & = -64 \end{array}$$

de aquí tenemos que $a_1 = -12$, $a_2 = -8$, $a_3 = 32$.

Por lo tanto la función afín es $\varphi_M(x) = -12x_1 - 8x_2 + 32$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad -12x_1 - 8x_2 + 32 \quad \text{s.a} \quad x \in D \cap M \cap x_2 \geq 2$$

La solución al programa lineal es: $\omega(M_{12}) = (8,2)$ y su valor óptimo es $\beta(M_{12}) = -80$

2) Eliminar toda $M \in \mathcal{M}_2$ tal que $\beta(M) \geq f(x^2) - \epsilon$, donde $f(x^2) - \epsilon = -75.5137$, como $-64 \leq -75.5137$ eliminamos M_{01}.

   Sea $\mathcal{R}_2 = \{M_{12}\}$

3) Como $\mathcal{R}_2 \neq \varnothing$ ir a 4

4) Elegir $M_2 \in argmin\{\beta(M) : M \in \mathcal{R}_2\}$.

   $M_2 = \{x : 4 \leq x_1 \leq 8, 2 \leq x_2 \leq 4\}$ subdividir de acuerdo al proceso elegido.

   $j_2 \in \{1, 2\}$, sea $j_2 = 2$ y $w^2 \in (r_2^1, s_2^1) = (2, 4)$.

   Así, usamos el hiperplano $x^2 = 3$ donde $w^2 = 3$. Ahora generamos la siguiente partición:

   $M_{21} = \{x \in M_2 : x_2 \leq 3\}$ y $M_{22} = \{x \in M_2 : x_2 \geq 3\}$

5) Sea $x^3 = (8,2)$ $f(x^3) = -75.3137$. Sea $\mathcal{N}_3 = \{M_{21}, M_{22}\}$. Sea $\mathcal{M}_3 = (\mathcal{R}_2 \diagdown \{M_2\}) \cup \mathcal{N}_\ni$, así, $\mathcal{M}_3 = \mathcal{N}_3$ pues $\mathcal{R}_2 = \{M_2\}$

   Sea $k \leftarrow 3$ ir a 1

Iteración 3.

1) Para cada $M \in \mathcal{N}_3$ construye la función afín $\varphi_M(x)$ que coincide con $f(x)$ en los vértices de M.

$\varphi_{M_{21}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{lllllll} 4a_1 & + & 2a_2 & + & a_3 & = -27.3137 \\ 4a_1 & + & 3a_2 & + & a_3 & = -36.7846 \\ 8a_1 & + & 2a_2 & + & a_3 & = -75.3137 \\ 8a_1 & + & 3a_2 & + & a_3 & = -84.7846 \end{array}$$

de aquí tenemos que $a_1 = -12$, $a_2 = -9.4709$, $a_3 = 39.6281$.

Por lo tanto la función afín es $\varphi_M(x) = -12x_1 - 9.4709x_2 + 39.6281$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad -12x_1 - 9.4709x_2 + 39.6281 \quad \text{s.a} \quad x \in D \cap M \cap x_2 \leq 3$$

La solución al programa lineal es: $\omega(M_{21}) = (8, 2)$ y su valor óptimo es $\beta(M_{21}) = -75.3137$

Veamos que sucede con el otro subsimplex: $\varphi_{M_{22}}(x_1, x_2) = a_1x_1 + a_2x_2 + a_3$, para obtener el valor numérico de las a_i $i = 1, 2, 3$. resolvemos el siguiente sistema:

$$\begin{array}{lllllll} 4a_1 & + & 3a_2 & + & a_3 & = -36.7846 \\ 4a_1 & + & 4a_2 & + & a_3 & = -48 \\ 8a_1 & + & 3a_2 & + & a_3 & = -84.7846 \\ 8a_1 & + & 4a_2 & + & a_3 & = -96 \end{array}$$

de aquí tenemos que $a_1 = -12$, $a_2 = -11.2154$, $a_3 = 44.8616$.

Por lo tanto la función afín es $\varphi_M(x) = -12x_1 - 11.2154x_2 + 44.8616$

Resolver el siguiente programa lineal

$$\text{mín} \quad \varphi_M(x) \quad \text{s.a} \quad x \in D \cap M$$

es decir

$$\text{mín} \quad -12x_1 - 11.2154x_2 + 44.8616 \quad \text{s.a} \quad x \in D \cap M \cap x_2 \geq 3$$

La solución al programa lineal es: $\omega(M_{22}) = (7,3)$ y su valor óptimo es $\beta(M_{22}) = -72.7846$

2) Eliminar toda $M \in \mathcal{M}_3$ tal que $\beta(M) \geq f(x^3) - \epsilon$, donde $f(x^3) - \epsilon = -75.5137$, como $-75.3137 > -75.5137$ y $-72.7846 > -75.5137$ entonces $\mathcal{R}_3 = \varnothing$.

3) Como $\mathcal{R}_3 = \varnothing$ Fin; x^3 es una solución ϵ global.

Es importante observar que en los dos primeros algoritmos de ramificación y acotamiento presentados en este capítulo; algoritmos normal cónico y simplicial a diferencia del rectangular, dentro de su criterio de decisión para realizar la división de la región factible a través de una partición, usan un conjunto denotado por Δ, el cual como vimos esta formado por una sucesión infinita de números naturales o múltiplos de éste. Debido a la naturaleza del conjunto, contendrá ciertos elementos y otros no, lo cuál hace que la rapidez con la que se llegue a concluir dependa fuertemente de la elección de dicho conjunto.

5 Aplicación

El objetivo de este capítulo es ejemplificar el uso de los métodos estudiados a lo largo del presente trabajo. Encontrando así, dos aplicaciones, la primera de ella a un problema de la vida real como lo es el de extraer agua de los acuíferos. La segunda de carácter teórico, es la aportación de [6] al desarrollar un algoritmo de ramificación y acotamiento para resolver problemas de flujo en redes con costos cóncavos.

5.1. Extracción de agua

El problema de extracción del agua de los mantos acuíferos es de general importancia, pues su propósito es tener agua disponible para satisfacer a las poblaciones que carecen de ello. Pero el problema que se estudia es la de extraer el agua al menor costo posible. Varias estudios a cerca de este tipo de problemas a través de optimización se han realizado entre los cuales se pueden destacar (Aguado, Remson 1974) por incluir a las variables del agua subterránea como variables de decisión directas en el modelo de programación lineal, usando el método de diferencias finitas para llevar las ecuaciones diferenciales parciales que modelan a las restricciones a unas de tipo lineal. Meses más tarde, los autores basados en (Aguado, Remson 1974) proponen la manera óptima de resolver el problemas lineal, determinando el número óptimo de pozos, su localización y la tasa de bombeo necesaria para mantener los niveles del agua por debajo de lo especificado por el estado estacionario. A diferencia de estas investigaciones, el artículo que lleva por título *Groundwater management using numerical simulation and outer approximation methods for global optimization* [9] considera ya que la función objetivo es de tipo cóncava. De este modo se expondrá a grandes rasgos la implementación del método de aproximación exterior utilizado.

5.1.1. Planteamiento del Problema

El problema que se desea resolver es sobre administración de aguas subterráneas con cargo fijos. De acuerdo a [9] los modelos sobre la administración de aguas subterráneas que se habían planteado antes, no consideraban dentro del modelo el costo de la instalación de pozos, por considerarse que tenía un costo pequeño comparado con los costos de operación y mantenimiento. En dicho artículo, si se incluyen los costos de instalación de pozos dentro de la función objetivo, la cuál resulta ser cóncava cuyas restricciones son un conjunto convexo compacto. Esto da lugar a un modelo de minimización cóncava. Veamos el ejemplo que presentan;

5.1.2. Extracción de agua de un acuífero

El acuífero se representa a través de un esquema de elementos finitos, las restricciones son impuestas sobre el máximo valor de la cabeza hidráulica en nodos específicos. Se introduce un número de pozos potenciales de bombeo y el objetivo es satisfacer las restricciones con el mínimo bombeo. El costo total incluye costos de tratamiento, mantenimiento y el costo de instalación de un pozo fijo. La formulación del problema es la siguiente:

Antes de mencionar el planteamiento es importante familiarizarnos un poco con ciertos términos que son del área de la hidráulica.

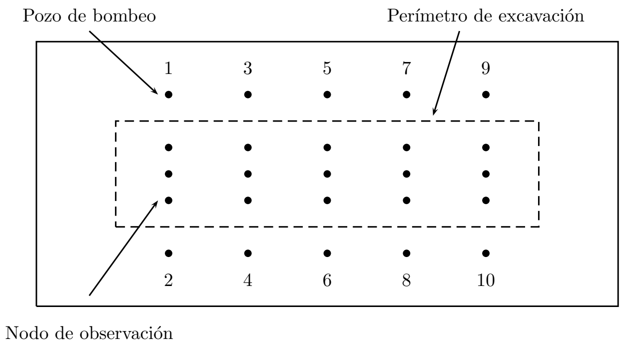
Cabeza hidráulica (Hydraulic head)(h): Es una medida específica de presión del agua o energía total por unidad de peso sobre el datum.

Datum: Un datum (plural datums or data) es un punto de referencia sobre el cual se hacen mediciones.

Conductividad hidráulica (K): Es una constante de proporcionalidad en la ley de Darcy, relaciona la cantidad de agua que fluirá a través de una unidad de área del acuífero bajo una unidad de gradiente de la cabeza hidráulica.

Isotrópico: Objeto que presenta uniformidad en toda dirección.

Un área de excavación rectangular de 200×30m es localizada en el centro de un acuífero libre cuyas dimensiones son $L_x = 400$m y $L_y = 100$m. El fondo del acuífero es considerado el datum. Existe una cabeza hidráulica constante de 36m por arriba del datum alrededor de toda la frontera del acuífero. La conductividad hidráulica del arena homogénea isotrópica del acuífero es 3.05m/d. Es necesario que la cabeza hidráulica junto con el sitio de excavación se mantengan en o por debajo de los 30.5m encima del datum a través de un bombeo en estado estacionario.



Lo que se desea es minimizar los costos de instalación de costo, los costo de funcionamientos, sujetos a las restricciones de la propia naturaleza del acuífero, adicionalmente se pide ciertas características sobre la cabeza hidráulica y el sitio de excavación. Para poder llevar este problema a un modelo matemático de minimización cóncava con restricciones lineales, primero hay que tomar en cuenta que el sistema de aguas profundas se describe a través de la siguiente ecuación diferencial parcial

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} - \frac{2}{K} q = 0$$

donde:

ϕ	h^2, (L^2);
h	Cabeza hidráulica encima del datum, (L)
K	Conductividad hidráulica, (LT^{-1})
q	Tasa de descarga por unidad de área,(LT^{-1})

Esta es la ecuación de flujo en estado estacionario para un acuífero libre homogéneo isotrópico. Para poder llevarlas a un modelo lineal los autores utilizaron la simulación numérica a través del método de elementos finitos; suponen que el flujo se encuentra en estado estable. El acuífero libre es aproximado en la simulación numérica con la ecuación de flujo de un acuífero confinado, dicha ecuación se representa por:

$$A\bar{\phi} = \{b + q\} \tag{5.1}$$

donde

A	matriz de coeficientes de flujo;
$\bar{\phi}$	vector de la cabeza hidráulica;
b	vector que tiene información sobre las condiciones de frontera;
q	vector de tasa de bombeo;

Resolviendo 5.1 para la cabeza hidráulica se obtiene

$$\bar{\phi} = [A]^{-1}\{b+q\} \tag{5.2}$$

donde $[A_r]$ es la matriz que contiene sólo los renglones de $[A]^{-1}$ que corresponden las restricciones de la cabeza hidráulica definidas como puntos de observación, es decir, en los nodos donde las restricciones fueron impuestas. De este modo

$$\phi = [A_r]\{b+q\}$$

donde ϕ es el vector que contiene los valores de la cabeza hidráulica en los nodos de observación. Entonces la ecuación anterior representa las restricciones, que para este problema son de tipo lineal. La tasa de bombeo q representa a las variables de decisión.

$$\begin{array}{ll} \text{mín}\, f(q) = & \Sigma_{i=1}^{n}\alpha_i q_i \,+\, \alpha_{0_i}(1-e^{-bq_i}) \\ \text{tal que} & \\ & \phi_j \leq \phi_j* \quad j \in J \\ & 0 \leq q_i \leq q_i* \quad i \in I \end{array}$$

donde

f	es una función continua cóncava
α_i	es el costo de bombeo por unidad
α_{0_i}	es el costo fijo de instalar un pozo en el nodo i
$b=1$	es un factor escalar
$\phi_j = h_i^2$	h es la cabeza hidráulica
ϕ_j*	es un valor fijo
q_i	es la taza de bombeo en el nodo i
q_i*	es la cota superior fija de q_i
I	es el conjunto de puntos cuyo número de elementos es igual al número de parámetros (pozos de bombeo)
J	es el conjunto de puntos cuyo número de elementos es igual al número de restricciones (nodos de observación).

El último término de la función objetivo representa la aproximación del consto fijo de instalación del pozo en el lugar i. Este término desaparece si no hay bombeo en el nodo i. En caso de que el bombeo se lleve a cabo en el nodo i, es necesario que el término e^{-bq_i} sea muy pequeño, de tal manera que la expresión $(1-e^{-bq_i})$ sea casi igual a 1. Ajustando el término bq_i a través del factor escalar b, es que se logra que el término e^{-bq_i} sea muy pequeño.

Una vez que se tiene el modelo y todos los datos, es posible aplicarle el algoritmo de aproximación exterior. Los resultados fueron; la ubicación óptima de los pozos y la tasa de bombeo óptima se muestran a continuación:

Pozo	1	2	3	4	5	6	7	8	9	10
q	6,190	–	–	–	–	5,469	–	–	6,190	–

donde q es la tasa de bombeo la cual se mide en m^3/d.

Cabe mencionar que en él artículo consultado, se compararon los resultados que arrojó el método de aproximación exterior con los que dio el mismo problema pero formulado como un problema de programación lineal entera mixta, siendo el primero un mejor método de solución.

5.2. Flujo en redes con costos cóncavos

Debido a la gran variedad de problemas que son resueltos a través de los algoritmos en redes es que encontrar nuevas y mejores formas de implementación resulta una herramienta poderosa, ampliando igualmente su rango de aplicación. Este es el caso del artículo que lleva por título *A Branch and Bound Algorithm for Concave Network Flow Problems* cuyos autores son Dalila B.M.M. Fontes, Elene Hadjiconstantinou y Nicos Christofides. En éste veremos cómo es implementado el método de ramificación y acotamiento para resolver problemas de flujo en redes cuyos costos están definidos por una función cóncava. Este tipo de problemas son considerados $NP-$difíciles; es decir, son problemas al menos tan difíciles como un problema de NP, no necesariamente siendo un problema NP. Es importante mencionar que los problemas de flujo en redes cuya función de costo es no lineal, puede ser transformado en uno donde la función sea cóncava, de ahí la importancia de encontrar una manera eficiente de resolverlo. Varios son los esfuerzos que se han hecho al respecto, entre los métodos utilizados para ello podemos mencionar a los métodos exactos, heurísticos, a la programación dinámica, etc. [6].

5.2.1. Planteamiento del problema

Sea $G = (W, A)$ la gráfica que deseamos optimizar, W es el conjunto de vértices, donde $|W| = n + 1$ y w_i con $i = 1, ..., n$ son los nodos de demanda siendo el último nodo $w_{n+1} = t$ el nodo de oferta o nodo fuente. A es el conjunto de arcos dirigidos, $A = \{(i, j) : i \in W,\ j \in W \setminus t\ \}$ Cada nodo demanda tiene asociado una demanda entera positiva r_i. La oferta en el nodo fuente es igual a la suma de las demandas requeridas en los n nodos demanda. Se considera una solución al flujo r_{ij} asociado a cada arco $(i, j) \in A$. Una función cóncava, no decreciente no negativa g_{ij} se asocia a cada arco (i, j) y cumple $g_{ij}(0) = 0$. El objetivo es encontrar un subconjunto de arcos y flujo sobre los arcos que satisfagan la demanda al costo mínimo. Este tipo de problemas se le conoce como Problemas cóncavos de flujo en redes con fuente única sin límite de capacidad y tendrán solución si existe un camino dirigido desde el nodo fuente hacia todos los nodos demanda, además se pide que no existan ciclos con costo negativos. Un flujo factible sin ciclos positivos, para el caso en que no hay límite de capacidad y hay sólo un único nodo fuente, es un árbol enraizado en el nodo fuente generando todos los vértices demanda.

La siguiente es una formulación de programación dinámica que desarrollaron (Fontes et al 2005a); hay independencia en cuanto al tipo de la función de costo, al número de arcos de costos no lineales, supuestos de separabilidad y aditividad.

Sea $S \subseteq W$ y un vértice $x \in S$. Sea $\{S', \bar{S'}\}$ una partición de S, donde $S' \subseteq S \diagdown \{x\}$ y $\bar{S'} = S \diagdown S'$. Para cada posible conjunto S' sea $z \in S'$ el vértice raíz de un árbol dirigido generando el conjunto S'. Sea $f(S', z)$ el costo mínimo de abastecer todos los vértices demanda en S' con la materia prima disponible en el vértice z a través de un árbol dirigido enraizado en z. El costo mínimo de abastecer un conjunto S' desde un vértice $x \notin S'$ con la materia prima disponible en algún vértice $z \in S'$ es encontrado determinando la mejor combinación del costo mínimo del árbol dirigido de S' enraizado en el vértice $z \in S'$ con el costo del arco (x, z), es decir, $\text{mín}_{z \in S'}\{f(S', z) + g_{xz}(\Sigma_{i \in S'} r_i)\}$.

El costo mínimo incurrido en abastecer los restantes vértices demanda del conjunto S que no están en S' desde x está dado por $f(\bar{S'}, x)$. Por lo dicho anteriormente, el costo mínimo $f(S, x)$ de abastecer todos los vértices demanda en S con la materia disponible en $x \in S$ es obtenida a través de examinar todos los posibles conjuntos $S' \subseteq S \diagdown \{x\}$ y dado por:

$$f(S, x) = \underset{S' \subseteq S \diagdown \{x\}}{\text{mín}} [f(S - S', x) + \underset{z \in S'}{\text{mín}} [f(S', z) + g_{xz}(\Sigma_{i \in S'} r_i)]] \tag{5.3}$$

Las condiciones iniciales para la ecuación 5.3 está dada por $f(\{x\}, x) = 0$ $\forall\ x \in W$. Así, el costo óptimo de abastecer todos los vértices demanda en el conjunto W desde el vértice fuente t está dado por:

$$f* \equiv f(W, t) = \underset{S' \subseteq W \diagdown \{t\}}{\text{mín}} [f(W - S', t) + \underset{z \in S'}{\text{mín}} [f(S', z) + g_{tz}(\Sigma_{i \in S'} r_i)]] \tag{5.4}$$

Dentro del algoritmo propuesto se usa como subrutina un procedimiento llamado: Relajación del espacio estado. Como el espacio estado es de gran tamaño, son pocos los problemas combinatorios que se han resuelto eficientemente sólo usando programación dinámica. La relajación del espacio estado es un procedimiento para relajar el estado del espacio asociado al proceso recursivo de programación dinámica. La solución relajada de la recursión da una cota inferior, en el caso de minimización, al valor de la solución óptima.

5.2.2. Implementando el método Ramificación y Acotamiento

Los principales elementos del método de ramificación y acotamiento son los subproblemas que debemos seleccionar, resolver acotando y revisando el criterio de eliminación y dividir, si es que el subproblema no ha sido dividido. La evolución del algoritmo es almacenada en un árbol binario.

Proceso de Ramificación

El proceso de ramificación divide un nodo árbol, correspondiente a un subproblema en dos subproblemas mutuamente excluyentes y exhaustivos. De este modo, el algoritmo va construyendo un árbol binario de subproblemas. La estrategia de ramificación involucra la elección de un arco del conjunto de arcos disponibles en un nodo árbol sobre el cual es próximo a ramificar. Supongamos que se escoge un arco (x, y) en un nodo árbol específico, donde el vértice x es un vértice ya fijo que forma parte de la solución o bien, es el vértice fuente que abastece directamente al vértice y generando así una solución parcial. De este modo, se garantiza que la gráfica de arcos fijos es conexa, de hecho un árbol. La primer ramificación se fija el arco en la solución. La alternativa es rechazar el arco (x, y).

La regla para ramificar escoge el arco próximo a ramificar entre los arcos obtenidos en la solución asociada con la cota inferior del actual nodo árbol. Los siguientes pasos son aplicados para elegir un arco (x, y) donde x representa el nodo fijo o el nodo fuente y y el nodo candidato a elegir.

a) Vértices con mayor sobre-oferta correspondiente a un arco que es usado con mucha frecuencia o bien a arcos asociados con ciclos,

$$\frac{\Sigma_{x \in W} r_{zy} - \Sigma_{z \in V} r_{yz}}{r_y}$$

donde r_{zy} es el flujo del arco (z, y) y r_y es la demanda del vértice y. Observemos que la proporción anterior es mayor o igual que uno sólo si el vértice y es abastecido más de una vez, ya sea por el uso repetido del arco (x, y) o bien por al menos dos diferentes arcos.

b) Si no hay arcos causantes de la no factibilidad para fijarlos, entonces se escoge al más cercano a tal arco. La distancia es medida en términos de número de arcos que separan a los dos arcos en cuestión.

c) Si es encontrado más de un arco (después de aplicar a) o b)) entonces se deben seguir los siguientes criterios para escoger un arco tal que el vértice que será fijo tenga: la demanda más alta, la profundidad más baja y el espacio reducido más grande.

Regla para elegir nodo

Otra regla de elección que hay que tomar en cuenta cuando se tiene más de un subproblema es cuál de éstos se elegirá. Se propone tomar el siguiente nodo cuya diferencia entre su cota inferior y la mejor cota superior disponible sea la más grande. En la búsqueda de la solución óptima a fuerza este nodo será visitado, pues su cota inferior es la más pequeña, de este modo no podrá ser eliminado. En esta implementación cuando un nodo es considerado para más particiones entonces ambos hijos (los nodos que corresponden a los dos subproblemas siguientes) se resuelven simultáneamente.

Proceso de Acotamiento

Las cotas se obtienen en cada nodo árbol a través de resolver el correspondiente subproblema relajado usando una versión modificada del procedimiento Espacio Estado Ascendente (Fontes 2005b). Cada uno de estos subproblemas tiene por característica dos tipos de arcos, los que no puede pertenecer a la solución y los que deben pertenecer a la solución, los cuales se consideran fijos. Dado que el problema que deseamos resolver es de minimización, una vez determinados los arcos que no serán parte de la solución, a través de las decisiones en el proceso de ramificación, se hace que el costo de dichos arcos sea muy grande. Así se asegura que no formarán parte de la solución.

Una manera de tomar los arcos fijos como parte de la solución es eliminar del problema resuelto la parte correspondiente a las decisiones ya tomadas. Es decir, si un arco (t, x) es fijado como parte de la solución entonces el problema a resolver es el abastecimiento de los vértices restantes. Un arco (t, x) no puede ser eliminado de ningún subproblema ya resuelto, pues es posible que en la solución óptima sea usado para que pase una cantidad más grande de flujo que la del flujo ya fijado dirigido hacia éste. La función de costo debe cambiar para que refleje el hecho de que el arco está siendo usado y cierta cantidad de flujo está pasando hacia él. Cuando el arco (t, x) es fijado como parte de la solución, r_x la demanda del vértice x es dirigida hacia éste. Supongamos que otro arco (x, y) es también fijado como parte de la solución, el flujo sobre estos arcos es $r_x + r_y$ y r_y respectivamente. Como la función de costos no es lineal, entonces se modifica del siguiente modo:

$$g''_{xy}(r_{xy}) = \begin{cases} g'_{xy}(r_{xy}) & si\ (x,y)\ no\ es\ fijo \\ g_{xy}(r_{xy}+r) - g_{xy}(r) + \lambda_y & en\ otro\ caso \end{cases}$$

donde r es el flujo acumulado previamente fijado en el arco (x, y) y $g'_{xy}(r_{xy}) = g_{xy}(r_{xy}) + \lambda_y$.

Sea u el actual nodo de búsqueda y F_u el conjunto de vértices correspondientes al extremo derecho del conjunto fijado como parte de la solución. La cota inferior se obtiene a través de

$$Z_{LB} = f(n + 1 - |F_u|, Q_u, R_u, t) - \Sigma_{i \in V \setminus F_u} \lambda_i$$

donde $Q_u = \Sigma_{i \in V \setminus F_u} q_i$, $R_u = \Sigma_{i \in V \setminus F_u} r_i$ y V es el conjunto de los n vértices demanda $V = W \setminus \{t\}$. Así, el espacio estado del actual subproblema es mucho más pequeño que el espacio estado del problema original en $|F_u|$ arcos, de este modo los vértices has sido fijados como parte de la solución final. [1]

A continuación veremos el algoritmo que desarrollaron [6] donde suponen que el nodos raíz ya ha sido resuelto.

1. Inicialización.

   $A_f = \varnothing$ y $A_e = \varnothing$

   Leer el nodo raíz.

   a) Leer el valor y estructura de la cota inferior.
   b) Leer los vectores de penalidades y pesos.
   c) Escoger el arco a ramificar (x, y).
   d) Almacenar el nodo raíz.

2. Resolver al hijo derecho

   a) Eliminar el arco (x, y), $A_e = A_e \cup \{(x, y)\}$.
   b) Calcular el número de arcos fijos que salen del nodo fuente (s), y calcular el número total de arcos fijos $(m = |A_f|)$.
   c) Si algún nodo demanda no está conectado entonces
      i) Elimina el nodo.
      ii) Ir a 3
   d) Calcular $Z_{LB} = f(P, Q, R, t, s) - \Sigma_{i \in V} \lambda_i$ como en la ecuación [2]
   e) Si la solución es factible entonces
      i) Si $Z_{LB} < Z_{UB}$ entonces actualizar Z_{UB}, el mejor encontrado hasta ahora.
      ii) Elimina el nodo.
      iii) Ir a 3
   f) Si $Z_{LB} > Z_{UB}$ entonces
      i) Elimina el nodo.
      ii) Ir a 3
   g) Si $Z_{LB} > Z*_{LB}$ actualiza $Z*_{LB}$
   h) Si el tiempo/iteraciones están en su límite entonces
      i) Actualiza las penalidades/peso
      ii) Ir a 2(d).

[1] detalles sobre la relajación ver [6]
[2] ecuación (28) de [6]

i) Almacena los datos del nodo

3. Resolver al hijo izquierdo

   a) Un Arco (x, y), $A_f = A_f \cup \{(x, y)\}$.
   b) Actualiza el número de arcos fijos que salen del vértice fuente. Actualiza el número de arcos fijos $m = m + 1$.
   c) Si $m = n$ entonces
      i) Calcula es costo de las estructuras fijas, esto es una solución factible.
      ii) Ir a 3(e).
   d) Calcula $Z_{LB} = f(P, Q, R, t, s) - \Sigma_{i \in V} \lambda_i$ como en la ecuación [3]
   e) Si la solución es factible entonces
      i) Si $Z_{LB} < Z_{UB}$ actualiza Z_{UB} el mejor encontrado hasta ahora
      ii) Elimina el nodo
      iii) Ir a 4
   f) Si $Z_{LB} > Z_{UB}$ entonces
      i) Elimina el nodo
      ii) Ir a 4
   g) Si $Z_{LB} > Z*_{LB}$ entonces actualiza $Z*_{LB}$.
   h) Si el tiempo/iteraciones están en su límite entonces
      i) Actualiza las penalidades/peso
      ii) Ir a 3(d).
   i) Almacena los datos del nodo

4. Elegir nodo

   Si existen subproblemas activos elegir uno de la lista.

   Sino parar.

5. Recuperar los datos del nodo

   a) Recuperar el vector de penalidades y pesos.
   b) Recuperar los arcos fijos y eliminados de la solución.

6. Escoger un arco para ramificar.

   De todos los posibles arcos escoger (x, y).

   Ir a 2.

En este capítulo solamente presentamos dos ejemplos de problemas que se presentan en otras áreas del conocimiento de dos de los métodos estudiados, pero la cantidad de éstos es muy grande.

[3]ecuación (28) de [6]

Conclusiones

En los ejemplos presentados y resueltos de manera detallada en cada uno de los capítulos, se hizo uso de subproblemas lineales, tal cual lo exigen los algoritmos, los cuales fueron resueltos con la ayuda del programa LINDO. Las aproximaciones para encontrar las $\alpha-$extensiones se hicieron a través del programa MATLAB.

Es importante mencionar que los métodos deterministas que estudiamos tienen un gran campo de aplicación, como pudimos ver en el capítulo 5 sólo presentamos dos ejemplos, pero en [5] es posible encontrar diferentes aplicaciones que sobre todo involucran procesos químicos, para resolver dichos procesos o mejorarlos hacen uso de los métodos de ramificación y acotamiento.

También vimos que los métodos de aproximación exterior son usados para resolver problemas de hidrogeología, como fue el caso del presentado en el capítulo 5, cabe mencionar que como consecuencia del trabajo de [9], tres años más tarde los mismo autores desarrollaron un algoritmo que combina el método de corte con el de aproximación exterior para resolver el problema de extracción de agua pero considerando que la región de soluciones factibles es convexa, no precisamente lineal.

Cabe hacer énfasis que durante este trabajo solamente estudiamos el caso en que la función es cóncava y las restricciones son un sistema de igualdades o desigualdades lineales, pero dentro del área de optimización global encontraremos también casos en que las restricciones son de tipo no lineales que por supuesto su aplicación en la vida real es basta y merece especial atención su estudio. Así mismo, hay que tomar en cuenta que este trabajo representa un primer acercamiento para aquellas personas interesadas en el comenzar el estudio de la optimización global determinista. Cabe resaltar que dentro de la optimización global hay diversas ramas de interés que reciben su nombre dependiendo del tipo de algoritmos que los caracteriza, una es la que ya hemos visto y se le conoce como determinista, pero hay otras como la estocástica y la heurística.

Otro aspecto importante a considerar en la implementación de los algoritmos en las diferentes áreas del conocimiento es la necesidad de estar al menos un poco familiarizado con el área en cuestión, por ejemplo; para el problema del acuífero existen ciertas restricciones que son modeladas a través de un sistema de ecuaciones diferenciales parciales, las cuales para poder transformarlas a expresiones lineales, usaron el método de elementos finitos, esto sin mencionar los modelos matemáticos usados para representar los fenómenos físicos. Otro ejemplo lo tenemos en el discutido en el capitulo 5, en donde es central saber un poco de teoría de redes y al menos una noción de programación dinámica.

Se pudo apreciar la importancia de la diversidad de aplicaciones que pueden tener estos métodos; por ejemplo en el método de ramificación y acotamiento implementado en los problemas de flujo en redes con costos cóncavos, la esencia del algoritmo permanece, no siendo así su implementación, los cuál hace que estos métodos sean herramientas poderosas y adaptables que nos permitirán resolver diferentes tipos de problemas.

De este modo podemos concluir que el campo de aplicación para estos métodos y su importancia en la minimización cóncava es grande, no sólo en su forma directa, sino que también se ven involucrados en otras áreas del conocimiento o bien en otras ramas de la optimización. Representando así los cimientos sobre los cuales es posible seguir desarrollando nuevos algoritmos y/o presentar mejoras a los ya existentes.

A

Conceptos básicos

Algunos de las definiciones de conjuntos que necesitaremos son: conjunto convexo, cerrado, abierto, compacto. El tipo de funciones con las que trabajaremos principalmente son funciones cóncavas, así que también daremos algunas de sus propiedades que serán de gran importancia a lo largo del presente trabajo.

A.1. Conjuntos Convexos

Definición 18. *(Línea) Sean los vectores x_1, $x_2 \in \mathbb{R}^n$. Sea la línea que pasa a través de los puntos x_1 y x_2 definida como el conjunto*

$$\{x|x = (1 - \lambda)x_1 + \lambda x_2, \quad \lambda \in \mathbb{R}^n\} \tag{A.1}$$

Definición 19. *(Segmento de Línea Cerrado) Sean los vectores x_1, $x_2 \in \mathbb{R}^n$. El segmento de línea cerrado que pasa a través de los puntos x_1 y x_2 definida como el conjunto*

$$\{x|x = (1 - \lambda)x_1 + \lambda x_2, \quad 0 \leq \lambda \leq 1\} \tag{A.2}$$

Los segmentos de línea abiertos, cerrados-abiertos y abiertos-cerrados pueden ser definidos de manera análoga modificando las desigualdades de λ.

Definición 20. *(Semiespacio) Sea el vector $c \in \mathbb{R}^n$, $c \neq 0$ y el escalar $z \in \mathbb{R}$. El semiespacio abierto en $\mathbb{R}^n$ es definido como el conjunto:*

$$\{x|c^t x < z, \quad x\mathbb{R}^n\} \tag{A.3}$$

El semiespacio cerrado en $\mathbb{R}^n$ es definido como el conjunto:

$$\{x|c^t x \leq z, \quad x\mathbb{R}^n\} \tag{A.4}$$

Definición 21. *(Hiperplano) El hiperplano en $\mathbb{R}^n$ es definido como el conjunto*

$$\{x|c^t x \leq z, \quad x \in \mathbb{R}^n\} \tag{A.5}$$

Definición 22. *(Hiperplano soporte) Sea S un conjunto no vacío en $\mathbb{R}^n$ y sea $\bar{x} \in \partial S$. Un hiperplano $H = \{x|c^t(x-\bar{x}) = 0\}$ es llamado un hiperplano soporte de S en $\bar{x}$ si $c^t(x-\bar{x}) \geq 0$ para todo $x \in S$, o bien, si $c^t(x-\bar{x}) \leq 0$ para todo $x \in S$.*

Definición 23. *(Politopo y Poliedro) La intersección de un número finito de semiespacios cerrados en $\mathbb{R}^n$ es definido como un politopo. Un politopo acotado es llamado un poliedro.*

Definición 24. *(Conjunto Convexo) Un conjunto $S \in \mathbb{R}^n$ es llamado convexo si el segmento de línea cerrado une a cualesquiera dos puntos x_1 y x_2 del conjunto S, es decir, $(1-\lambda)x_1 + \lambda x_2$ pertenece al conjunto S para cualquier λ con $0 \leq \lambda \leq 1$*

Algunos ejemplos de conjunto convexos son:

1. La línea
2. Semiespacios abiertos y cerrados
3. Poliedros, Politopos
4. Todos los puntos dentro o sobre el círculo
5. Todos los puntos dentro o sobre un polígono

Lema 1. *(Propiedades sobre los conjuntos convexos) Sea S_1 y S_2 conjuntos convexos en $\mathbb{R}^n$. Entonces*

1. *La intersección $S_1 \cap S_2$ es un conjunto convexo*
2. *La suma $S_1 + S_2$ de dos conjuntos convexos es un conjunto convexo*
3. *El producto θS_1 de un número real θ y el conjunto S_1 es un conjunto convexo*
4. *La diferencia $S_1 - S_2$ de dos conjuntos convexos es un conjunto convexo*

Definición 25. *(Combinación Convexa) Sea $\{x_1, ..., x_r\}$ un conjunto finito de puntos en $\mathbb{R}^n$. Una combinación convexa de este conjunto es un punto de la forma:*

$$\lambda_1 x_1 + \cdots + \lambda_r x_r \tag{A.6}$$
$$\lambda_1 + \cdots + \lambda_r = 1 \tag{A.7}$$
$$\lambda_1, \cdots, \lambda_r \geq 0 \tag{A.8}$$

Definición 26. *(Envolvente Convexa) Sea S un conjunto (convexo o no convexo) en $\mathbb{R}^n$. La envolvente convexa $H(S)$ de S es definida como la intersección de todos los conjuntos convexos en $\mathbb{R}^n$ que contiene a S como un subconjunto.*

Teorema 10. *La envolvente convexa $H(S)$ de S es definida como el conjunto de todas las combinaciones de S. Entonces $x \in H(S)$ si y sólo si x puede ser representado como:*

$$x = \Sigma_{i=1}^{r} \lambda_i x_i \tag{A.9}$$
$$\Sigma_{i=1}^{r} \lambda_i = 1 \tag{A.10}$$
$$\lambda_i \geq 0, \quad i = 1, ..., r \tag{A.11}$$
$$x_i \in S, \quad i = 1, ..., r \tag{A.12}$$

donde r es un entero positivo.

Observación. Cualquier punto x de la envolvente convexa de un conjunto S en $\mathbb{R}^n$ puede ser escrito como una combinación convexa de cuando mucho $n+1$ puntos de S como lo establece el siguiente teorema

Teorema 11. *(Caratheodory) Sea S un conjunto (convexo o no convexo) en $\mathbb{R}^n$. Si $x \in H(S)$ entonces puede ser representado como*

$$x = \Sigma_{i=1}^{n+1} \lambda_i x_i \tag{A.13}$$
$$\Sigma_{i=1}^{n+1} \lambda_i = 1 \tag{A.14}$$
$$\lambda_i \geq 0, \quad i = 1, ..., n+1 \tag{A.15}$$
$$x_i \in S, \quad i = 1, ..., n+1 \tag{A.16}$$

Decimos que x es un punto cerradura de un subconjunto de X de $\mathbb{R}^n$ si existe una sucesión $\{x_k\} \subset X$ tal que converge a x. La cerradura de X, se denota como $cl(X)$, es el conjunto de todos los puntos cerradura de X.

Definición 27. *(Cerrado) Un subconjunto X de $\mathbb{R}^n$ es llamado cerrado si es igual a su cerradura.*

Definición 28. *(Abierto) Un subconjunto X de $\mathbb{R}^n$ es llamado abierto si su complemento $\{x|x \notin X\}$ es cerrado.*

Definición 29. *(Acotado) Un subconjunto X de $\mathbb{R}^n$ es llamado acotado si existe un escalar c tal que $||x|| \leq c$ para todo $x \in X$.*

Definición 30. *(Compacto) Un subconjunto X de $\mathbb{R}^n$ es llamado compacto si es cerrado y acotado.*

A.2. Funciones cóncavas y algunas de sus propiedades

Definición 31. *(Función Cóncava) Sea S un subconjunto convexo de $\mathbb{R}^n$ y sea $f(x)$ una función real sobre S. La función $f(x)$ se dice que es cóncava si para cualesquiera x_1, $x_2 \in S$ y con $0 \leq \lambda \leq 1$ tenemos:*

$$f((1-\lambda)x_1 + \lambda x_2) \geq (1-\lambda)f(x_1) + \lambda f(x_2) \tag{A.17}$$

Definición 32. *(Función Estrictamente Cóncava) Sea S un subconjunto convexo de $\mathbb{R}^n$ y sea $f(x)$ una función real sobre S. La función $f(x)$ se dice que es estrictamente cóncava si para cualesquiera x_1, $x_2 \in S$ y con $0 < \lambda < 1$ tenemos:*

$$f((1-\lambda)x_1 + \lambda x_2) > (1-\lambda)f(x_1) + \lambda f(x_2) \tag{A.18}$$

Definición 33. *(Función Convexa) Sea S un subconjunto convexo de $\mathbb{R}^n$ y sea $f(x)$ una función real sobre S. La función $f(x)$ se dice que es convexa si para cualesquiera x_1, $x_2 \in S$ y con $0 \leq \lambda \leq 1$ tenemos:*

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \tag{A.19}$$

Definición 34. *(Función Estrictamente Convexa) Sea S un subconjunto convexo de $\mathbb{R}^n$ y sea $f(x)$ una función real sobre S. La función $f(x)$ se dice que es estrictamente convexa si para cualesquiera x_1, $x_2 \in S$ y con $0 < \lambda < 1$ tenemos:*

$$f((1-\lambda)x_1 + \lambda x_2) < (1-\lambda)f(x_1) + \lambda f(x_2) \tag{A.20}$$

Observación. La función $f(x)$ es convexa sobre S si y sólo si $-f(x)$ es cóncava. Así, los resultados obtenidos para las funciones cóncavas pueden ser modificados en resultados de funciones convexas multiplicando por -1 y viceversa.

Los siguientes son algunas combinaciones de funciones cóncava que dan lugar a nuevas funciones cóncavas :

1. Sean $f_1(x), f_2(x), ..., f_n(x)$ funciones cóncava sobre un subconjunto convexo S de $\mathbb{R}^n$. Entonces, su suma
$$f_1(x) + f_2(x) + \cdots + f_n(x) \tag{A.21}$$
es cóncava.

2. Sea $f(x)$ cóncava (estrictamente cóncava) sobre un subconjunto S de $\mathbb{R}^n$ y λ un número positivo. Entonces, $\lambda f(x)$ es cóncava (estrictamente cóncava).

3. Sean $f_1(x), f_2(x), ..., f_n(x)$ funciones cóncavas y acotadas inferiormente sobre un subconjunto convexo S de $\mathbb{R}^n$. Entonces, la función ínfimo
$$f(x) = \min\{f_1(x), f_2(x), ..., f_n(x)\} \tag{A.22}$$
es una función cóncava sobre S.

4. Sean $f_1(x), f_2(x), ..., f_n(x)$ funciones cóncavas sobre un subconjunto convexo S de $\mathbb{R}^n$. Entonces, la función media geométrica
$$f(x) = (\Pi_i f_i(x))^{1/n} \tag{A.23}$$
es una función cóncava sobre S.

Continuidad y semicontinuidad

Definición 35. *(Función Continua) Sea S un subconjunto de $\mathbb{R}^n$, $x^0 \in S$ y $f(x)$una función real definida sobre S. $f(x)$ es continua sobre x^0 si alguna de las condiciones siguientes que son equivalentes se satisface:*

1. *Para cada $\epsilon_1 > 0$ existe un $\epsilon_2 > 0$: $||x - x^0|| < \epsilon_2$, $x \in S$ implica que*
$$-\epsilon_1 < f(x) - f(x^0) < \epsilon_1 \tag{A.24}$$

2. *Para cada sucesión $x^1, x^2, ..., x^n$ $(x \in S)$ convergente a x^0*
$$\lim_{n\to\infty} f(x^n) = f(\lim_{n\to\infty} x^n) = f(x^0) \tag{A.25}$$

$f(x)$ es continua sobre S si es continua sobre cada punto $x^0 \in S$.

Definición 36. *(Función Semi-continua inferiormente) $f(x)$ es semi-continua inferiormente en x^0 si alguna de las condiciones siguientes que son equivalentes se satisface:*

1. *Para cada $\epsilon_1 > 0$ existe un $\epsilon_2 > 0$: $||x - x^0|| < \epsilon_2$, $x \in S$ implica que*
$$-\epsilon_1 < f(x) - f(x^0) \tag{A.26}$$

2. Para cada sucesión $x^1, x^2, ..., x^n$ $(x \in S)$ convergente a x^0

$$\lim_{n\to\infty} inf f(x^n) \geq f(\lim_{n\to\infty} x^n) = f(x^0) \tag{A.27}$$

donde $\lim_{n\to\infty} inf f(x^n)$ es el ínfimo del punto límite de la sucesión $f(x^1), f(x^2), ..., f(x^n)$.

$f(x)$ es semi-continua inferiormente sobre S si es semi-continua inferiormente para cada $x^0 \in S$.

Teorema 12. *Sea S un conjunto no vacío convexo en $\mathbb{R}^n$ y $f(x)$ una función cóncava. Entonces $f(x)$ es continua en el interior de S.*

Observación. Es posible que las funciones cóncava y convexas no sean continuas en cualquier lugar, pero los puntos de discontinuidad tienen que estar en la frontera de S.

Teorema 13. *Sea $f_i(x)$ una familia de funciones semi-continuas inferiormente sobre S. Entonces,*

(i) Su mínima cota superior $f(x) = \sup_i f_i(x)$ es semi-continua inferiormente sobre S.

(ii) Si la familia $f_i(x)$ es finita, su máxima cota inferior $f'(x) = \inf_i f_i(x)$ es semi-continua inferiormente sobre S.

Teorema 14 (Teorema de Weierstrass)**.** *Si una función f es continua en un intervalo cerrado $[a, b]$ entonces hay al menos dos puntos $x1, x2$ pertenecientes a $[a, b]$ donde f alcanza valores extremos absolutos, es decir $f(x_1) \leq f(x) \leq f(x_2)$, para cualquier $x \in [a, b]$.*

Diferenciabilidad

Propiedad 5. *Suponga que $S = R^n$ o bien A es un subconjunto convexo abierto de R^n, si $x \in S$ entonces f es diferenciable en x, o x es el límite de una sucesión de puntos en S con la propiedad que f es diferenciable en cada punto de la sucesión o ambas. Más aún, el conjunto de puntos en S donde f no es diferenciable tiene medida cero.*

Propiedad 6. *Suponga que f es diferenciable en cada punto $x \in S$ entonces para cualquier $y \in S$ se cumple*

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle$$

donde $\langle \cdot, \cdot \rangle$ es el producto interior.

Propiedad 7. *Si $S = R^n$ o S es un subconjunto convexo abierto de R^n entonces al menos un subgradiente de f existe en cada punto $x \in S$. Más aún, para cada $x \in A$ si f es diferenciable en x entonces $\nabla f(x)$ es el único subgradiente de f en x; de otro modo, el conjunto de subgradientes de f en x contiene más de un elemento.*

A.3. Envolventes Convexas

Definición 37. *(Envolvente convexa) Sea $f(x)$ una función semi-continua inferiormente definida sobre un conjunto no vacío convexo $S \in \mathbb{R}^n$. Entonces, la envolvente convexa $f(x)$ sobre $S \subset \mathbb{R}^n$ es una función $\phi(x)$ que satisface:*

(i) $\phi(x)$ *es convexa sobre* S.

(ii) $\phi(x) \leq f(x)$ *para todo* $x \in S$.

(iii) *Si $h(x)$ es una función convexa definida sobre S tal que $h(x) \leq f(x)$ para todo $x \in S$, entonces, $h(x) \leq \phi(x)$ para todo $x \in S$.*

Observaciones. La condición (i) establece la convexidad de ϕ, mientras que la condición (ii) estable que $\phi(x)$ es la estimación inferior de $f(x)$ sobre todo el dominio. Finalmente la condición (iii) afirma que no existe otra función convexa de estimación inferior que se ajuste más que la envolvente convexa $\phi(x)$.

Teorema 15. *Sea $f(x)$ una función semi-continua inferiormente definida sobre un conjunto convexo compacto $S \in \mathbb{R}^n$, y $\phi(x)$ la envolvente convexa de $f(x)$ sobre S. Entonces,*

(i) $\text{mín}_{x\in S} f(x) = \text{mín}_{x\in S} \phi(x) = f*$.

(ii) $\{y \in S : f(y) = f*\} \subseteq \{y \in S : \phi(y) = f*\}$.

Observación. Este teorema establece que para cada problema de optimización no convexa que tiene una región factible convexa, es posible asociarle un problema de optimización convexa, el cual tiene la misma solución óptima que el problema de optimización no convexa.

La envolvente convexa de una función cóncava sobre un politopo

La envolvente convexa de una función cóncava sobre un politopo se basa en el siguiente teorema.

Teorema 16. *Sea $f(x)$ una función cóncava definida sobre un politopo P y sea $v_1, v_2, ..., v_m$ los vértices del politopo. Entonces, la envolvente convexa $\phi(x)$ de la función cóncava $f(x)$ sobre el politopo P puede ser escrito de la siguiente manera:*

$$\begin{aligned} \phi(x) = \min_{\lambda} \Sigma_{i=1}^{m} \lambda_i f(v_i) \\ s.a \Sigma_{i=1}^{m} \lambda_i v_i = x \\ \Sigma_{i=1}^{m} \lambda_i = 1 \\ \lambda = (\lambda_1, ..., \lambda_m) \geq 0 \end{aligned}$$

La envolvente convexa de una función cóncava sobre un simplex

La envolvente convexa de una función cóncava sobre un simplex es una función afín que puede ser determinada de manera única por un sistema de ecuaciones lineales, como el siguiente teorema lo establece.

Teorema 17. *Sea $f(x)$ una función cóncava definida sobre un simplex S y sea $v_1, v_2, ..., v_m$ los vértices del simplex. Entonces, la envolvente convexa $\phi(x)$ de la función cóncava $f(x)$ sobre el simplex S es la función afín:*

$$\phi(x) = c^T x + b$$

donde el vectos de coeficientes c, b son determinados de manera única por el siguiente sistema ecuaciones lineales

$$f(v_i) = c^T v_i + b \quad i = 1, ..., n.$$

A.4. Mínimo local, global y $\epsilon-$global

Consideremos el problema de minimizar $f(x)$ sujeto a $x \in S$ donde el conjunto S se define como:

$$S = \{x \in X | h(x) = 0, g(x) \leq 0\}$$

donde X es conjunto convexo cerrado.

Definición 38. *(Solución Factible) Un punto $x \in S$ es una solución factible a este problema.*

Definición 39. *(Mínimo local) Supongamos que $x* \in S$ y que existe un $\epsilon > 0$ tal que*

$$f(x) \geq f(x*) \ \forall x \in S : ||x - x*|| < \epsilon$$

entonces, $x*$ es un mínimo local.

Definición 40. *(Mínimo global) Supongamos que $x* \in S$ y que*

$$f(x) \geq f(x*) \ \forall x \in S$$

entonces, $x*$ es un mínimo global.

Definición 41. *(Mínimo global único) Supongamos que $x* \in S$ y que existe un $\epsilon > 0$ tal que*

$$f(x) > f(x*) \ \forall x \in S : ||x - x*|| < \epsilon$$

entonces, $x*$ es un mínimo global único.

En varios aplicaciones prácticas, calcular el mínimo global exacto puede ser una tarea difícil computacionalmente hablando o no necesaria. Con mayor frecuencia se tiene más interés en identificar el mínimo global con una tolerancia $\epsilon \geq 0$ dada.

Definición 42. *(Mínimo $\epsilon-$global) Supongamos que $x* \in S$ es una solución factible, $\epsilon \leq 0$ es una pequeña tolerancia dada y*

$$f(x) \geq f(x*) - \epsilon \ \forall x \in S : ||x - x*|| < \epsilon$$

entonces, $x*$ es un mínimo $\epsilon-$global.

Bibliografía

[1] Abbas, M., and Bellahcene, F. Cutting plane method for multiobjective. *European Journal of Operational Research 168* (2006), 967–984.

[2] Bazaraa, M., Sherali, H., and C., S. *Nonlinear Programming; Theory and Algorithms*, third ed. John Wiley & Sons, 2006.

[3] Benson, H. P. Concave minimization: Theory, applications and algorithms. 43–148.

[4] Bergamini, M. L., Grossmann, I., Scenna, N., and Aguirre, P. An improved piecewise outer-approximation algorithm for the global optimization of minlp models involving concave and bilinear terms. *Computers and Chemical Engineering 32* (2008), 477–493.

[5] Floudas, C. A. *Deterministic Global Optimization: Theory, Methods and Applications*. Kluwer Academic Publishers, 2000.

[6] Fontes, D. B., Hadjiconstantinou, E., and Christofides, N. A branch and bound algorithm for concave network flow problems. *Journal of Global Optimization 34* (2006), 127–155.

[7] Horst, R., Pardalos, P. M., and Thoai, N. V. *Introduction to Global Optimization*, segunda edición ed. Kluwer Academic Publisher, 2000.

[8] Horst, R., and Tuy, H. *Global Optimization: Deterministic Approaches*, tercera edición ed. Springer, 1995.

[9] Karatzas, G. P., and Pinder, G. F. Groundwater manangement using numerical simulation and the oter approsimation method for global optimization. *Water Resources Research 29* (1993), 3371–3378.

[10] Lawphongpanich, S. A nontangential cutting plane algorithm. *Operations Research Letters 30* (2002), 49–56.

[11] Locatelli, M., and Thoai, H. V. Finite exact branch-and-bound algorithms for concave minimization over polytopes. *Journal of Global Optimization 18* (2000), 107–128.

[12] O.Barrientos, R. Correa, P. R. A. V. A branch and bound method for solving integer separable concave problems. *Computational Optimization and Applications 26* (2003), 155–171.

[13] P.M.Pardalos, J. R. Methods for global concave minimization: A bibliographic survey. *Society of Industrial and Applied Mathematics 28* (1986), 367–379.

[14] Porembski, M. Finitely convergent cutting planes for concave minimization. *Journal of Global Optimization 20* (2001), 113–136.

[15] Porembski, M. Cone adaptation strategies for a finite and exact cutting plane algorithm for concave minimization. *Journal of Global Optimization 24* (2002), 89–107.

[16] Spiliotopoulos, A. A., Karatzas, G. P., and Pinder, G. F. A multiperiod approach to the solution of groundwater management problems using an outer approximation method. *European Journal of Operational Research 157* (2004), 514–525.

[17] Still, C., and Westerlund, T. A sequential cutting plane algorithm for solving convex nlp problems. *European Journal of Operational Research 173* (2006), 444–464.