UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMÁTICAS
Departamento de Estadística e Investigación Operativa

TESIS DOCTORAL

Optimalidad en la frontera de las restricciones en Programación Polinómica

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Juan Crisóstomo Almaraz Simón

DIRECTOR:

Juan Béjar Alamo

Madrid, 2015

UNIVERSIDAD COMPLUTENSE DE MADRID
Facultad de Ciencias Matemáticas
Departamento de Estadística e Investigación Operativa

OPTIMALIDAD EN LA FRONTERA DE LAS RESTRICCIONES EN PROGRAMACION POLINOMICA

Juan Crisostomo Almaraz Simón

Madrid, 1989

Colección Tesis Doctorales. N.º 31/89

Edita e imprime la Editorial de la Universidad
Complutense de Madrid. Servicio de Reprografía
Escuela de Estomatología. Ciudad Universitaria
Madrid, 1989
Ricoh 3700
Depósito Legal: M-3434-1989

El presente trabajo ha sido realizado bajo la dirección del Profesor Doctor Don Juan Béjar Alamo, a quien quiero expresar mi agradecimiento más entrañable por su inestimable ayuda, estímulo y bondadosa comprensión.

También quiero manifestar mi agradecimiento a los integrantes del Departamento de Estadística e Investigación Operativa que de una u otra forma me han prestado su apoyo.

Juan Crisóstomo Almaraz Simón

I N D I C E

I N T R O D U C C I O N

El objeto inicial de este trabajo era realizar un estudio sobre la Programación Polinómica en general, tema poco estudiado en la literatura de la Programación Matemática, ya que si bien algunos aspectos parciales de la misma cuales son la Programación Lineal, la Cuadrática, la Polinómica Entera y la Programación Polinómica 0-1 han tenido un extenso y profundo tratamiento, no así, en cambio, la Programación Polinómica en general.

Al profundizar en el estudio de la Programación no Lineal, observamos que existían algunas parcelas de la Programación Matemática que habían sido tratadas muy de pasada en la literatura existente. Fruto de esta observación y del tema propuesto por el Director del trabajo, es la Memoria que ahora presentamos.

La Memoria consta de tres Capítulos y un Apéndice.

El Capítulo 1, está dedicado a la Programación no Lineal. En el mismo, se aportan resultados originales acerca de tres puntos:

1°) Optimo en la frontera de las restricciones.

2°) Unicidad del óptimo.

3°) Equivalencia de algunos teoremas al de Kuhn y Tucker en Programas Cuadráticos.

Con respecto al punto primero, se estudia:

a) El problema de Programación con función objetivo convexa y conjunto de restricciones cerrado, estableciéndose dos teoremas: uno que es una condición necesaria y suficiente para que el óptimo esté en la frontera de las restricciones y el otro que es una condición suficiente.

b) La Programación Convexa, estableciéndose dos teoremas análogos a los del apartado a). De lo obtenido para este tipo de Programación, se derivan los resultados correspondientes para programas cuadráticos (cuando los términos de segundo grado de la función objetivo constituyen una forma cuadrática definida positiva o negativa).

En relación con el segundo punto de este Capítulo, se establece un nuevo teorema de unicidad del óptimo en la frontera en Programación Convexa, en el que la función objetivo se le exige únicamente que sea convexa (o cóncava) en contraposición al teorema existente (teorema A.3. del Apéndice) en el que se formula la hipótesis de convexidad (o concavidad) estricta para la función objetivo.

Como una consecuencia de lo establecido en los puntos primero y segundo y en el teorema A.3. del Apéndice, se formulan, para la Programación Convexa, dos teoremas de unicidad y optimalidad en la frontera de las restricciones.

Finaliza el primer Capítulo, estableciéndo teoremas equivalentes al de Kuhn y Tucker para programas cuadráticos, cuando el óptimo está en la frontera de las restricciones. En estos teoremas, se ponen de manifiesto las dos condiciones que se han de verificar en el punto de óptimo:

1ª) Que tal punto de óptimo es la intersección única de la hipersuperficie de nivel de la función objetivo con la frontera de las restricciones.

2ª) La relación existente entre el vector gradiente de las restricciones saturadas linealmente indpendientes con las restantes matrices del programa.

La finalidad del Capítulo 2 es el estudio de la Programación Polinómica. Las ideas originales de este Capítulo se refieren a:

1º) Estudio sobre el óptimo y la optimalidad en la frontera de las restricciones en algunos programas polinómicos.

2º) Algoritmo para la obtención del óptimo en Programación Polinómica con restricciones lineales.

Este Capítulo, se inicia indicando las condiciones bajo las cuales se pueden particularizar a la Programación Polinómica los resultados obtenidos en el Capítulo 1.

Seguidamente, y en relación con el primer punto, se estudia la optimalidad en programas polinómicos con función objetivo de coeficientes positivos y conjunto de restricciones compacto en R_+^n , particularizándose los resultados obtenidos al caso en el que las restricciones son lineales y forman un poliedro convexo. El resultado que establece la optimalidad en la frontera de las restricciones se hace de dos formas diferentes, aunque es más preciso el primer método pues nos indica qué condición adicional ha de verificar la región de la frontera en la que se ubica el óptimo. Se muestra, mediante un contraejemplo, que tales resultados sobre Programación Polinómica Positiva no pueden extenderse a la Programación Posinomial.

Con respecto al segundo punto, se propone un método para la obtención del óptimo cuando el conjunto de las restricciones es lineal y la función objetivo polinómica es convexa en dicho conjunto de restricciones. El citado método extiende y desarrolla una idea formulada por Beale, E.M.L. [7] para el caso cuadrático. El algoritmo establecido por nosotros generaliza la metodología del simplex con modificaciones esenciales que pasamos a exponer. Como se sabe, el algoritmo del simplex, en un problema de mínimo, consiste en determinar la solución óptima (que está en al menos un vértice del politopo de las restricciones) de un programa lineal; partiendo de un vértice se pasa, en cada iteración, a otro adyacente tomando como nueva variable básica la de coste reducido negativo más pequeño.

Reiterando el proceso, se llega a un vértice que es solución óptima, el cual se alcanza cuando los costes reducidos de las variables libres son todos positivos. En Programación Polinómica con restricciones lineales y función objetivo polinómica convexa en tal conjunto de restricciones, la solución mínima, en general, no se encuentra en un vértice. El algoritmo que formulamos permite, para estos programas, avanzar por vértices adyacentes mientras sea posible. Es decir, mientras que en cada iteración se pase de una solución básica factible a otra solución del mismo tipo. La dirección de avance, en cada iteración, viene dada por la variable libre que verifica que la derivada de la función objetivo respecto de la misma, en el vértice de partida, es la más negativa. Cuando el punto al que se llega no es vértice, se transforma en tal mediante la ampliación del programa. La ampliación mencionada, se consigue añadiendo una nueva ecuación lineal con una nueva variable sin restricción en el signo, logrando así que el punto de llegada se transforme en vértice del nuevo programa. En este programa, se reitera la metodología anteriormente citada.

Se estudian las propiedades inherentes a tal método, viéndose las relaciones existentes entre los sucesivos programas con respecto a: 1) Las soluciones factibles, soluciones óptimas, soluciones básicas factibles y soluciones básicas factibles no degeneradas, 2) Acotación de las soluciones factibles, 3) Caracterización del óptimo. Asimismo, se analiza la convergencia del algoritmo.

Finaliza el Capítulo indicando en qué sentido se puede extender el método a programas polinómicos no convexos con restricciones lineales.

El Capítulo 3 está dedicado a la Programación Polinómica. Se enfoca la misma desde un nuevo punto de vista, transformándola en un problema equivalente de Programación Multiobjetivo, en el cual se lleva a cabo el estudio del problema original.

Las ventajas de este nuevo enfoque, debidas a las funciones multiobjetivo que se establecen y a la relación de preferencia definida, se refieren a la sencilla caracterización de los elementos eficientes y a la precisión con la que se da la ubicación de los mismos.

Una vez formulado el equivalente problema multiobjetivo, se estudian las propiedades de la relación de preferencia $\succ$ ($\succ$ "más preferido que"). Seguidamente, se ven las relaciones $\sim$ ($\sim$ "indiferente a") y $\succsim$ ($\succsim$ "más preferido o indiferente que") asociadas a la anterior. Posteriormente, se definen el conjunto eficiente y la estructura de dominación D relativos a la relación $\succ$. A continuación, se estudian las propiedades de la estructura D así como la existencia de soluciones eficientes. Se dan dos caracterizaciones del conjunto eficiente: la primera como transformado, mediante las funciones multiobjetivo, del conjunto de soluciones óptimas del programa polinómico inicial, y la segunda mediante un producto escalar en el que interviene el cono polar positivo estricto D^{so} de la estructura D. Debido a la sencilla expresión de D^{so}, se profundiza aún más sobre el conjunto de elementos eficientes y sobre su ubicación.

Finaliza este Capítulo, estableciéndose un teorema sobre la optimalidad en la frontera del conjunto factible del programa multiobjetivo formulado.

El estudio realizado en el Capítulo 3, se ha hecho para el problema de mínimo; un tratamiento análogo puede formularse para el caso de máximo.

En el Apéndice, se incluyen resultados de Programación Convexa que sirven de apoyo para la demostración de algunos teoremas de la Memoria, asimismo se adjunta la demostración original de un resultado topológico conocido que es necesario para la demostración de un teorema esencial del Capítulo 1.

Madrid, Enero de mil novecientos ochenta y ocho.

C A P I T U L O 1

OPTIMALIDAD EN LA FRONTERA DE LAS RES TRICCIONES EN PROGRAMACION NO LINEAL

1.1. INTRODUCCION

Se presentan en este Capítulo conclusiones nuevas acerca de algunos temas de la Programación no Lineal. Dichos resultados se refieren a:

1) Optimalidad en la frontera de las restricciones
2) Unicidad del óptimo
3) Teoremas equivalentes al de Kuhn y Tucker en ciertos programas cuadráticos

El primer punto se trata en los epígrafes 1.2. y 1.3.

En el apartado 1.2. se consideran Programas no Lineales, bajo las hipótesis de función objetivo convexa o cóncava y conjunto de restricciones cerrado. Se establecen dos teoremas: el teorema 1.2.1. que da una condición necesaria y suficiente y el teorema 1.2.2. que formula una condición suficiente de óptimo en la frontera para este tipo de programas.

En la sección 1.3., se estudian programas convexos (es decir, cuando la función objetivo es convexa o cóncava y las restricciones son convexas). Se particularizan al caso convexo los teoremas formulados en el epígrafe 1.2. De la condición suficiente dada en el teorema 1.3.2., se derivan en cascada una serie de teoremas relativos a programas convexos cuando las funciones objetivo son cuadráticas. Se comprueba entonces, que la optimalidad en la frontera en programas cuadráticos, con restricciones lineales, no son sino casos particulares del resultado general establecido en el teorema 1.2.1. También se puede derivar, del citado teorema, la optimalidad en la frontera para el caso de Programación Lineal. Por tanto, de los resultados aquí establecidos más los ya conocidos de Programación Convexa en el caso de $\left\{\begin{matrix}\text{Maximizar}\\ \text{Minimizar}\end{matrix}\right\}$ una función objetivo $\left\{\begin{matrix}\text{convexa}\\ \text{cóncava}\end{matrix}\right\}$ sujeta a restricciones convexas, y en particular a restricciones lineales, se puede formular una teoría unificada de óptimo en la frontera de las restricciones en Programación Convexa.

Con relación al segundo punto, se establece un nuevo teorema de unicidad en programas convexos (apartado 1.4.). Como una consecuencia de lo establecido por separado para la optimalidad en la frontera y unicidad del óptimo en las secciones 1.3. y 1.4., así como del teorema A.3. de unicidad (Apéndice), se formulan los teoremas del epígrafe 1.5. en los que se dan condiciones suficientes de unicidad y optimalidad en la frontera de las restricciones.

En lo referente al tercer punto, se dan teoremas equivalentes al de Kuhn y Tucker para programas cuadráticos de mínimo con restricciones lineales, cuando el óptimo está en la frontera (si la función objetivo es forma cuadrática definida positiva y las restricciones lineales son del tipo $Ax \geqslant b$, $x \geqslant 0$, en el teorema 1.3.5. hemos demostrado que el óptimo está en la frontera; pero si la función objetivo es cuadrática con términos lineales y con las restricciones lineales anteriores, la optimalidad no necesariamente se produce en la frontera). En estos teoremas, se formulan las condiciones necesarias y suficientes de óptimo, que nos indican:

1°) Que en el punto x^o de mínimo la hipersuperficie de nivel de la función objetivo (relativa a x^o) y la frontera de las restricciones han de tener como intersección única dicho punto x^o.

2°) La relación que debe existir entre el vector gradiente en x^o de las restricciones saturadas linealmente independientes con las restantes matrices del programa.

1.2. OPTIMALIDAD EN LA FRONTERA EN ALGUNOS PROGRAMAS NO LINEALES

En este apartado, se formulan dos teoremas de optimalidad en la frontera de las restricciones para programas con función objetivo convexa o cóncava y conjunto de restricciones cerrado. El primer teorema da una condición necesaria y suficiente y el segundo una condición suficiente.

Teorema 1.2.1.

Dado el programa:

$$(P_1)\begin{cases}\begin{Bmatrix}\text{Min } F(x)\\ \text{Max } F(x)\end{Bmatrix}\\ \text{bajo las restricciones:}\\ x \in Y \subset D \subset R^n\end{cases}$$

siendo $F(x)$ función $\begin{Bmatrix}\text{convexa}\\ \text{cóncava}\end{Bmatrix}$ en D, Y conjunto cerrado para la topología usual de R^n, $Int(Y) \neq \emptyset$, S_o y M_a, respectivamente, los conjuntos de soluciones óptimas de (P_1) y de $\begin{Bmatrix}\text{mínimo}\\ \text{máximo}\end{Bmatrix}$ absoluto de $F(x)$ en D. Entonces, $S_o \subset Fr(Y)$ si y sólo si:

$$\left[M_a \subset Int(Y)^c \cap D \quad \text{ó} \quad \exists\, x^* \in Y^c \cap D \text{ tal que} \begin{Bmatrix}F(x^*) < F(x^o)\\ F(x^*) > F(x^o)\end{Bmatrix},\ x^o \in S_o\right]$$

Demostración

$\Rightarrow$ Realizamos, en primer lugar, la demostración cuando $F(x)$ es convexa.

1) Supongamos, primeramente, que $F(x)$ posee mínimo absoluto en D.

Vamos a probar que $\forall x^1 \in M_a \Rightarrow x^1 \in (Int(Y))^c \cap D$

En efecto, si $\exists\, x^1 \in M_a$ tal que $x^1 \in Int(Y)$, pueden presentarse dos posibilidades:

a) $F(x^1) < F(x^o)$, $x^o \in S_o$, entonces como $x^1 \in Y$ se sigue que $x^o \notin S_o$, en contra de la hipótesis del teorema. Contradicción, que tiene su origen en suponer que $\exists\, x^1 \in M_a$ tal que

$x^1 \in \text{Int}(Y)$. Luego, si se presenta la posibilidad a), se concluye que $\forall x^1 \in M_a$ es $x^1 \in (\text{Int}(Y))^c \cap D$.

b) $F(x^1) = F(x^o)$, con $x^o \in S_o$, ello implicaría que $x^1 \in S_o$ y $x^1 \in \text{Int}(Y)$, contradicción con la hipótesis de que $S_o \subset \text{Fr}(Y)$, que nace de suponer que $\exists x^1 \in M_a$ tal que $x^1 \in \text{Int}(Y)$. Luego, si ocurre b), concluímos que $M_a \subset (\text{Int}(Y))^c \cap D$.

2) Supongamos, en segundo lugar, que $F(x)$ no tiene mínimo absoluto en D.

Probemos, para este supuesto, que:

$\exists x^* \in Y^c \cap D$ tal que $F(x^*) < F(x^o)$, con $x^o \in S_o$

En efecto, si $\forall x^* \in Y^c \cap D$ fuera $F(x^*) \geq F(x^o)$ ello implicaría que $x^o \in M_a$, en contra de lo supuesto de que $F(x)$ no tiene mínimo absoluto en D. Contradicción que se origina al suponer que $\forall x^* \in Y^c \cap D$ es $F(x^*) \geq F(x^o)$. Luego, si $F(x)$ no tiene mínimo absoluto en D, entonces $\exists x^* \in Y^c \cap D$ tal que $F(x^*) < F(x^o)$, con $x^o \in S_o$.

$\Longleftarrow$ Demostremos, ahora, la proposición recíproca.

1) Supongamos que $F(x)$ tiene mínimo absoluto en D y que $M_a \subset (\text{Int}(Y))^c \cap D$.

Vamos a probar que $S_o \subset \text{Fr}(Y)$.

En efecto, si $S_o \not\subset \text{Fr}(Y)$, ello equivale a que $\exists x^o \in S_o$ tal que $x^o \in \text{Int}(Y)$. Y entonces, pueden ocurrir dos posibilidades:

a) $F(x^1) < F(x^o) \qquad \forall x^1 \in M_a$

Como $x^o \in Int(Y)$ e Y es cerrado en R^n, se sigue que:

$$\left.\begin{array}{l} \exists x^2 \in Fr(Y) \quad \text{tal que:} \\ x^2 = \lambda x^o + (1-\lambda)\, x^1 \qquad 0 < \lambda < 1 \end{array}\right\} (*)$$

Luego,

$$F(x^2) = F(\lambda x^o + (1-\lambda)x^1) \underset{\substack{\uparrow \\ F \text{ convexa}}}{\leq} \lambda F(x^o) + (1-\lambda)F(x^1) \underset{\substack{\uparrow \\ F(x^1) < F(x^o)}}{<}$$

$$< \lambda F(x^o) + (1-\lambda)F(x^o) = F(x^o)$$

Es decir:

$$F(x^2) < F(x^o)$$

Luego,

$$\exists x^2 \in Fr(Y) \quad \text{tal que} \quad F(x^2) < F(x^o)$$

ello implica que $x^o \in S_o$. Contradicción, que nace de suponer que $\exists x^o \in S_o$ tal que $x^o \in Int(Y)$. Luego,

$\forall x^o \in S_o$ es $x^o \in Fr(Y)$. Es decir, $S_o \subset Fr(Y)$.

(*) Esta propiedad no es evidente. Aunque es un resultado topológico conocido se da una demostración original en el Apéndice (teorema A.6.).

b) $F(x^1) = F(x^o) \qquad \forall x^1 \in M_a$

Ello implica que $x^o \in M_a$. Pero, $x^o \in \text{Int}(Y)$, lo que contradice la hipótesis de ser $M_a \subset Y^c \cap D$. La contradicción, tiene su origen en suponer que $\exists x^o \in S_o$ tal que $x^o \in \text{Int}(Y)$.

Luego, si ocurre b), $\forall x^o \in S_o$ es $x^o \in \text{Fr}(Y)$.

Es decir, $S_o \subset \text{Fr}(Y)$.

2) Supongamos que $\exists x^* \in Y^c \cap D$ tal que $F(x^*) < F(x^o)$ con $x^o \in S_o$.

Vamos a probar que $S_o \subset \text{Fr}(Y)$.

En efecto,

si $S_o \not\subset \text{Fr}(Y) \iff \exists x^o \in S_o$ tal que $x^o \in \text{Int}(Y)$

De la condición anterior y de ser Y cerrado en R^n, se sigue que:

$$\exists x^2 \in \text{Fr}(Y) \text{ tal que:}$$

$$x^2 = \lambda x^o + (1-\lambda) x^* \qquad 0 < \lambda < 1$$

Luego,

$$F(x^2) = F(\lambda x^o + (1-\lambda)x^*) \underset{\substack{\uparrow \\ F \text{ convexa}}}{\leq} \lambda F(x^o) + (1-\lambda) F(x^*) \underset{\substack{\uparrow \\ F(x^*) < F(x^o)}}{<}$$

$$< \lambda F(x^o) + (1-\lambda) F(x^o) = F(x^o)$$

Es decir,

$$F(x^2) < F(x^o)$$

Luego, $\exists x^2 \in \text{Fr}(Y)$ tal que $F(x^2) < F(x^o)$

ello implica que $x^o \in S_o$. Contradicción, que nace de suponer que $\exists x^o \in S_o$ tal que $x^o \in Int(Y)$.

Luego,

$\forall x^o \in S_o$ es $x^o \in Fr(Y)$. Es decir, $S_o \subset Fr(Y)$.

Con lo que termina la demostración, cuando la función objetivo es convexa.

Consideremos, en segundo lugar, el caso en el que la función objetivo es cóncava.

Teniendo en cuenta la relación existente entre las funciones convexa y cóncava, el programa anterior (P_1) puede transformarse en el siguiente programa equivalente:

$$(P_1') \left\{ \begin{array}{l} -\text{Min} \left\{-F(x)\right\} \\ \text{bajo las restricciones:} \\ x \in Y \subset D \subset R^n \end{array} \right.$$

siendo $-F(x)$ función convexa en D, Y conjunto cerrado en R^n, $Int(Y) \neq \emptyset$, S_o y M_a los conjuntos de soluciones mínimas de $-F(x)$ en Y y de mínimo absoluto de $-F(x)$ en D, respectivamente. Entonces, $S_o \subset Fr(Y)$ si y sólo si:

$M_a \subset (Int(Y))^c \cap D$ ó $\exists x^* \in Y^c \cap D$

tal que: $-F(x^*) < -F(x^o)$ con $x^o \in S_o$.

Obviamente, el programa (P_1') verifica el teorema anterior cuando la función objetivo es convexa. Por tanto, en virtud de la equivalencia de los programas (P_1) y (P_1'), se cumple para (P_1) en el caso en el que $F(x)$ sea cóncava, la proposición formulada.

Queda, con ello, completamente demostrado el teorema.

Nota 1.2.1.

En el teorema 1.2.1., suponemos que $Int(Y) \neq \emptyset$, ya que si $Int(Y) = \emptyset$ es evidente que el óptimo, si existe, pertenece a $Fr(Y) = Y$.

Nota 1.2.2.

De las hipótesis del teorema 1.2.1., se sigue que

$F(x)$ no puede ser constante en Y, pues caso de serlo,
$\forall x^o \in S_o \implies (x^o \in \mathrm{Int}(Y)$ ó $x^o \in \mathrm{Fr}(Y))$.

Basándonos, en parte, en el teorema 1.2.1., vamos a establecer una condición suficiente de óptimo en la frontera de las restricciones.

Teorema 1.2.2.

Dado el programa:

$$(P_2)\left\{\begin{array}{l} \left\{\begin{array}{l}\text{Min } F(x)\\ \text{Max } F(x)\end{array}\right\} \\ \\ \text{bajo las restricciones:} \\ \\ x \in Y \subset D \subset R^n \end{array}\right.$$

siendo $F(x)$ función $\left\{\begin{array}{l}\text{convexa}\\ \text{cóncava}\end{array}\right\}$ en D, Y conjunto cerrado en R^n para la topología usual, $\mathrm{Int}(Y) \neq \emptyset$, S_o el conjunto de soluciones óptimas del programa (P_2). Si $\exists x^* \in Y^c \cap D$ tal que $\left\{\begin{array}{l}F(x^*) < F(x^o)\\ F(x^*) > F(x^o)\end{array}\right\}$, con $x^o \in S_o$, entonces $S_o \subset \mathrm{Fr}(Y)$.

Demostración

Consideraremos, en primer lugar, el caso en el que $F(x)$ es convexa.

Haremos la demostración por reducción al absurdo.

Supongamos que:

$$\exists x^o \in S_o \text{ tal que } x^o \notin Fr(Y)$$

Ello implica que, $x^o \in Int(Y)$

Entonces, por ser Y cerrado (y en virtud de la Proposición 1.2.i.),

$$\exists x^1 \in Fr(Y) \text{ tal que } x^1 = \lambda x^o + (1-\lambda)x^* \qquad 0 < \lambda < 1$$

Por tanto,

$$F(x^1) = F(\lambda x^o + (1-\lambda)x^*) \underset{\substack{\uparrow \\ F \text{ convexa en } D}}{\leq} \lambda F(x^o + (1-\lambda)F(x^*) \underset{\substack{\uparrow \\ \text{Por ser } F(x^*) < F(x^o)}}{<}$$

$$< \lambda F(x^o) + (1-\lambda)F(x^o) = F(x^o)$$

Luego,

$$F(x^1) < F(x^o)$$

De donde se concluye, que $x^o \notin S_o$.

Contradicción con la hipótesis del teorema, que tiene su origen en suponer que $\exists x^o \in S_o$ tal que $x^o \notin Fr(Y)$.

Luego,

$$\forall x^o \in S_o \text{ es } x^o \in Fr(Y)$$

Es decir,

$$S_o \subset Fr(Y)$$

Se concluye, así, la demostración en el supuesto en el que $F(x)$ es convexa.

Consideraremos, ahora, el caso en el que $F(x)$ es cóncava.

Teniendo en cuenta la relación existente entre las funciones convexa y cóncava, el programa:

$$(P_2) \begin{cases} \text{Max } F(x) \\ \text{bajo las restricciones:} \\ x \in Y \subset D \subset R^n \end{cases}$$

siendo $F(x)$ función cóncava en D, Y conjunto cerrado en R^n, $\text{Int}(Y) \neq \emptyset$, S_o el conjunto de soluciones óptimas de (P_2). Si, $\exists x^* \in Y^c \cap D$ tal que $F(x^*) > F(x^o)$, con $x^o \in S_o$, entonces $S_o \subset \text{Fr}(Y)$,

puede transformarse en el siguiente equivalente:

$$(P_2') \begin{cases} -\text{Min } -F(x) \\ \text{bajo las restricciones:} \\ x \in Y \subset D \subset R^n \end{cases}$$

siendo $-F(x)$ convexa en D, Y cerrado en R^n, $\text{Int}(Y) \neq \emptyset$, S_o el conjunto de soluciones óptimas de $-F(x)$ en Y. Si $\exists x^* \in Y^c \cap D$ tal que $-F(x^*) < -F(x^o)$ con $x^o \in S_o$, entonces $S_o \subset \text{Fr}(Y)$.

Obviamente, el programa (P_2') verifica el teorema 1.2.2., para función objetivo convexa. Por tanto, en virtud de la equivalencia de los programas (P_2) y (P_2'), se cumple para (P_2), cuando $F(x)$ es cóncava, la proposición formulada.

Se termina así, la demostración del teorema 1.2.2.

Nota 1.2.3.

En el Teorema 1.2.2., imponemos la condición de que sea $\mathrm{Int}(Y) \neq \emptyset$, puesto que si $\mathrm{Int}(Y) = \emptyset$, evidentemente el óptimo, si existe, se alcanza necesariamente en la frontera.

1.3. OPTIMALIDAD EN LA FRONTERA EN PROGRAMAS CONVEXOS

Se establecen en este apartado, una serie de teoremas para la caracterización del óptimo en la frontera de las restricciones en programas convexos (con función objetivo convexa o cóncava).

Teorema 1.3.1.

Dado el programa:

$$(P_3) \begin{cases} \begin{Bmatrix} \mathrm{Min}\ F(x) \\ \mathrm{Max}\ F(x) \end{Bmatrix} \\ \text{bajo las restricciones:} \\ \qquad Y \begin{cases} g_i(x) \leq 0 & i \in I = \{1,\dots,m\} \\ x \in R^n \end{cases} \end{cases}$$

siendo $F(x) \left\{ \begin{matrix} \text{convexa} \\ \text{cóncava} \end{matrix} \right\}$ en $D \subset R^n$, $Y \subset D$, $g_i(x)$ $i \in I$ convexas en R^n, $Int(Y) \neq \emptyset$, S_o y M_a los conjuntos de soluciones óptimas de (P_3) y de $\left\{ \begin{matrix} \text{mínimo} \\ \text{máximo} \end{matrix} \right\}$ absoluto de $F(x)$ en D, respectivamente. Entonces, $S_o \subset Fr(Y)$ si y sólo si:

$$M_a \subset (Int(Y))^c \cap D \quad \text{ó} \quad \exists x^* \in Y^c \cap D \quad \text{tal que} \left\{ \begin{matrix} F(x^*) < F(x^o) \\ F(x^*) > F(x^o) \end{matrix} \right\}, x^o \in$$

Demostración

Es una consecuencia del teorema 1.2.1.

En efecto, por la forma de definir el conjunto Y y por ser $g_i(x)$ $i \in I$ convexas, el conjunto Y es la intersección de cerrados en R^n para la topología usual y, por tanto, Y es cerrado en R^n. Luego, se verifican las hipótesis del teorema 1.2.1. y por consiguiente la tesis.

Concluye así, la demostración.

El teorema que vamos a establecer a continuación, se puede considerar como una consecuencia, en parte, del teorema anterior o bien como un corolario del teoréma 1.2.2.

Teorema 1.3.2.

Dado el programa:

$$(P_4) \begin{cases} \begin{Bmatrix} \text{Min } F(x) \\ \text{Max } F(x) \end{Bmatrix} \\ \text{bajo las restricciones:} \\ \quad y \begin{cases} g_i(x) \leqslant 0 & i \in I = \{1,\dots,m\} \\ x \in R^n \end{cases} \end{cases}$$

siendo $F(x)$ función $\begin{Bmatrix} \text{convexa} \\ \text{cóncava} \end{Bmatrix}$ en un conjunto $D \subset R^n$, $Y \subset D$, $Int(Y) \neq \emptyset$, S_o el conjunto de soluciones óptimas del programa (P_4). Si $\exists x^* \in Y^c \cap D$ tal que $\begin{Bmatrix} F(x^*) < F(x^o) \\ F(x^*) > F(x^o) \end{Bmatrix}$ con $x^o \in S_o$, entonces $S_o \subset Fr(Y)$.

Demostración

Es una consecuencia del teorema 1.2.2., ya que por la forma de definir el conjunto Y mediante desigualdades no estrictas y por ser $g_i(x)$ $i \in I$ convexas, el conjunto Y es la intersección de cerrados en R^n con la topología usual y, por tanto, Y es cerrado en R^n. Luego, se verifican las hipótesis del teorema 1.2.2. y por consiguiente la conclusión.

Veamos, ahora, algunos casos particulares de funcio nes objetivo $\left\{\begin{matrix}\text{convexas}\\ \text{cóncavas}\end{matrix}\right\}$ a las que podemos aplicar total o parcialmente el teorema 1.3.2.

Teorema 1.3.3.

Dado el programa:

$$(P_5)\quad \left\{\begin{array}{l} \left\{\begin{array}{l}\text{Min}\left\{F(x) = x^T C x\right\}\\ \text{Max}\left\{F(x) = x^T C x\right\}\end{array}\right\} \\ \text{bajo las restricciones:} \\ Y \left\{\begin{array}{ll} g_i(x) \leqslant 0 & i \in I = \{1,\dots,m\} \\ x \in R^n \end{array}\right. \end{array}\right.$$

siendo $F(x)$ forma cuadrática definida $\left\{\begin{matrix}\text{positiva}\\ \text{negativa}\end{matrix}\right\}$ en R^n, con $C \in M(n,n)$ matriz en R, $g_i(x)$ $i \in I$ funciones convexas en R^n. Si $0 \notin Y$ y x^o es solución óptima de (P_5), entonces $x^o \in Fr(Y)$.

Demostración

Al ser $F(x) = x^T C x$ forma cuadrática definida $\left\{\begin{matrix}\text{positiva}\\ \text{negativa}\end{matrix}\right\}$, se sigue que $\left\{\begin{matrix} F(x) = x^T C x > 0 \\ F(x) = x^T C x < 0\end{matrix}\right\} \forall x \neq 0$.

Como $0 \notin Y$, se concluye que $\exists\, x^* = 0 \notin Y$ tal que $\left\{\begin{matrix} F(x^*) = F(0) < F(x^o) \\ F(x^*) = F(0) > F(x^o)\end{matrix}\right\}$ y aplicaríamos el teorema 1.3.2.

Nota 1.3.1.

Bajo las hipótesis del teorema anterior, si las funciones $g_i(x)$ $i \in I$ son lineales, como son $\left\{\begin{matrix}\text{cónvexas}\\ \text{cóncavas}\end{matrix}\right\}$, se concluye por el teorema 1.3.3. que el $\left\{\begin{matrix}\text{mínimo}\\ \text{máximo}\end{matrix}\right\}$ está en la frontera de las restricciones.

Teorema 1.3.4.

Dado el programa:

$$(P_6) \begin{cases} \left\{\begin{matrix} \text{Min } \left\{F(x) = x^T C\, x\right\} \\ \text{Max } \left\{F(x) = x^T C\, x\right\} \end{matrix}\right\} \\ \text{bajo las restricciones:} \\ Y \begin{cases} g_i(x) \leq 0 & i \in I = \left\{1,\ldots,m\right\} \\ x \geq 0 \\ x \in R^n \end{cases} \end{cases}$$

siendo $F(x)$ forma cuadrática definida $\left\{\begin{matrix}\text{positiva}\\ \text{negativa}\end{matrix}\right\}$ en R^n, con $C \in M(n,n)$ matriz en R, $g_i(x)$ $i \in I$ funciones convexas en R^n. Si x^o es solución óptima de (P_6), entonces $x^o \in Fr(Y)$.

Demostración

Si $0 \in Y$, por definición de Y se sigue que $0 \in Fr(Y)$

y obviamente el $\left\{\begin{matrix}\text{mínimo}\\ \text{máximo}\end{matrix}\right\}$, por ser $F(x)$ forma cuadrática definida $\left\{\begin{matrix}\text{positiva}\\ \text{negativa}\end{matrix}\right\}$, es $F(0) = 0$. Ahora bien, como la solución óptima es única, según el teorema A.3. (Apéndice), se sigue que $x^o = 0 \in Fr(Y)$.

Si $0 \notin Y$, ello implica que:

$$\exists\, x^* = 0 \in Y^c \quad \text{tal que} \quad \left\{\begin{matrix} F(x^*) = F(0) < F(x^o) \\ F(x^*) = F(0) > F(x^o) \end{matrix}\right\}$$

y se aplicaría el teorema 1.3.2.

Nota 1.3.2.

En particular, si las funciones $g_i(x)$ $i \in I$ son lineales, como son $\left\{\begin{matrix}\text{convexas}\\ \text{cóncavas}\end{matrix}\right\}$, se concluye por este teorema 1.3.4. que el $\left\{\begin{matrix}\text{mínimo}\\ \text{máximo}\end{matrix}\right\}$ del programa está en la frontera de las restricciones.

Teorema 1.3.5.

Dado el programa:

$$(P_7)\left\{\begin{array}{l} \left\{\begin{matrix} \text{Min}\left\{F(x) = x^T\, C\, x + p^T\, x\right\} \\ \text{Max}\left\{F(x) = x^T\, C\, x + p^T\, x\right\} \end{matrix}\right\} \\ \text{bajo las restricciones:} \\ \quad y \left\{\begin{array}{ll} g_i(x) \leq 0 & i \in I = \{1,\dots,m\} \\ x \in R^n & \end{array}\right. \end{array}\right.$$

siendo $x^T\, C\, x$ forma cuadrática definida $\left\{\begin{matrix}\text{positiva}\\ \text{negativa}\end{matrix}\right\}$ en R^n, $p^T\, x$ forma lineal en R^n, con $C \in M(n,n)$ y $p \in M(n,1)$

matrices en R, y $g_i(x)$ $i \in I$ funciones convexas en R^n. Si $\forall x \in Y$ es $\left\{\begin{matrix} F(x) > 0 \\ F(x) < 0 \end{matrix}\right\}$ y si x^o es solución óptima de (P_7), entonces $x^o \in Fr(Y)$.

Demostración

Como $\forall x \in Y$ es $\left\{\begin{matrix} F(x) > 0 \\ F(x) < 0 \end{matrix}\right\}$, se sigue que $0 \notin Y$. Luego, $\exists x^* = 0 \in Y^c$ tal que $\left\{\begin{matrix} F(x^*) = F(0) = 0 < F(x^o) \\ F(x^*) = F(0) = 0 > F(x^o) \end{matrix}\right\}$, y aplicaríamos el teorema 1.3.2.

Teorema 1.3.6.

Dado el programa:

$$(P_8) \left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{Min} \left\{ F(x) = x^T C x + p^T x \right\} \\ \text{Max} \left\{ F(x) = x^T C x + p^T x \right\} \end{array} \right\} \\ \text{bajo las restricciones:} \\ \quad Y \left\{ \begin{array}{ll} g_i(x) \leqslant 0 & i \in I = \left\{ 1, \ldots, m \right\} \\ x \geqslant 0 & \\ x \in R^n & \end{array} \right. \end{array} \right.$$

siendo $x^T C x$ forma cuadrática definida $\left\{\begin{matrix} \text{positiva} \\ \text{negativa} \end{matrix}\right\}$ en R^n, $p^T x$ forma lineal en R^n, $C \in M(n,n)$ y $p \in M(n,1)$ matrices en R, y $g_i(x)$ $i \in I$ funciones convexas en R^n. Si $\forall x \in Y$, $x \neq 0$, es $\left\{\begin{matrix} F(x) > 0 \\ F(x) < 0 \end{matrix}\right\}$ y si x^o es solución óptima de (P_8), entonces $x^o \in Fr(Y)$.

Demostración

Si $0 \in Y$, se sigue por la forma de definir Y que $0 \in Fr(Y)$. Además, el valor $\begin{Bmatrix} \text{mínimo} \\ \text{máximo} \end{Bmatrix}$ de $F(x)$ en Y es $F(0) = 0$ y, por tanto, 0 es solución óptima de (P_8). Ahora bien, al ser $F(x)$ estrictamente $\begin{Bmatrix} \text{convexa} \\ \text{cóncava} \end{Bmatrix}$, se concluye por el teorema A.3. (Apéndice), que el óptimo es único. Luego, $x^o = 0 \in Fr(Y)$.

Si $0 \notin Fr(Y)$, se deduce que:

$$\exists x^* = 0 \in Y^c \quad \text{tal que} \begin{Bmatrix} F(x^*) = F(0) = 0 < F(x^o) \\ F(x^*) = F(0) = 0 > F(x^o) \end{Bmatrix}$$

y se aplicaría el teorema 1.3.2.

Nota 1.3.3.

El teorema 1.3.2. obtenido por nosotros, también se puede aplicar a programas lineales, concluyéndose de dicho teorema que en programas lineales el óptimo está en la frontera de las restricciones.

1.4. UNICIDAD DEL OPTIMO EN PROGRAMAS CONVEXOS

De la teoría existente acerca de la unicidad del óptimo se sabe, teorema A.3. (Apéndice), que el $\begin{Bmatrix} \text{mínimo} \\ \text{máximo} \end{Bmatrix}$ es único si la función objetivo es $\begin{Bmatrix} \text{convexa} \\ \text{cóncava} \end{Bmatrix}$ estricta y las restricciones son convexas. Nosotros vamos a probar un nuevo teorema de unicidad, en el que la función objetivo sólo se le exige que sea $\begin{Bmatrix} \text{convexa} \\ \text{cóncava} \end{Bmatrix}$.

Teorema 1.4.1.

Dado el programa:

$$(PG_1) \left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{Min } F(x) \\ \text{Max } F(x) \end{array} \right\} \\ \text{bajo las restricciones:} \\ Y \left\{ \begin{array}{ll} g_i(x) \leq 0 & i \in I = \{1,..,m\} \\ x \in R^n \end{array} \right. \end{array} \right.$$

siendo $F(x)$ función $\left\{ \begin{array}{l} \text{convexa} \\ \text{cóncava} \end{array} \right\}$ en $D \subset R^n$, $Y \subset D$, y $g_i(x)$ $i \in I$ funciones convexas estrictas en R^n. Si el conjunto de soluciones óptimas de (PG_1) S_o es distinto del vacío y está contenido en la frontera de Y, entonces S_o se reduce a un punto.

Demostración

Veamos el siguiente:

Lema previo:

Cualquier subconjunto U, $U \neq \emptyset$ y U no reducido a un punto, contenido en la frontera de un conjunto convexo Y definido por restricciones convexas estrictas, no es un conjunto convexo. En particular, como $Fr(Y) \subset Fr(Y)$, se concluye que dicha frontera no es un conjunto convexo.

Demostración

Sea:

$Y=\{x\in R^n \ / \ g_i(x)\leqslant 0 \quad i\in I=\{1,..,m\},\ g_i \quad i\in I$ convexas estrictas

Se sabe, teorema A.1. (Apéndice), que Y es un conjunto convexo.

Sea $\forall U\subset Fr(Y)$, $U\neq\emptyset$ y U no unitario.

Si U posee puntos aislados, ello implica que U no es convexo y el lema queda demostrado.

Si U no tiene puntos aislados, supongamos que U fuera convexo.

Ello equivale:

$$\forall x^1, x^2\in U,\quad \forall\lambda\in[0,1] \quad \Rightarrow \quad \lambda x^1 + (1-\lambda)x^2\in U$$

Tomemos $\quad 0<\lambda<1$

Como $U\subset Fr(Y)$, ello implica

$$\exists\, i\in I \text{ tal que } g_i\left[\lambda x^1 + (1-\lambda)x^2\right] = 0 \qquad (1)$$

Por otra parte, para tal $i\in I$ por ser g_i función convexa estricta, se verifica que:

$$g_i\left[\lambda x^1 + (1-\lambda)x^2\right] < \lambda g_i(x^1) + (1-\lambda)\, g_i(x^2) \qquad (2)$$

Pero,

$$g_i(x^1)\leqslant 0 \quad \text{por ser } x^1\in U\subset Fr(Y)$$

$$g_i(x^2)\leqslant 0 \quad \text{por ser } x^2\in U\subset Fr(Y)$$

Luego, en (2):

$$g_i\left[\lambda x^1 + (1-\lambda)x^2\right] < \lambda g_i(x^1) + (1-\lambda)\, g_i(x^2)\leqslant 0$$

Es decir,

$$g_i\left[\lambda x^1 + (1-\lambda)\, x^2\right] < 0 \qquad (3)$$

De (1) y (3) se sigue una contradicción, que nace de suponer que U es convexo. Luego, U no es convexo.

En particular, Fr(Y) no es un conjunto convexo.

Queda probado el lema.

Demostración del teorema

La haremos únicamente cuando la función objetivo es convexa, ya que cuando es cóncava el razonamiento es completamente análogo.

Supongamos que S_o no se reduce a un punto.

Sabemos, por el teorema A.2. del Apéndice, que:

$$S_o \text{ es convexo} \qquad (4)$$

De otra parte, por hipótesis $S_o \subset Fr(Y)$ y de acuerdo con el lema previo, se sigue que:

$$S_o \text{ no es convexo} \qquad (5)$$

De (4) y (5) surge una contradicción, que se origina al suponer que S_o no es unitario. Luego, la solución óptima es única.

Queda así, demostrado el teorema.

1.5. UNICIDAD Y OPTIMALIDAD EN LA FRONTERA EN PROGRAMAS CONVEXOS

Como una consecuencia de lo establecido en los apartados 1.3. y 1.4., formulamos los siguientes resultados:

Teorema 1.5.1.

Dado el programa:

$$(PG_2)\begin{cases} \begin{Bmatrix} \text{Min } F(x) \\ \text{Max } F(x) \end{Bmatrix} \\ \text{bajo las restricciones:} \\ \text{Y} \quad \begin{cases} g_i(x) \leqslant 0 & i \in I = \{1,\ldots,m\} \\ x \in R^n \end{cases} \end{cases}$$

siendo $F(x)$ función $\begin{Bmatrix} \text{convexa} \\ \text{cóncava} \end{Bmatrix}$ estricta en $D \subset R^n$, $Y \subset D$, $g_i(x)$ $i \in I$ funciones convexas en R^n, S_o el conjunto de soluciones óptimas de (PG_2) y M_a el conjunto de soluciones de $\begin{Bmatrix} \text{mínimo} \\ \text{máximo} \end{Bmatrix}$ absoluto de $F(x)$ en D (que puede ser vacío). Si $x^o \in S_o$ y

$M_a \subset \mathring{Y}^c \cap D$ ó $\exists x^* \in Y^c \cap D$ tal que $\begin{Bmatrix} F(x^*) < F(x^o) \\ F(x^*) > F(x^o) \end{Bmatrix}$,

entonces $x^o \in Fr(Y)$ y x^o es solución óptima única.

Demostración

Por el teorema 1.3.1., se sigue que $x^o \in Fr(Y)$ y por el teorema A.3. del Apéndice, se concluye que x^o es solución óptima única.

Teorema 1.5.2.

Dado el programa:

$$(PG_3)\begin{cases}\begin{Bmatrix}\text{Min } F(x)\\ \text{Max } F(x)\end{Bmatrix}\\ \text{bajo las restricciones:}\\ \text{y } \begin{cases} g_i(x) \leqslant 0 & i \in I = \{1,\dots,m\}\\ x \in R^n\end{cases}\end{cases}$$

siendo $F(x)$ función $\begin{Bmatrix}\text{convexa}\\ \text{cóncava}\end{Bmatrix}$ en $D \subset R^n$, $Y \subset D$, $g_i(x)$ $i \in I$ funciones convexas estrictas en R^n, S_o el conjunto de soluciones óptimas de (PG_3) y M_a el conjunto de soluciones de $\begin{Bmatrix}\text{mínimo}\\ \text{máximo}\end{Bmatrix}$ absoluto de $F(x)$ en D (M_a puede ser vacío). Si $x^o \in S_o$ y

$M_a \subset \mathring{Y}^c \cap D$ ó $\exists x^* \in Y^c \cap D$ tal que $\begin{Bmatrix}F(x^*) & < F(x^o)\\ F(x^*) & > F(x^o)\end{Bmatrix}$,

entonces $x^o \in Fr(Y)$ y x^o es solución óptima única.

Demostración

Por el teorema 1.3.1. se concluye que $S_o \subset Fr(Y)$ y por el teorema 1.4.1. se sigue que $S_o = \{x^o\}$.

1.6. TEOREMAS EQUIVALENTES AL DE KUHN Y TUCKER EN PROGRAMAS CUADRATICOS

Establecemos, en este epígrafe, teoremas que caracterizan la solución óptima. Dichos teoremas ponen de manifiesto la relación existente entre la hipersuperficie de nivel (relativa a la función objetivo) y la frontera de las restricciones en el punto de óptimo, así como la conexión entre las distintas matrices del programa.

Teorema 1.6.1.

Dado el programa:

$$(PQ_1) \begin{cases} \text{Min}\left\{F(x) = x^T C x\right\} \\ \text{bajo las restricciones:} \\ \quad Y \begin{cases} A x \geq b \\ x \geq 0 \\ x \in R^n \end{cases} \end{cases}$$

siendo $F(x)$ forma cuadrática definida positiva, $C \in M(n,n)$, $A \in M(m,n)$ con $rg(A) = m$, $b \in M(m,1)$ y todas las matrices de números reales. Entonces, x^o es óptimo de (PQ_1) si y sólo si:

1°) $\{x^o\} = \{x \in R^n \;/\; F(x^o) = x^T C x\} \cap Fr(Y)$

2°) $S^{-1} \bar{b} \geq 0, \quad S^{-1} = \frac{1}{2} \bar{A} C^{-1} \bar{A}^T$

siendo $\bar{A} \in M(h+k,n)$ y $\bar{b} \in M(h+k,1)$, respectivamente, las matrices de coeficientes y de términos independientes correspondientes a las restricciones saturadas (*) linealmente independientes en el punto x^o.

Demostración

$\Rightarrow$ Si x^o es solución óptima de (PQ_1), ello equivale a que x^o verifica las condiciones de Kuhn y Tucker.

La función Lagrangiana es:

$$L(x,\lambda,\mu) = x^T C x + \lambda^T(-Ax+b) - \mu^T x$$

Las condiciones de Kuhn y Tucker relativas al punto de silla $(x^o, \lambda^o, \mu^o) \in R^{n+m+n}$ de la función L que proporciona el x^o óptimo de (PQ_1) son:

$$\exists \lambda^o \geq 0, \ \lambda^o \in R^m, \ \exists \mu^o \geq 0, \ \mu^o \in R^n \text{ tales que:}$$

1°) $\nabla_x L(x^o, \lambda^o, \mu^o) = 0$

2°) $\lambda_i^o (-A_{i*} x^o + b_i) = 0 \qquad i = 1,\dots, m$

$\mu_j^o x_j^o = 0 \qquad j = 1,\dots, n$

(*) Definición (de restricción saturada)

Dado el programa $\begin{cases} \text{Min } F(x) \\ \text{bajo las restricciones:} \\ Y \begin{cases} g_i(x) \leq 0 \\ x \in R^n \end{cases} \end{cases}$

si en cierto punto $x^o \in Y$ se verifica que $g_i(x^o)=0$, se dice que $g_i(x)$ es una restricción saturada o activa en x^o.

En este caso, dichas condiciones se convierten en:

1°) $2\ C\ x^{o} - A^{T}\ \lambda^{o} - \bar{\mu}^{o} = 0$

2°) $\lambda_{i}^{o}\ (-\ A_{i*}\ x^{o} + b_{i}) = 0 \qquad i = 1,\ldots,m$

$\mu_{j}^{o}\ x_{j}^{o} = 0 \qquad j = 1,\ldots,n$

Por el teorema 1.3.1.4. del apartado 1.3., se sigue que:

$$x^{o} \in Fr(Y) \qquad (1)$$

Por otra parte, dada la hipersuperficie de nivel $F(x^{o}) = x^{T}\ C\ x$ es obvio que:

$$x^{o} \in \{x \in R^{n}\ /\ F(x^{o}) = x^{T}\ C\ x\} \qquad (2)$$

De (1) y (2) se concluye que:

$$x^{o} \in \{x \in R^{n}\ /\ F(x^{o}) = x^{T}\ C\ x\} \cap Fr(Y)$$

Por el teorema A.3. del Apéndice, se sigue que x^{o} es único. Luego,

$$\{x \in R^{n}\ /\ F(x^{o}) = x^{T}\ C\ x\} \cap Fr(Y) = \{\ x^{o}\ \}$$

Es decir, se verifica la primera hipótesis del teorema.

Vamos a demostrar, ahora, que el teorema de Kuhn y Tucker implica la segunda hipótesis del teorema.

Supongamos que las restricciones saturadas o activas linealmente independientes en el punto x^{o} de óptimo son:

$$\begin{array}{ll} a_{11}\, x_1^o + \ldots\ldots\ldots + a_{1n}\, x_n^o = b_1 & \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{h1}\, x_1^o + \ldots\ldots\ldots + a_{hn}\, x_n^o = b_h & (3) \\ x_1^o = 0 & \\ x_k^o = 0 & \end{array}$$

siendo $h \leqslant m$ y $k \leqslant n$ y $h+k \leqslant n$ por ser ecuaciones linealmente independientes.

Y en forma matricial:

$$\begin{pmatrix} a_{11} \ldots\ldots a_{1n} \\ \ldots\ldots\ldots\ldots \\ a_{h1} \ldots\ldots a_{hn} \\ 1 \ldots\ldots\ldots 0 \\ \ldots\ldots\ldots\ldots \\ 0 \ldots 1 \ldots 0 \end{pmatrix} \begin{pmatrix} x_1^o \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n^o \end{pmatrix} = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ b_h \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

Es decir,

$$\bar{A}\ x^o = \bar{b} \qquad (4)$$

$\bar{A} \in M(h+k,n)$, $x^o \in M(n,1)$, $\bar{b} \in M(h+k,1)$

Teniendo en cuenta la segunda condición de Kuhn y Tucker, para las restricciones no saturadas en x^o, se cumple que:

$$\begin{array}{ll} \mu^o_{k+1} = 0\ , \ldots\ldots\ldots ,\ \mu^o_n = 0 & \\ & (5) \\ \lambda^o_{h+1} = 0\ , \ldots\ldots\ldots ,\ \lambda^o_m = 0 & \end{array}$$

La primera condición de Kuhn y Tucker, queda:

$$2\ C\ x^{o} = A^{T}\ \lambda^{o} + \mu^{o}$$

Luego,

$$2\ C\ x^{o} = \begin{pmatrix} a_{11} \cdots\cdots a_{m1} \\ \vdots \qquad\quad \vdots \\ a_{1n} \cdots\cdots a_{mn} \end{pmatrix} \begin{pmatrix} \lambda^{o}_{1} \\ \\ \lambda^{o}_{m} \end{pmatrix} + \begin{pmatrix} \mu^{o}_{1} \\ \\ \mu^{o}_{n} \end{pmatrix}$$

Y teniendo en cuenta las expresiones (3) y (5):

$$2\ C\ x^{o} = \begin{pmatrix} a_{11} \cdots\cdots a_{h1} \\ \cdots\cdots\cdots \\ a_{1n} \cdots\cdots a_{hn} \end{pmatrix} \begin{pmatrix} \lambda^{o}_{1} \\ \vdots \\ \lambda^{o}_{h} \end{pmatrix} + \begin{pmatrix} \mu^{o}_{1} \\ \vdots \\ \mu^{o}_{k} \\ \vdots \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} \cdots & a_{h1} & 1 \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{1k} \cdots & a_{hk} & 0 \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} \cdots & a_{hn} & 0 \cdots & 0 \end{pmatrix} \begin{pmatrix} \lambda^{o}_{1} \\ \vdots \\ \lambda^{o}_{h} \\ \mu^{o}_{1} \\ \vdots \\ \mu^{o}_{k} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Y sintetizando:

$$2\ C\ x^{o} = \bar{A}^{T}\ \ \bar{\lambda}^{o} \qquad (6)$$

siendo $\bar{A}^{T} \in M(n,h+k)$, $\bar{\lambda}^{o} \in M(h+k,1)$

Por ser C matriz simétrica definida positiva, se sigue que $\det(C) \neq 0$ lo que implica que $\exists\, C^{-1}$.

Por tanto, despejando de (6)

$$x^o = \frac{1}{2} C^{-1} \bar{A}^T \bar{\lambda}^o$$

Sustituyendo en (4)

$$\frac{1}{2} \bar{A} C^{-1} \bar{A}^T \bar{\lambda}^o = \bar{b}$$

Llamando $S = \frac{1}{2} \bar{A} C^{-1} \bar{A}^T$, nos queda:

$$S \bar{\lambda}^o = \bar{b} \qquad (7)$$

Pero (7) es un sistema lineal en $\bar{\lambda}^o$ que tiene solución única (por ser linealmente independientes los vectores gradientes de las restricciones saturadas en x^o), luego aplicando el teorema de Rouché-Frobenius, se sigue que:

$$rg(S) = h + k \iff \det(S) \neq 0 \iff \exists\, S^{-1} \implies$$

$$\implies \bar{\lambda}^o = S^{-1} \bar{b}$$

Pero, de acuerdo con el teorema de Kuhn y Tucker, es $\bar{\lambda}^o \geqslant 0$. Luego, $S^{-1} \bar{b} \geqslant 0$ y queda probada la segunda hipótesis.

$\Longleftarrow$ Recíprocamente, probemos que las hipótesis primera y segunda del enunciado del teorema implican las condiciones de Kuhn y Tucker en x^o.

De la primera hipótesis se sigue que $x^o \in Fr(Y)$.

Supongamos que las restricciones saturadas o activas linealmente independientes en el punto x^o son:

$$
\begin{array}{ll}
a_{11}\,x_1 + \dots\dots + a_{1n}x_n = b_1 & \\
\dots\dots\dots\dots\dots\dots\dots\dots & \\
a_{h1}\,x_1 + \dots\dots + a_{hn}x_n = b_h & \\
x_1 \dots\dots\dots\dots\dots = 0 & (8) \\
\ddots & \\
\quad x_k \dots\dots\dots = 0 &
\end{array}
$$

siendo $h \leqslant m$, $k \leqslant n$ y por suponer que son linealmente independientes $h+k \leqslant n$.

Llamando $H_i (i = 1,\dots,h)$ y $C_j (j = 1,\dots,k)$ a los hiperplanos de R^n determinados, respectivamente, por

$$\sum_{d=1}^{n} a_{id}\, x_d = b_i \qquad i = 1,\dots,h$$

$$x_j = 0 \qquad j = 1,\dots,k$$

se obtiene que la variedad afín que determinan las ecuaciones (8) es:

$$D = H_1 \cap \dots \cap H_k \cap C_1 \cap \dots \cap C_k$$

cuya dimensión, por suponer las ecuaciones (8) linealmente independientes, es $n - (h+k)$.

Pueden darse dos posibilidades:

a) $n - (h+k) > 0 \iff D$ no se reduce al punto x^o

b) $n - (h+k) = 0 \iff D$ se reduce al punto x^o

Sean:

$$a_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \quad y \quad e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow (j) \\ \\ \\ \end{matrix}$$

los vectores asociados a los hiperplanos

$H_i \quad i = 1,..,h$

$C_j \quad j = 1,..,k$

Como $D = H_1 \cap \ldots \cap H_h \cap C_1 \cap \ldots \cap C_k \Rightarrow$

$$\Rightarrow \begin{cases} a_i \text{ ortogonal a } D & i = 1,\ldots,h \\ e_j \text{ ortogonal a } D & j = 1,\ldots,k \end{cases}$$

Luego,

$$D \text{ es ortogonal a } < a_1,\ldots,a_h,e_1,\ldots,e_k >$$

Si ocurre la posibilidad a), $\dim D = n - (h+k) > 0$ podemos expresar R^n como la siguiente suma directa:

$$R^n = D' \oplus <a_1,\ldots,a_h,e_1,\ldots,e_k>$$

siendo D' el subespacio vectorial subyacente de D.

Si consideramos el hiperplano tangente en x^o a la hipersuperficie de nivel $S = \{x \in R^n / F(\hat{x}^o) = x^T C x\}$, se verifica que dicho hiperplano tangente H contiene a $D^{(*)}$.

(*) Dicha propiedad no es obvia. Se demuestra más adelante en la Nota 1.6.1.

Por otra parte, $F(x^o)$ es ortogonal a H.

Como $D \subset H$, se sigue que:

$\nabla F(x^o)$ es ortogonal a D

Por tanto,

$$\nabla F(x^o) \in < a_1,\ldots,a_h,e_1,\ldots,e_k>$$

Si ocurre la posibilidad b):

$$\dim D = n - (h+k) = 0 \iff n = h+k$$

Luego,

$< a_1,\ldots,a_h,e_1,\ldots,e_k >$ es una base de R^n (ya que hemos supuesto que tales vectores son linealmente independientes).

Por tanto, cualquier vector de R^n se puede expresar como combinación lineal de la misma, en particular $\nabla F(x^o)$. Luego, $\nabla F(x^o) \in <a_1,\ldots,a_h,e_1,\ldots,e_k >$.

Así, pues, para ambas posibilidades existen números reales $\exists \lambda_i^o$, $i = 1,\ldots,h$ y $\exists\, \mu_j^o$, $j = 1,\ldots,k$ únicos, tales que:

$$\nabla F(x^o) = 2\ C\ x^o = \lambda_1^o\ a_1 + \ldots + \lambda_h^o\ a_h + \mu_1^o\ e_1 + \ldots + \mu_k^o\ e_k =$$

$$= \lambda_1^o \begin{pmatrix} a_{11} \\ \cdot \\ \cdot \\ \cdot \\ a_{1n} \end{pmatrix} + \ldots + \lambda_h^o \begin{pmatrix} a_{h1} \\ \cdot \\ \cdot \\ \cdot \\ a_{hn} \end{pmatrix} + \mu_1^o \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \cdot \\ 0 \end{pmatrix} + \ldots + \mu_k^o \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \cdot \\ 0 \end{pmatrix} \leftarrow (k$$

$$= \begin{pmatrix} a_{11} \cdots\cdots a_{h1} \\ \cdots\cdots\cdots \\ \cdots\cdots\cdots \\ a_{1n} \cdots\cdots a_{hn} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \vdots \\ \vdots \\ \lambda_h^o \end{pmatrix} + \begin{pmatrix} \mu_1^o \\ \vdots \\ \mu_k^o \\ 0 \\ \vdots \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} \cdots a_{h1} & a_{h+1,1} \cdots a_{m1} \\ \cdots\cdots & \cdots\cdots \\ a_{1n} \cdots a_{hn} & a_{h+1,n} \cdots a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \vdots \\ \lambda_h^o \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \mu_1^o \\ \vdots \\ \mu_k^o \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Es decir:

$$2\, C\, x^o = A^T\, \lambda^o + \mu^o \qquad (9)$$

que es la primera condición de Kuhn y Tucker si demostramos que $\lambda^o \geqslant 0$ y $\mu^o \geqslant 0$.

Por otra parte, para las restricciones saturadas (8), se verifica que:

$$\left.\begin{aligned} \lambda_i^o\, (-a_{i1}\, x_1^o - \ldots - a_{in}\, x_n^o + b_i) = 0 \qquad & i = 1,..,h \\ \mu_j^o\, x_j^o = 0 \qquad & j = 1,..,k \end{aligned}\right\} \quad (10)$$

Y para las restricciones no saturadas en x^o, se cumple también que:

$$\lambda_r^o \; (-a_{r1}x_1^o - \ldots - a_{rn}x_n^o + b_r) = 0 \qquad r=h+1,\ldots,m$$

$$\mu_t^o \; x_t^o = 0 \qquad t = k+1,\ldots,n \qquad (11)$$

ya que por la forma de definir λ^o y μ^o es:

$$\lambda_r^o = 0 \qquad r = h+1,\ldots,m$$

$$\mu_t^o = 0 \qquad t = k+1,\ldots,n$$

Pero las ecuaciones (10) y (11) son la segunda condición de Kuhn y Tucker, si probamos que $\lambda^o \geqslant 0$ y $\mu^o \geqslant 0$.

Demostremos que $\lambda^o \geqslant 0$ y que $\mu^o \geqslant 0$.

Obviamente, para llegar a la expresión (9), según hemos visto antes, es:

$$\lambda_r^o = 0 \qquad r = h+1,\ldots,m$$

$$\mu_t^o = 0 \qquad t = k+1,\ldots,n$$

Luego, sólo nos falta probar que $\lambda_i^o \geqslant 0$ $i = 1,..,h$ $\mu_j^o \geqslant 0$ $j = 1,\ldots,k$

La expresión (9) puede reagruparse así:

$$2\,C\,x^o = \begin{pmatrix} a_{11} \ldots a_{h1} \\ \ldots\ldots \\ a_{1n} \ldots a_{hn} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \vdots \\ \lambda_h^o \end{pmatrix} + \begin{pmatrix} \mu_1^o \\ \vdots \\ \mu_k^o \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad =$$

$$= \begin{pmatrix} a_{11} \cdots a_{h1} & 1 \ldots 0 \\ \cdots\cdots\cdots\cdots & \\ a_{1k} \cdots a_{hk} & 0 \ldots 1 \\ \cdots\cdots\cdots\cdots & \\ a_{1n} \cdots a_{hn} & 0 \ldots 0 \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \vdots \\ \lambda_h^o \\ \mu_1^o \\ \vdots \\ \mu_k^o \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Y resumiendo:

$$2 \; C \; x^o = \bar{A}^T \; \bar{\lambda}^o \qquad (12)$$

siendo $\bar{A}^T \in M(n,h+k)$ y $\bar{\lambda}^o \in M(h+k,1)$.

Por otra parte, las restricciones saturadas linealmente independientes en el punto x^o, que aparecen en la expresión (8), podemos agruparlas así:

$$\begin{pmatrix} a_{11} \cdots a_{1k} \cdots a_{1n} \\ \cdots\cdots\cdots\cdots \\ a_{h1} \cdots a_{hk} \cdots a_{hn} \\ 1 \ldots\ldots 0 \ldots\ldots 0 \\ \cdots\cdots\cdots\cdots \\ 0 \ldots\ldots 1 \ldots\ldots 0 \end{pmatrix} \begin{pmatrix} x_1^o \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n^o \end{pmatrix} = \begin{pmatrix} b_1 \\ \cdot \\ b_h \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

Es decir:

$$\bar{A} \; x^o = \bar{b} \qquad (13)$$

siendo $\bar{A} \in M(h+k,n)$ y $\bar{b} \in M(h+k,1)$

Vamos a determinar el valor de $\bar{\lambda}^o$, para lo cual despejaremos x^o en la ecuación (12) y lo sustituiremos en la ecuación (13).

En (12), como C es matriz simétrica definida positiva, se sigue que:

$$\det(C) \neq 0 \iff \exists\, C^{-1}$$

Luego,

$$x^{o} = \frac{1}{2}\, C^{-1}\, \bar{A}^{T}\ \bar{\lambda}^{o}$$

Y sustituyendo este valor en (13):

$$\frac{1}{2}\, \bar{A}\, C^{-1}\, \bar{A}^{T}\ \bar{\lambda}^{o} = \bar{b}$$

Haciendo $S = \frac{1}{2}\, \bar{A}\, C^{-1}\, \bar{A}^{T}$, queda:

$$S\ \bar{\lambda}^{o} = \bar{b} \qquad (14)$$

siendo S M(h+k,h+k) y además simétrica.

Como tal sistema tiene solución única $\bar{\lambda}^{o}$ (según se ha visto al obtener la igualdad (9)), se sigue que:

$$\det(S) \neq 0 \iff \exists\, S^{-1}$$

Premultiplicando por S^{-1} en (14), obtenemos:

$$\bar{\lambda}^{o} = S^{-1}\, \bar{b}$$

Ahora bien, por la segunda hipótesis es:

$$S^{-1}\, \bar{b} \geq 0$$

de lo que se deduce:

$$\bar{\lambda}^{o} \geq 0 \iff \lambda^{o} \geq 0 \ \text{ y } \ \mu^{o} \geq 0$$

Luego, λ^{o} y μ^{o} verifican el teorema de Kuhn y Tucker.

Queda, por tanto, probada la equivalencia.

Corolario 1.6.1.

Dado el programa cuadrático:

$$(PQ_2)\begin{cases} \text{Min } F(x) = x^T C x \\ \text{bajo las restricciones:} \\ \quad \text{y} \begin{cases} A x \geqslant b \\ x \geqslant 0 \\ x \in R^n \end{cases} \end{cases}$$

siendo $F(x)$ forma cuadrática definida positiva, $C \in M(n,n)$, $A \in M(m,n)$ con $rg(A) = m$, $b \in M(m,1)$ y todas las matrices de números reales. Si x^o verifica:

1°) $x^o \geqslant 0$

2°) $\{x^o\} = \{x \in R^n / F(x^o) = x^T C x\} \cap D$, siendo D la variedad afín solución del sistema $A x = b$.

3°) $(A C^{-1} A^T)^{-1} b \geqslant 0$

entonces, x^o es solución óptima de (PQ_2).

Demostración

Basta aplicar el teorema anterior, teniendo en cuenta que las condiciones primera y segunda implican la primera hipótesis del teorema 1.6.1., y la tercera condición corresponde a la segunda hipótesis del teorema 1.6.1. cuando A y b son las matrices de coeficientes y de términos independientes correspondientes a las restricciones saturadas linealmente independientes en el punto x^o.

Una variante del teorema 1.6.1., con restricciones más generales, es la siguiente:

Teorema 1.6.2.

Dado el programa:

$$(PQ_3)\begin{cases} \text{Min } F(x) = x^T C x \\ \text{bajo las restricciones:} \\ \qquad Y \begin{cases} A x \geq b \\ x \in R^n \end{cases} \end{cases}$$

siendo F(x) forma cuadrática definida positiva, $C \in M(n,n)$, $A \in M(m,n)$, con $rg(A) = m$, $b \in M(m,1)$ y todas las matrices de números reales. Además $0 \notin Y$. Entonces, x^o es solución óptima de (PQ_3) si y sólo si:

1°) $\{ x^o \} = \{ x \in R^n \ / \ F(x^o) = x^T C x \} \cap \ Fr(Y)$

2°) $S^{-1} \bar{b} \geq 0$, $\quad S^{-1} = \frac{1}{2} \bar{A} C^{-1} \bar{A}^T$

siendo $\bar{A}$ la submatriz de A correspondiente a las restricciones saturadas linealmente independientes en x^o y $\bar{b}$ la submatriz de b relativa a los términos independientes de dichas restricciones saturadas linealmente independientes en x^o.

Demostración

Por ser $0 \notin Y$, sabemos por el teorema 1.3.3. del apartado 1.3., que $x^o \in Fr(Y)$. El resto del razonamiento es análogo al del teorema 1.6.1.

Nota 1.6.1.

Pasamos a demostrar la propiedad mencionada en el (*) de la página 34.

Proposición 1.6.1.

Dado el programa:

$$
(PQ_1)\left\{\begin{array}{l} \text{Min}\{F(x) = x^T C x\} \\ \text{bajo las restricciones:} \\ \qquad Y \left\{\begin{array}{l} A x \geq b \\ x \geq 0 \\ x \in R^n \end{array}\right. \end{array}\right.
$$

siendo $F(x)$ forma cuadrática definida positiva, $C \in M(n,n)$, $A \in M(m,n)$, $b \in M(m,1)$, $rg(A) = m$ y todas las matrices de números reales. Sea $x^o \in Y$ verificando la condición primera del teorema 1.7.1., $S = \{x \in R^n / F(x^o) = x^T C x\}$, D la variedad afín que determinan las restricciones saturadas linealmente independientes en x^o, H el hiperplano tangente en x^o a la hipersuperficie de nivel S. Entonces, $D \subset H$.

Demostración

Pueden presentarse dos posibilidades:

1ª) $\dim(D) = 0 \iff D = \{x^o\}$, obviamente, en este caso es $D \subset H$.

2^{a}) $\dim(D) > 0$

La tesis para este caso, la probaremos por reducción al absurdo.

Si $D \not\subset H$, se sigue que:

$\exists d \neq 0,\ d \in D$ tal que $d \notin H$

Veamos que la recta r de vector de dirección d que pasa por x^{o}, corta a S en otro punto $x^{1} \neq x^{o}$.

En efecto, la ecuación de r sería:

$$x = x^{o} + \lambda d \qquad \forall \lambda \in R$$

Hallemos la intersección de dicha recta con S.

$$\begin{aligned} x^{T} C x &= (x^{o} + \lambda d)^{T} C (x^{o} + \lambda d) = \\ &= (x^{o})^{T} C x^{o} + 2 \lambda d^{T} C x^{o} + \lambda^{2} d^{T} C d \end{aligned} \qquad (1)$$

Pero,

$$x^{T} C x = x^{o} C x^{o} \qquad (2)$$

De (1) y (2) se sigue que:

$$\lambda^{2} d^{T} C d + 2 \lambda d^{T} C x^{o} = 0$$

Es decir,

$$\lambda \left[\lambda d^{T} C d + 2 d^{T} C x^{o} \right] = 0$$

cuyas raíces son:

$$\lambda = 0 \quad y \quad \lambda = - \frac{2\, d^T\, C\, x^o}{d^T\, C\, d}$$

Si $\lambda = 0 \Rightarrow x = x^o$

Si $\lambda = - \dfrac{2\, d^T\, C\, x^o}{d^T\, C\, d}$

Para esta segunda raíz, puede ocurrir que:
$\lambda = 0$ ó $\lambda \neq 0$

Si $\lambda = 0$, entonces la recta r es tangente a S en x^o, ello implica que $d \in H$. Por tanto, para este caso, se cumple que $D \subset H$.

Si $\lambda \neq 0$, la recta r corta a S en otro punto $x^1 \neq x^o$. (3)

Ahora bien, como:

$$D \cap S = \{ x^o \} \qquad (4)$$

de (3) y (4) surge una contradicción, que se origina al suponer que $D \not\subset H$. Luego, $D \subset H$.

Queda así demostrada la proposición.

<u>Observación</u>

Una proposición análoga a la anterior, puede establecerse para el teorema 1.6.2.

Teorema 1.6.3.

Dado el programa:

$$(PQ_4)\begin{cases} \text{Min}\left\{Q(x) = x^T C\, x + p^T x\right\} \\ \text{bajo las restricciones:} \\ \qquad Y \begin{cases} A\,x \geq b \\ x \geq 0 \\ x \in R^n \end{cases} \end{cases}$$

siendo C matriz simétrica definida positiva, $C \in M(n,n)$, $p \in M(n,1)$, $A \in M(m,n)$, $rg(A) = m$, $b \in M(m,1)$ y todas las matrices de números reales. Sea $x^o \in Fr(Y)$. Entonces, x^o es óptimo de (PQ_4) si y sólo si:

1°) $\{x^o\} = \{x \in R^n \,/\, Q(x^o) = Q(x)\} \cap Fr(Y)$

2°) $S^{-1}\, T \geq 0$, siendo $S^{-1} = \bar{A}\, C^{-1}\, \bar{A}^T$ y $T = 2\,\bar{b} + \bar{A}\, C^{-1}\, p$, y $\bar{A} \in M(h+k,n)$ y $\bar{b} \in M(h+k,1)$, respectivamente, las matrices de coeficientes y de términos independientes correspondientes a las restricciones saturadas linealmente independientes en x^o.

Demostración

$\Rightarrow$ Si x^o es solución óptima de (PQ_4), ello equivale a que x^o verifica las condiciones de Kuhn y Tucker.

La función de Lagrange asociada al programa (PQ_4) es:

$$L(x, \lambda, \mu) = x^T C x + p^T x + \lambda(-A x + b) - \mu x$$

Las condiciones de Kuhn y Tucker relativas al punto de silla $(x^o, \lambda^o, \mu^o) \in R^{n+m+n'}$ de la función L que proporciona el x^o óptimo de (PQ_4) son:

$\exists \lambda^o \geq 0, \quad \lambda^o \in R^m, \exists \mu^o \geq 0, \ \mu^o \in R^n$ tales que:

1ª) $\nabla_x L(x^o, \lambda^o, \mu^o) = 0 \iff 2 C x^o + p - \bar{A} \quad \lambda^o - \mu^o = 0$

2ª) $\lambda_i^o(-A_{i*} x^o + b_i) = 0 \qquad i=1,..,m$

$\mu_j^o x_j^o = 0 \qquad j = 1,.., n$

Siendo A_{i*} la fila i-ésima de la matriz A

Por hipótesis, $x^o \in Fr(Y)$ (1)

Por otra parte, dada la hipersuperficie de nivel $S = \{ x \in R^n / Q(x^o) = Q(x) \}$ es claro que:

$$x^o \in S \qquad (2)$$

De (1) y (2), se concluye que:

$$x^o \in \{ x \in R^n / Q(x^o) = Q(x) \} \cap Fr(Y)$$

Por el teorema A.3. del Apéndice, se sigue que x^o es único. Luego,

$$\{ x^o \} = \{x \in R^n / Q(x^o) = Q(x) \} \cap Fr(Y)$$

Es decir, se verifica la primera hipótesis del teorema.

Demostremos, ahora, que el teorema de Kuhn y Tucker implica la segunda hipótesis del teorema.

Supongamos que las restricciones saturadas linealmente independientes en el punto x^o del óptimo son:

$$
\begin{array}{l}
a_{11}\, x_1^o + \dots\dots + a_{1n}\, x_n^o = b_1 \\
\dots\dots\dots\dots\dots\dots\dots\dots \\
a_{h1}\, x_1^o + \dots\dots + a_{hn}\, x_n^o = b_h \\
\quad x_1^o \qquad\qquad\qquad\qquad = 0 \\
\qquad\quad x_k^o \qquad\qquad\qquad = 0
\end{array}
$$

siendo $h \leqslant m$, $k \leqslant n$ y $h+k \leqslant n$ por ser ecuaciones linealmente independientes.

Y en forma simplificada:

$$
\begin{pmatrix}
a_{11} \dots\dots & a_{1n} \\
a_{h1} \dots\dots & a_{hn} \\
1 \dots\dots & 0 \\
\dots\dots & \dots \\
0 \dots 1 \dots & 0
\end{pmatrix}
\begin{pmatrix}
x_1^o \\ \vdots \\ \vdots \\ x_n^o
\end{pmatrix}
=
\begin{pmatrix}
b_1 \\ \vdots \\ b_h \\ 0 \\ \vdots \\ 0
\end{pmatrix}
\qquad (3)
$$

Es decir,

$$\bar{A}\, x^o = \bar{b} \qquad (4)$$

De la segunda condición de Kuhn y Tucker, para las restricciones no saturadas en x^o, se sigue que:

$$\begin{aligned} \lambda^{o}_{h+1} &= 0,\ldots\ldots, \lambda^{o}_{m} = 0 \\ \mu^{o}_{k+1} &= 0,\ldots\ldots, \mu^{o}_{n} = 0 \end{aligned} \qquad (5)$$

La primera condición de Kuhn y Tucker, queda:

$$2\,C\,x^{o} = A^{T}\,\lambda^{o} + \mu^{o} - p$$

$$2\,C\,x^{o} = \begin{pmatrix} a_{11}\cdots\cdots a_{m1} \\ \cdots\cdots\cdots \\ a_{1n}\cdots\cdots a_{mn} \end{pmatrix} \begin{pmatrix} \lambda^{o}_{1} \\ \vdots \\ \lambda^{o}_{m} \end{pmatrix} + \begin{pmatrix} \mu^{o}_{1} \\ \vdots \\ \mu^{o}_{n} \end{pmatrix} - \begin{pmatrix} p_{1} \\ \vdots \\ p_{n} \end{pmatrix}$$

Y teniendo en cuenta las expresiones (3) y (5):

$$2\,C\,x^{o} = \begin{pmatrix} a_{11}\cdots a_{h1} \\ \cdots\cdots \\ a_{1n}\cdots a_{hn} \end{pmatrix} \begin{pmatrix} \lambda^{o}_{1} \\ \vdots \\ \lambda^{o}_{h} \end{pmatrix} + \begin{pmatrix} \mu^{o}_{1} \\ \vdots \\ \mu^{o}_{k} \end{pmatrix} - \begin{pmatrix} p_{1} \\ \vdots \\ p_{n} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}\cdots a_{h1} & 1\cdots\ 0 \\ \cdots\cdots\cdots & \cdots\cdots \\ a_{1k}\cdots a_{hk} & 0\cdots\ 1 \\ \cdots\cdots\cdots & \cdots\cdots \\ a_{1n}\cdots a_{hn} & 0\cdots\ 0 \end{pmatrix} \begin{pmatrix} \lambda^{o}_{1} \\ \vdots \\ \lambda^{o}_{h} \\ \mu^{o}_{1} \\ \vdots \\ \mu^{o}_{k} \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} p_{1} \\ \vdots \\ \vdots \\ \vdots \\ p_{n} \end{pmatrix}$$

Y simplificando:

$$2\,C\,x^{o} = \bar{A}^{T}\,\bar{\lambda}^{o} - p \qquad (6)$$

Pero en (6), por ser C matriz simétrica definida positiva, se sigue que:

$$\det(C) \neq 0 \iff \exists\, C^{-1}$$

Luego,

$$x^o = \frac{1}{2} C^{-1} (\bar{A}^T \bar{\lambda}^o - p) \qquad (7)$$

Sustituyendo (7) en (4):

$$\frac{1}{2} \bar{A} C^{-1} (A^{-T} \bar{\lambda}^o - p) = \bar{b}$$

Operando, se obtiene:

$$\bar{A} C^{-1} \bar{A}^T \bar{\lambda}^o = 2 \bar{b} + \bar{A} C^{-1} p \qquad (8)$$

Haciendo $S = \bar{A} C^{-1} \bar{A}^T$ y $T = 2 \bar{b} + \bar{A} C^{-1} p$ y sustituyendo en (8), nos queda:

$$S \bar{\lambda}^o = T \qquad (9)$$

Pero (9) es un sistema lineal en $\bar{\lambda}^o$ que tiene solución única (por ser linealmente independientes los vectores gradientes de las restricciones saturadas en x^o), luego por el teorema de Rouché-Fröbenius, se sigue que:

$$rg(S) = h+k \iff \det(S) \neq 0 \iff \exists\, S^{-1}$$

Por tanto, de (9) concluímos:

$$\bar{\lambda}^o = S^{-1} T \qquad (10)$$

Ahora bien, por el teorema de Kuhn y Tucker es $\bar{\lambda}^o \geq 0$. Luego, en (10):

$$S^{-1} T \geq 0$$

Queda así probada la segunda hipótesis del teorema.

$\Longleftarrow$ Recíprocamente, probemos que las hipótesis primera y segunda implican las condiciones de Kuhn y Tucker en x^o.

. Tanto de la hipótesis previa del teorema como de la primera hipótesis, se sigue que $x^o \in Fr(Y)$.

Supongamos que las restricciones saturadas linealmente independientes en el punto x^o son:

$$\left.\begin{array}{l} a_{11}\, x_1 + \cdots\cdots + a_{1n}\, x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{h1}\, x_1 + \cdots\cdots + a_{hn}\, x_n = b_h \\ \quad x_1 \qquad\qquad\qquad\quad = 0 \\ \qquad \ddots \\ \qquad\quad x_k \qquad\qquad\quad = 0 \end{array}\right\} \quad (11)$$

siendo $h \leq m$, $k \leq n$ y por suponer que son linealmente independientes tales restricciones saturadas, $h+k \leq n$.

Llamando H_i $(i = 1,\ldots, h)$ y C_j $(j = 1,\ldots, k)$ a los hiperplanos de R^n determinados, respectivamente, por

$$\sum_{d=1}^{n} a_{id}\, x_d = b_i \qquad i = 1,\ldots\ldots, h$$

$$x_j = 0 \qquad j = 1,\ldots\ldots, k$$

se obtiene que la variedad afín que determinan las ecuaciones (11) es:

$$D = H_1 \cap \ldots\ldots\ldots \cap H_k \cap C_1 \cap \ldots\ldots \cap C_k$$

cuya dimensión, por suponer las ecuaciones (11) linealmente independientes, es $n-(h+k)$.

Pueden darse dos posibilidades:

a) $n-(h+k) > 0 \Leftrightarrow D$ no se reduce al punto x^0

b) $n-(h+k) = 0 \Leftrightarrow D$ se reduce al punto x^0

Sean:

$$a_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \quad \text{y} \quad e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \to (j) \\ \\ \\ \end{matrix}$$

los vectores asociados a los hiperplanos

$H_i \qquad i = 1,\ldots, h$

$C_j \qquad j = 1,\ldots,k$

Como $D = H_1 \cap \ldots \cap H_h \cap C_1 \cap \ldots\ldots \cap C_k \Rightarrow$

$$\Rightarrow \begin{cases} a_i \text{ ortogonal a } D & i = 1,\ldots, h \\ e_j \text{ ortogonal a } D & j = 1,\ldots, k \end{cases}$$

Luego,

D es ortogonal a $< a_1,\ldots, a_h, e_1,\ldots, e_k >$

Si ocurre la posibilidad a) $(\dim D = n - (h+k) > 0)$, podemos expresar R^n como la siguiente suma directa:

$$R^n = D' \oplus < a_1,\ldots, a_h, e_1,\ldots, e_k >$$

siendo D' el subespacio vectorial subyacente de D.

Si consideramos el hiperplano tangente H en x^o a la hipersuperficie de nivel $S = \{x \in R^n / Q(x^o) = Q(x)\}$, se verifica que:

$$D \subset H \qquad (13) \qquad (*)$$

Por otra parte,

$$Q(x^o) \text{ es ortogonal a } H \qquad (14)$$

De (13) y (14), se concluye:

$$\nabla Q(x^o) \text{ es ortogonal a } D \underset{\text{por } (12)}{\Longleftrightarrow}$$

$$\nabla Q(x^o) \in \langle a_1, \ldots, a_h, e_1, \ldots, e_k \rangle$$

Si $\dim(D) = n - (h+k) = 0 \Longleftrightarrow D = \{x^o\}$

Y, en este caso, $\{a_1, \ldots, a_h, e_1, \ldots, e_k\}$ es una base de R^n (ya que hemos visto anteriormente que tales vectores son linealmente independientes). Por tanto, cualquier vector de R^n se puede expresar como combinación lineal de la misma, en particular $\nabla Q(x^o)$. Luego,

$$\nabla Q(x^o) \in \langle a_1, \ldots, a_h, e_1, \ldots, e_k \rangle$$

Por tanto, cualquiera que sea la dimensión de D, es:

(*) Dicha propiedad, que no es obvia, se demuestra de manera análoga a lo realizado en la proposición 1.6.1. (Nota 1.6.1.) para el programa (PQ_1).

$$\nabla Q(x^o) = 2\,C\,x^o + p = \lambda_1^o\, a_1 + \dots + \lambda_h^o\, a_h + \mu_1^o\, e_1 + \dots + \mu_k^o\, e_k =$$

$$= \lambda_1^o \begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix} + \dots + \lambda_h^o \begin{pmatrix} a_{h1} \\ \vdots \\ a_{hn} \end{pmatrix} + \mu_1^o \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + \mu_k^o \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow (k) =$$

$$= \begin{pmatrix} a_{11} \dots\dots\dots a_{h1} \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ a_{1n} \dots\dots\dots a_{hn} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \cdot \\ \cdot \\ \cdot \\ \lambda_h^o \end{pmatrix} + \begin{pmatrix} \mu_1^o \\ \vdots \\ \mu_k^o \\ \vdots \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} \dots a_{h1}\; a_{h+1,1} \dots a_{m1} \\ \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots \\ a_{1n} \dots a_{hn}\; a_{h+1,n} \dots a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \vdots \\ \lambda_h^o \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \mu_1^o \\ \vdots \\ \mu_k^o \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

siendo λ^o y μ^o únicos, por ser únicas las coordenadas del vector $\nabla Q(x^o)$ respecto de la base

$\{a_1, \dots, a_h, e_1, \dots, e_k\}$

Es decir:

$$2\,C\,x^o + p = A^T\ \lambda^o + \mu^o \qquad (15)$$

que es la primera condición de Kuhn y Tucker si demostramos que $\lambda^o \geqslant 0$ y $\mu^o \geqslant 0$.

Por otra parte, para las restricciones saturadas en

x^o (11), se verifica que:

$$\left.\begin{array}{ll} \lambda_i^o \,(- A_{i*}\, x^o + b_i) = 0 & i = 1,.., h \\ \mu_j^o \; x_j^o = 0 & j = 1,.., k \end{array}\right\} \quad (16)$$

Y para las restricciones no saturadas en x^o, se cumple que:

$$\left.\begin{array}{ll} \lambda_r^o \,(- A_{r*}\, x^o + b_r) = 0 & r = h+1,..., m \\ \mu_t^o \; x_t^o = 0 & t = k+1,...,n \end{array}\right\} \quad (17)$$

por ser $\lambda_r^o = 0 \qquad r = h+1,..., m$ y

$\mu_t^o = 0 \qquad t = k+1,..., n$

Pero las ecuaciones (16) y (17), son la segunda condición de Kuhn y Tucker si probamos que $\lambda^o \geqslant 0$ y $\mu^o \geqslant 0$.

Demostremos que $\lambda^o \geqslant 0$ y que $\mu^o \geqslant 0$.

Evidentemente, para llegar a la expresión (15), es

$\lambda_r^o = 0 \qquad r = h+1,...., m$

$\mu_t^o = 0 \qquad t = k+1,...., n$

Por tanto, sólo nos falta probar que $\lambda_i^o \geqslant 0$ $i = 1,..., h$ y que $\mu_j^o \geqslant 0 \quad j = 1,...., k$

La igualdad (15), puede expresarse así:

$$2\,C\,x^{o} = \begin{pmatrix} a_{11} \cdots a_{h1} & 1 \cdots 0 \\ \cdots\cdots & \cdots\cdots \\ a_{1k} \cdots a_{hk} & 0 \cdots 1 \\ \cdots\cdots & \cdots\cdots \\ a_{1n} \cdots a_{hn} & 0 \cdots 0 \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \vdots \\ \lambda_h^o \\ \mu_1^o \\ \vdots \\ \mu_k^o \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} p_1 \\ \vdots \\ \vdots \\ p_n \end{pmatrix}$$

Y resumiendo:

$$2\,C\,x^{o} = \bar{A}^{T}\ \bar{\lambda}^{o} - p \qquad (18)$$

siendo $\bar{A}^{T} \in M(n,h+k)$, $\bar{\lambda}^{o} \in M(h+k,1)$ y $p \in M(n,1)$.

Por otra parte, las ecuaciones (11) pueden agruparse así:

$$\begin{pmatrix} a_{11} \cdots a_{1k} \cdots a_{1n} \\ \cdots\cdots\cdots \\ a_{h1} \cdots a_{hk} \cdots a_{hn} \\ 1 \cdots\cdots 0 \cdots\cdots 0 \\ \cdots\cdots\cdots \\ 0 \cdots\cdots 1 \cdots\cdots 0 \end{pmatrix} \begin{pmatrix} x_1^o \\ \vdots \\ \vdots \\ x_n^o \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_h \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Es decir,

$$\bar{A}\ x^{o} = \bar{b} \qquad (19)$$

siendo $\bar{A} \in M(h+k,n)$ y $\bar{b} \in M(h+k,1)$.

Vamos a determinar $\bar{\lambda}^{o}$ para lo cual despejaremos

x^o en (18) y lo sustituiremos en (19).

En (18), como C es matriz simétrica definida positiva, se sigue que:

$$\det(C) \neq 0 \iff \exists\, C^{-1}.$$

Luego:

$$x^o = \frac{1}{2} C^{-1} (\bar{A}^T \ \bar{\lambda}^o - p)$$

Sustituyendo tal valor en (19):

$$\frac{1}{2} \bar{A} \ C^{-1} (\bar{A}^T \ \bar{\lambda}^o - p) = \bar{b}$$

Por tanto:

$$\bar{A} \ C^{-1} \ \bar{A}^T \ \bar{\lambda}^o = 2 \ \bar{b} + \bar{A} \ C^{-1} \ p \qquad (20)$$

Haciendo $\bar{A} \ C^{-1} \ \bar{A}^T = S$ y $2 \ \bar{b} + \bar{A} \ C^{-1} \ p = T$, nos queda en (20):

$$S \ \bar{\lambda}^o = T \qquad (21)$$

Como tal sistema lineal tiene solución única $\bar{\lambda}^o$ (según se indicó al deducir la igualdad (15)), se concluye que:

$$\det(S) \neq 0 \iff \exists\, S^{-1}$$

Por tanto,

$$\bar{\lambda}^o = S^{-1} \ T$$

Ahora bien, por la segunda hipótesis, es:

$$S^{-1} \ T \geq 0$$

lo que implica;

$$\bar{\lambda}^o \geq 0 \iff \lambda^o \geq 0 \quad y \quad \mu^o \geq 0$$

Luego, λ^o y μ^o verifican el teorema de Kuhn y Tucker.

Queda, por tanto, probada la equivalencia.

Corolario 1.6.2.

Dado el programa:

$$(PQ_5) \begin{cases} \text{Min}\{Q(x) = x^T\,C\,x + p^T\,x\} \\ \text{bajo las restricciones:} \\ \quad Y \quad \begin{cases} A\,x \geq b \\ x \geq 0 \\ x \in R^n \end{cases} \end{cases}$$

siendo C matriz simétrica definida positiva, $C \in M(n,n)$, $p \in M(n,1)$, $A \in M(m,n)$, $rg(A) = m$ y $b \in M(m,1)$, todas ellas matrices de números reales.

Si x^o verifica que:

1ª) $x^o \geq 0$

2ª) $\{x^o\} = \{x \in R^n \ / \ Q(x^o) = Q(x)\} \cap D$

siendo D la variedad afín solución del sistema $A\,x = b$

3ª) $(A\,C^{-1}\,A^T)^{-1}\,(2\,b + A\,C^{-1}\,p) \geq 0$

Entonces, x^o es solución óptima de (PQ_5).

Demostración

Es consecuencia del teorema 1.6.3..

En efecto, la primera y segunda condición de este corolario, implican que $x^o \in Fr(Y)$.

La segunda condición, es un caso particular de la hipótesis primera del teorema 1.6.3., cuando las restricciones saturadas linealmente independientes en x^o están dadas por $A\,x = b$.

Y la tercera condición corresponde a la segunda hipótesis del teorema 1.6.3., para este caso particular de restricciones saturadas en x^o.

C A P I T U L O 2

OPTIMALIDAD EN LA FRONTERA DE LAS RESTRICCIONES EN PROGRAMACION POLINOMICA

2.1. INTRODUCCION

En este Capítulo, se establecen resultados originales sobre Programación Polinómica. Dichas aportaciones se refieren a:

1) Optimalidad en la frontera de las restricciones

2) Algoritmo para la obtención del óptimo

Respecto del primer punto, se indican en el epígrafe 2.2. las particularizaciones que pueden realizarse de los apartados 1.2. á 1.5., ambos incluídos, del Capítulo 1. En la sección 2.3., se formulan nuevos resultados de optimalidad en la frontera para programas polinómicos con función objetivo de coeficientes positivos y conjunto de restricciones compacto en R_+^n . El resultado que establece la optimalidad en la frontera en tales programas, se hace de dos formas distintas: una mediante los conceptos de hipercubos superior e inferior n-dimensionales, la otra mediante una demostración directa. El primer método es más preciso, pues además de establecer que el óptimo está en la frontera nos indica qué condición adicional ha de verificar la región de la misma en la que se encuentra el óptimo. Se muestra, mediante un contraejemplo, que tales resultados no pueden extenderse a la Programación Posinomial.

Con relación al segundo punto, se establece (apartado 2.4.) un método para la obtención del óptimo en programas polinómicos con restricciones lineales y función polinómica objetivo convexa en la región de las restricciones. El citado método extiende y desarrolla una idea formulada por Beale, E.M.L. [7] para el caso cuadrático.

El algoritmo que establecemos consiste en una modificación esencial de la metodología del simplex. Pasamos a exponer el fundamento del algoritmo que formulamos.

Como se sabe, el algoritmo del simplex,en un problema de mínimo, consiste en determinar la solución óptima (que está en al menos un vértice del politopo de las restricciones) de un programa lineal mediante un proceso iterativo que permite pasar, en cada iteración, de un vértice a otro adyacente, tomando como nueva variable básica la de coste reducido negativo más pequeño hasta llegar al óptimo, que se alcanza cuando los costes reducidos de las variables libres son mayores o iguales que cero. En la Programación Polinómica con restricciones lineales, el óptimo no tiene por qué encontrarse en un vértice; el algoritmo formulado por nosotros permite, para estos programas, avanzar por vértices adyacentes, mientras sea posible, tomando como dirección de avance, en cada iteración, la dada por la variable libre para la cual la derivada de la función objetivo respecto de la misma sea la más negativa en el vértice de partida. Cuando el punto al que se llega no es vértice, se transforma en tal mediante la ampliación del programa y se reitera, para este nuevo programa, la metodología anterior.

La ampliación del programa se realiza añadiendo una nueva ecuación lineal con una nueva variable sin restricción en el signo. La ecuación añadida ha de cumplir dos condiciones esenciales para que el punto no vértice se transforme en vértice del nuevo programa.

En el subapartado 2.4.1., se expone con todo detalle el algoritmo.

En la subsección 2.4.2., se formulan las propiedades inherentes al citado método. Se establecen, en primer lugar, los teoremas que relacionan las soluciones de los sucesivos programas: soluciones factibles, soluciones básicas factibles, soluciones básicas factibles no degeneradas y soluciones óptimas (teoremas 2.4.2.1. y corolarios correspondientes, así como los teoremas 2.4.2.2., 2.4.2.3. y corolario 2.4.2.4.). Se estudia, seguidamente, la acotación de las soluciones factibles (teorema 2.4.2.4.), la caracterización del óptimo (teorema 2.4.2.5.) y la convergencia del algoritmo (teorema 2.4.2.6.).

En la subsección 2.4.3., se indica la propiedad adicional del algoritmo para el caso cuadrático y en la 2.4.4. se resuelve un ejemplo mediante la aplicación del citado algoritmo.

En el apartado 2.5., finalmente, se indica en qué sentido se puede aplicar el algoritmo a la Programación Polinómica cuando la función objetivo no es convexa en el conjunto que definen las restricciones.

2.2. CONSIDERACIONES SOBRE OPTIMALIDAD EN LA FRONTERA Y UNICIDAD DEL OPTIMO

A todos los programas polinómicos que verifican las hipótesis formuladas en el Capítulo 1, apartados 1.2. á 1.5. ambos incluídos, les podemos aplicar los teoremas que se establecen en dichos apartados.

2.3. OPTIMALIDAD EN LA FRONTERA EN ALGUNOS PROGRAMAS POLINOMICOS

Vamos a establecer nuevos resultados sobre optimalidad en la frontera de las restricciones para programas con función objetivo polinómica de coeficientes positivos y conjunto de restricciones Y compacto, con $Y \subset R_+^n$ (*), en la doble vertiente de máximo y de mínimo; particularizamos, además, tales resultados a programas con restricciones lineales.

Necesitamos, para ello, una definición previa.

Definición 2.3.1.

$\forall x^* = (x_1^*, \ldots, x_n^*) \in R^n$, llamaremos hipercubo superior de x^* y lo notaremos por $H_s^{x^*}$ al siguiente conjunto:

$$H_s^{x^*} = \left\{ (x_1, \ldots, x_n) \in R^n \;/\; x_i \geq x_i^* \quad i \in \{1, \ldots, n\} \text{ y } \exists\, j \in \{1, \ldots, n\} \text{ tal que } x_j > x_j^* \right\}$$

(*) Notación: $R_+^n = \left\{ x \in R^n \;/\; x \geq 0 \right\}$

Análogamente, llamaremos hipercubo inferior de $x^* \in R^n$ y lo denotaremos por $H_I^{x^*}$ al siguiente conjunto:

$$H_I^{x^*} = \left\{ (x_1,\ldots,x_n) \in R^n \;/\; x_i \leq x_i^* \quad \forall i \in \{1,\ldots,n\} \text{ y } \exists j \in \{1,\ldots,n\} \text{ tal que } x_j < x_j^* \right\}$$

Obviamente,

$$\mathring{H}_s^{x^*} = \left\{ (x_1,\ldots,x_n) \in R^n \;/\; x_i > x_i^* \quad \forall i \in \{1,\ldots,n\} \right\}$$

$$\mathring{H}_I^{x^*} = \left\{ (x_1,\ldots,x_n) \in R^n \;/\; x_i < x_i^* \quad \forall i \in \{1,\ldots,n\} \right\}$$

Teorema 2.3.1.

Dado el programa:

$$(PP_1) \begin{cases} \text{Max } P(x) \\ \text{bajo las restricciones:} \\ x \in Y \subset R_+^n \end{cases}$$

siendo $P(x)$ función polinómica de coeficientes positivos e Y conjunto compacto en R^n para la topología usual. El máximo del programa (PP_1) se alcanza en un punto de Y cuyo hipercubo superior no contiene en su interior punto alguno de Y.

Demostración

La haremos por reducción al absurdo.

Sea $x^* \in Y$ un punto de máximo de (PP_1) tal que:

$$\mathring{H}_s^{x^*} \cap Y \neq \emptyset$$

Ello implica que $\exists w \in \mathring{H}_s^{x^*} \cap Y$ cuyas coordenadas, por ser $w \in \mathring{H}_s^{x^*}$, serán todas estrictamente mayores que las del punto x^* y por ser $P(x)$ de coeficientes positivos e $Y \subset R_+^n$, concluímos que $P(x^*) < P(w)$. Contradicción con la hipótesis de que x^* es un punto de máximo de (PP_1). La contradicción se origina al suponer que $\mathring{H}_s^{x^*} \cap Y \neq \emptyset$. Luego, si x^* es un punto de máximo de (PP_1), entonces $\mathring{H}_s^{x^*} \cap Y = \emptyset$.

Queda probado el teorema.

<u>Corolario 2.3.1.</u>

Dado el programa:

$$(PP_2) \begin{cases} \text{Max } P(x) \\ \text{bajo las restricciones:} \\ Y \begin{cases} A\,x \leqslant b \\ x \in R_+^n \end{cases} \end{cases}$$

siendo $P(x)$ función polinómica de coeficientes positivos, Y poliedro convexo(*), $A \in M(m,n)$, $b \in M(m,1)$ y ambas matrices en R. Si x^* es un máximo de (PP_2), entonces $\mathring{H}_s^{x^*} \cap Y = \emptyset$.

(*) Siguiendo a Minoux, M. [40], damos la siguiente

<u>Definición</u>: Un conjunto convexo de la forma $Y = \left\{ x \in R^n \,/\, A\,x \leqslant b,\ x \geqslant 0 \right\}$ se llama politopo, siendo $A \in M(m,n)$, $b \in M(m,1)$ y ambas matrices en R.

Un politopo acotado no vacío se denomina poliedro convexo.

Demostración

Es una consecuencia del teorema 2.3.1., pues al ser Y poliedro convexo, se sigue que Y es un conjunto compacto para la topología usual de R^n y no vacío.

Teorema 2.3.2.

Dado el programa:

$$(PP_3) \begin{cases} \text{Min } P(x) \\ \text{bajo las restricciones:} \\ x \in Y \subset R^n_+ \end{cases}$$

siendo $P(x)$ función polinómica de coeficientes positivos e Y conjunto compacto para la topología usual de R^n. El mínimo del programa (PP_3) se alcanza en un punto de Y cuyo hipercubo inferior no contiene en su interior punto alguno de Y.

Demostración

Sea $x^* \in Y$ mínimo del programa (PP_3) tal que $\mathring{H}_I^{x^*} \cap Y \neq \emptyset$.

Ello implica que $\exists\, w \in \mathring{H}_I^{x^*} \cap Y$ cuyas coordenadas, por ser $w \in \mathring{H}_I^{x^*}$, serán todas estríctamente menores que las correspondientes de x^*, y por ser $P(x)$

de coeficientes positivos e $Y \subset R^n_+$, concluímos que $P(w) < P(x^*)$. Contradicción con la hipótesis de ser x^* un mínimo del programa (PP_3). Esta contradicción, tiene su origen en suponer que $\overset{\circ}{H}{}^{x^*}_I \cap Y \neq \emptyset$. Luego, $\forall x^*$ punto de mínimo de (PP_3) se verifica que $\overset{\circ}{H}{}^{x^*}_I \cap Y = \emptyset$.

Concluye así, la demostración del teorema 2.3.2.

Corolario 2.3.2.

Dado el programa:

$$(PP_4) \begin{cases} \text{Min } P(x) \\ \text{bajo las restricciones:} \\ Y \begin{cases} A\,x \leq b \\ x \in R^n_- \end{cases} \end{cases}$$

siendo $P(x)$ función polinómica de coeficientes positivos, Y poliedro convexo, $A \in M(m,n)$, $b \in M(m,1)$ y ambas matrices en R. Si x^* es un mínimo de (PP_4), entonces $\overset{\circ}{H}{}^{x^*}_I \cap Y = \emptyset$.

Demostración

Es consecuencia del teorema 2.3.2., pues al ser Y poliedro convexo es un compacto para la topología usual de R^n y no vacío.

Basándonos en los teoremas 2.3.1. y 2.3.2., establecemos, también, el siguiente:

Teorema 2.3.3.

Dado el programa:

$$(PP_5)\begin{cases} \text{Optimizar (Max. ó Min.)} \quad P(x) \\ \text{bajo las restricciones:} \\ x \in Y \subset R^n_+ \end{cases}$$

siendo $P(x)$ función polinómica de coeficientes positivos e Y conjunto compacto en R^n con la topología usual. Si $x^* \in Y$ es una solución óptima del programa (PP_5), entonces $x^* \in Fr(Y)$.

Demostración

Lo probaremos, en primer lugar, para el máximo.

Pueden ocurrir dos posibilidades:

a) $Int(Y) = \emptyset$, entonces es obvio que $x^* \in Fr(Y)$, ya que $Y = Fr(Y)$.

b) $Int(Y) \neq \emptyset$

Razonaremos por reducción al absurdo.

Si $\exists x^*$ máximo del programa (PP_5) tal que $x^* \notin Fr(Y)$, entonces se verificará que:

$$d(x^*, Fr(Y)) = \varepsilon > 0$$

Ello implica que:

$\exists\, \mathring{B}(x^*;\ \delta) \subset \mathrm{Int}(Y)$

De donde se sigue que:

$\exists\, \bar{x} \in \mathring{B}(x^*;\ \delta)$ tal que $\bar{x}_i > x_i^* \quad \forall i \in \{1,\dots,n\}$

(bastaría tomar, por ejemplo, $\bar{x} = x^* + \left(\frac{\delta_1}{2},\dots,\frac{\delta_n}{2}\right)$ siendo $\sum_{i=1}^{n} \delta_i^2 = \delta^2$ y $\delta_i > 0 \quad \forall i \in \{1,\dots,n\}$)

Entonces, por la definición 2.3.1., se sigue que:

$$\bar{x} \in \mathring{H}_s^{x^*} \qquad (1)$$

Como,

$$\bar{x} \in \mathrm{Int}(Y) \subset Y \qquad (2)$$

concluímos de (1) y (2) que:

$$\bar{x} \in \mathring{H}_s^{x^*} \cap Y \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \qquad \mathring{H}_s^{x^*} \cap Y \neq \emptyset$$

Por el teorema 2.3.1., se concluye que x^* no es máximo del programa (PP_5). Contradicción, que tiene su origen en suponer que $\exists\, x^*$ máximo de (PP_5) tal que $x^* \notin \mathrm{Fr}(Y)$. Luego, $\forall x^*$ máximo del programa (PP_5) se verifica que $x^* \in \mathrm{Fr}(Y)$.

Análogamente razonaremos para el mínimo.

a) Si $\mathrm{Int}(Y) = \emptyset$, entonces obviamente $x^* \in \mathrm{Fr}(Y) = Y$.

b) Si $\mathrm{Int}(Y) \neq \emptyset$, entonces si $\exists\, x^*$ mínimo de (PP_5) tal que $x^* \notin \mathrm{Fr}(Y)$, se verificará que:

$$d(x^*; Fr(Y)) = \delta > 0$$

Ello implica que:

$$\exists\, \mathring{B}(x^*; \delta) \subset Int(Y)$$

De donde se sigue que:

$$\exists\, \bar{x} \in \mathring{B}(x^*; \delta) \quad \text{tal que} \quad \bar{x}_i < x_i^* \qquad \forall i \in \{1, \ldots, n\}$$

(tomaríamos, por ejemplo, $\bar{x} = x^* - \left(\frac{\delta_1}{2}, \ldots, \frac{\delta_n}{2}\right)$

siendo $\delta_1^2 + \ldots + \delta_n^2 = \delta^2$ y $\delta_i > 0 \quad \forall i \in \{1, \ldots, n\}$)

Entonces, por la definición 2.3.1., se sigue que:

$$\bar{x} \in \mathring{H}_I^{x^*} \qquad (3)$$

Por otra parte,

$$\bar{x} \in Int(Y) \subset Y \qquad (4)$$

De (3) y (4) se concluye que:

$$\bar{x} \in \mathring{H}_I^{x^*} \cap Y \quad \Longleftrightarrow$$

$$\Longleftrightarrow \quad \mathring{H}_I^{x^*} \cap Y \neq \emptyset$$

De donde concluímos, por el teorema 2.3.2., que x^* no es mínimo del programa (PP_5). **Contradicción que se** origina al suponer que $\exists\, x^*$ mínimo de (PP_5) **tal que** $x^* \notin Fr(Y)$. Luego, $\forall x^*$ mínimo de (PP_5) **se verifica** que $x^* \in Fr(Y)$.

Finaliza, así, la demostración **del teorema** 2.3.3.

Nota 2.3.1.

Los teoremas 2.3.1., 2.3.2. y 2.3.3., además de establecer la optimalidad en la frontera nos proporcionan una información adicional acerca de la región de la misma en la que se encuentra el óptimo.

Una demostración directa del teorema 2.3.3. es la siguiente:

Por reducción al absurdo.

La haremos para el caso de máximo.

Sea x^* una solución máxima de (PP_5) tal que $x^* \notin Fr(Y)$. Ello implicaría que $x^* \in Int(Y)$, de donde se sigue que: $\exists\, \mathring{B}(x^*, r) \subset Int(Y)$. Por tanto,

$\exists\, \bar{x} = x^* + \left(\frac{r_1}{2}, \ldots, \frac{r_n}{2} \right) \in \mathring{B}(x^*, r)$ (siendo

$\sum_{i=1}^{n} r_i^2 = r^2$ y $r_i > 0$ $\forall i = 1,\ldots,n$). Luego, $\exists\, \bar{x} \in Int(Y)$ cuyas coordenadas son estrictamente mayores que las correspondientes de x^*; por tanto, al ser $P(x)$ de coeficientes positivos e $Y \subset R_+^n$, se concluye que $P(\bar{x}) > P(x^*)$. Ello implicaría que x^* no es solución máxima de (PP_5). Contradicción que se origina al suponer que $x^* \notin Fr(Y)$. Luego, si x^* es una solución máxima de (PP_5), entonces $x^* \in Fr(Y)$

Análogamente razonaríamos para el programa de mínimo.

Termina así, esta demostración del teorema 2.3.3.

Corolario 2.3.3.

Dado el programa:

$$(PP_6)\begin{cases}\text{Optimizar (Max ó Min) } P(x)\,,\\ \text{bajo las restricciones:}\\ Y\begin{cases} A\,x \leqslant b\\ x \in R^n_+\end{cases}\end{cases}$$

siendo $P(x)$ función polinómica de coeficientes positivos, Y poliedro convexo, $A \in M(m,n)$, $b \in M(m,1)$ y ambas matrices en R. Si x^* es un óptimo de (PP_6), entonces $x^* \in Fr(Y)$.

Demostración

Es una consecuencia del teorema 2.3.3., ya que por ser Y poliedro convexo es un conjunto no vacío y compacto en R^n para la topología usual.

Nota 2.3.2.

Los resultados obtenidos en este epígrafe 2.3. para la Programación Polinómica Positiva no se pueden generalizar a la Programación Posinomial(*), como nos muestra el siguiente contraejemplo:

(*) Véase la definición de programa posinomial en la página siguiente.

Sea el programa:

$$\text{Min } P(x_1,x_2) = x_1^2 + x_2^2 + \frac{2}{x_1 x_2}$$

sujeto a las restricciones:

$$Y \begin{cases} \frac{1}{2} \leq x_1 \leq 2 \\ \frac{1}{2} \leq x_2 \leq 2 \end{cases}$$

La función $P(x_1,x_2)$ es convexa en Y.

En efecto, la matriz hessiana de $P(x_1,x_2)$ es:

$$H = \begin{pmatrix} P''_{x_1^2} & P''_{x_1 x_2} \\ P''_{x_2 x_1} & P''_{x_2^2} \end{pmatrix} = \begin{pmatrix} \dfrac{2x_1^3 x_2 + 4}{x_1^3 x_2} & \dfrac{2}{x_1^2 x_2^2} \\ \dfrac{2}{x_1^2 x_2^2} & \dfrac{2x_1 x_2^3 + 4}{x_1 x_2^3} \end{pmatrix}$$

Vemos que:

$$D_1 = \frac{2x_1^3 x_2 + 4}{x_1^3 x_2} > 0$$

$$D_2 = |H| = \frac{12 + 4x_1^4 x_2^4 + 8(x_1 x_2^3 + x_1^3 x_2)}{x_1^4 x_2^4} > 0$$

ya que $x_1 > 0$, $x_2 > 0$.

(*) <u>Definición</u>: Una función $F(x) = U_1(x) + \ldots + U_p(x)$ con $U_i(x) = c_i x_1^{a_{1i}} x_2^{a_{2i}} \ldots x_n^{a_{ni}}$, siendo $c_i \geq 0$, $a_{ji} \in R_+$ $\forall j = 1,\ldots,n$ $\forall i = 1,\ldots,p$, $x_j \geq 0$ $\forall j = 1,\ldots,n$ se denomina posinomial.

Un problema de Programación Matemática en el que tanto la función objetivo como las restricciones son posinomios se denomina programa posinomial.

Luego, H es una matriz simétrica definida positiva en un abierto convexo que contiene a Y, ello es equivalente a que $P(x_1,x_2)$ es convexa en dicho abierto y, por tanto, es convexa en Y.

Calculemos, ahora, los extremos relativos de $P(x_1,x_2)$ sin restricciones:

$$\left.\begin{aligned} P'_{x_1} &= 2x_1 - \frac{2}{x_1^2\, x_2} = 0 \\ P'_{x_2} &= 2x_2 - \frac{2}{x_1\, x_2^2} = 0 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} 2x_1^3\, x_2 &= 2 \\ 2x_1\, x_2^3 &= 2 \end{aligned}\right\} \Rightarrow x_1 = \pm 1,\ x_2 = \pm 1$$

Luego, los posibles extremos relativos de $P(x_1,x_2)$, sin restricciones, tienen de coordenadas: $x_1 = \pm 1$, $x_2 = \pm 1$.

Veamos qué ocurre en $(1,1) \in Y$.

$$H(1,1) = \begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix} \quad \text{como } P''_{x_1^2} = 6 > 0 \ \text{ y } \ |H(1,1)| > 0$$

se sigue que (1,1) es un mínimo relativo, como $P(x_1,x_2)$, es convexa en Y se concluye que tal mínimo local es un mínimo global de $P(x_1,x_2)$ en Y. Por otra parte, $(1,1) \in \mathrm{Int}(Y)$. Es decir, el mínimo del programa se alcanza en un punto de Y que no pertenece a la frontera de Y.

2.4. ALGORITMO PARA LA OPTIMIZACION DE PROGRAMAS POLINOMICOS

En este epígrafe, se desarrolla un algoritmo para programas polinómicos cuando las restricciones son lineales y la función objetivo polinómica es convexa en el citado conjunto de restricciones.

2.4.1. Exposición del algoritmo

Sea el programa:

$$(PP^{(o)}) \begin{cases} \text{Min } P(x) \\ \text{bajo las restricciones:} \\ Y \begin{cases} A\,x = b \\ x \geq 0 \\ x \in R^n \end{cases} \end{cases}$$

donde $P(x)$ es función polinómica convexa en Y, $A \in M(m,n)$, $b \in M(m,1)$, $rg(A) = m$ y ambas matrices en R.

El algoritmo que vamos a formular, para este programa es una extensión con modificaciones esenciales del algoritmo del simplex.

Se parte de una solución básica factible (o vértice si lo consideramos desde el punto de vista geométrico), se elige la dirección de avance que viene dada, en cada iteración, por la variable libre para la cual la derivada de la función objetivo respecto de la misma, en el vértice de partida, sea la más negativa. Puede ocurrir que el punto de llegada sea vértice del programa inicial o que no lo sea. Si ocurre lo primero, se seguiría aplicando la metodología anterior para elegir la nueva dirección de avance. Si el punto de llegada no es vértice del programa dado, se transforma en solución básica factible de un programa ampliado, obtenido mediante la adición de una nueva ecuación lineal con una nueva variable sin restricción en el signo. Se precisa, más adelante, cuál es la ecuación que se añade.

El programa ampliado, cuya función objetivo es la misma y que consta de las restricciones del programa anterior más la ecuación añadida, está definido en un espacio euclídeo que tiene de dimensión una unidad más que el correspondiente al programa anterior. Se continúa aplicando a este nuevo programa la metodología anterior. Si ocurre en alguna iteración que el punto de llegada no es vértice del programa en cuestión, se amplía nuevamente el mencionado programa según el método anteriormente expuesto.

Pasos del algoritmo

1) Se elige una solución básica factible del primer programa $PP^{(o}$:

$$\{x^o_{b_1}, \ldots, x^o_{b_m}\}$$

2) Se expresan las variables básicas en función de las libres:

$$x_{b_i} = \sum_{j=1}^{n-m} a_{b_i f_j} \, x_{f_j} + x^o_{b_i} \qquad i = 1, \ldots, m$$

3) Se explicita la función objetivo a través de las variables libres únicamente:

$$P(x_1, \ldots, x_n) = P^{(o}(x_{f_1}, \ldots, x_{f_{n-m}})$$

4) Partiendo del punto $(x_{f_1},\dots,x_{f_{n-m}}) = (0,\dots,0)$, se elige la dirección, según una variable libre, que mejore la solución anterior. Para ello, se calcula:

$$\left(\frac{\partial P^{(o}}{\partial x_{f_j}}\right)_{(x_{f_1}=0,\dots,x_{f_{n-m}}=0)} \qquad j=1,\dots,n-m$$

Si $\left(\dfrac{\partial P^{(o}}{\partial x_{f_j}}\right)_{(0)} \geq 0 \qquad \forall j \in \{1,\dots,n-m\}$,

hemos llegado al óptimo, según demostraremos posteriormente, y se finaliza el proceso.

Si $\left(\dfrac{\partial P^{(o}}{\partial x_{f_j}}\right)_{(0)} < 0$ para algún $j \in \{1,\dots,n-m\}$,

la solución se puede mejorar. Y, en este caso, debemos elegir aquella variable x_{f_j} tal que $\left(\dfrac{\partial P^{(o}}{\partial x_{f_j}}\right)_{(0)}$ sea lo más negativo posible. Si hubiera más de una derivada parcial verificando la condición anterior, elegiríamos la variable libre de mayor índice.

5) Elegida la dirección x_{f_j}, que podemos suponer sin pérdida de generalidad que es x_{f_1}, hemos de determinar hasta qué valor $x_{f_1}^1$ podemos avanzar. Pueden ocurrir los siguientes casos:

a) Que alguna de las x_{b_i} se anule.

b) Que $\left(\dfrac{\partial P^{(o}}{\partial x_{f_1}} \right)_{(x_{f_1}, x_{f_2} = 0, \ldots, x_{f_{n-m}} = 0)} = 0$

en el punto $(x_{f_1}^1, 0, \ldots, 0)$.

c) Que se den las condiciones a) y b) simultáneamente.

6) Si se da el caso a), seguimos la metodología del simplex, entrando en la base x_{f_1} y saliendo de la misma la variable x_{b_i} que se anula. Si hubiera más de un x_{b_i} que se hiciera cero, elegiríamos la de menor índice. Podemos suponer que la variable que se anula es x_{b_1}, entonces las nuevas variables básicas serían $\{ x_{b_2}, \ldots, x_{b_m}, x_{f_1} \}$ con solución básica $\{x_{b_2}^1, \ldots, x_{b_m}^1, x_{f_1}^1 \}$. Las restricciones seguirían siendo las mismas y las variables básicas, expresadas en función de las libres, quedarían ahora:

$$x_{b_i} = \sum_{j=2}^{n-m} a'_{b_i f_j} x_{f_j} + a'_{b_i b_1} x_{b_1} + x_{b_i}^1 \qquad i = 2, \ldots, m$$

$$x_{f_1} = \sum_{j=2}^{n-m} a'_{f_1 f_j} x_{f_j} + a'_{f_1 b_1} x_{b_1} + x_{f_1}^1$$

Asimismo, se expresaría la función objetivo a través de las nuevas variables libres.

Y el algoritmo para este caso a), se repetiría nuevamente, sin que haya que ampliar el programa.

Si ocurre b), sería:

$$\left(\frac{\partial P^{(o}}{\partial x_{f_1}}\right)_{(x_{f_1}=x^1_{f_1},\, x_{f_2}=0,\ldots,\, x_{f_{n-m}}=0)} = 0$$

para cierto $x_{f_1} = x^1_{f_1}$.

En este caso, no se anulan las variables básicas iniciales $\{x_{b_1},\ldots,x_{b_m}\}$ y llegaríamos al punto $\{x^1_{b_1},\ldots,x^1_{b_m},x^1_{f_1}\}$ que mejora la primera solución básica y que sería solución factible, pero no básica, para las restricciones del primer programa. Para conseguir que tal punto se convierta en vértice hemos de pasar a un programa ampliado, para ello añadimos la siguiente ecuación lineal:

$$\nabla\left(\frac{\partial P^{(o}}{\partial x_{f_1}}\right)^{T}_{(x_{f_1}=x^1_{f_1},\, x_{f_2}=0,\ldots,\, x_{f_{n-m}}=0)} \begin{pmatrix} x_{f_1}-x^1_{f_1} \\ x_{f_2} \\ \vdots \\ x_{f_{n-m}} \end{pmatrix} = w_1$$

con w_1 sin restricción en el signo.

Dicha ecuación explicitada tiene la forma:

$$h_{f_1f_1}\left(x_{f_1} - x^1_{f_1}\right) + \sum_{j=2}^{n-m} h_{f_1f_j}\, x_{f_j} = w_1$$

Las restricciones del nuevo programa serían:

$$A\,x = b$$
$$h_{f_1f_1}\left(x_{f_1} - x^1_{f_1}\right) + \sum_{j=2}^{n-m} h_{f_1f_j}\, x_{f_j} = w_1$$
$$x \geq 0$$
$$(x,w_1) \in R^{n+1}$$

En este programa, de acuerdo con la solución a la que hemos llegado, es:

$$\operatorname{rg}\begin{pmatrix} & & A & & \\ 0\ldots.0 & h_{f_1f_1} & h_{f_1f_2} & \ldots. & h_{f_1f_{n-m}}\end{pmatrix} = m+1.$$

y por ser x_{f_1} la variable que entra en la base, se sigue que $h_{f_1f_1} \neq 0$.

Además, la ecuación añadida es un hiperplano que pasa por el punto $(x_{f_1}=x^1_{f_1}, x_{f_2}=0,\ldots,x_{f_{n-m}}=0, w_1=0)$.

El programa resultante, para este caso b), sería:

$$(PP^{(1)})\begin{cases} \text{Min. } P(x) \\ \text{bajo las restricciones:} \\ Y^{(1)}\begin{cases} A\,x = b \\ h_{f_1f_1}\left(x_{f_1}-x^1_{f_1}\right) + \displaystyle\sum_{j=2}^{n-m} h_{f_1f_j}\, x_{f_j} = w_1 \\ x \geq 0 \\ (x,w_1) \in R^{n+1} \end{cases} \end{cases}$$

Para este programa $(PP^{(1})$, el punto:

$$(x_{b_1}=x^1_{b_1},\ldots,x_{b_m}=x^1_{b_m},x_{f_1}=x^1_{f_1},x_{f_2}=0,\ldots,x_{f_{n-m}}=0,w_1=0)$$

constituye una solución básica factible.

A partir de este momento, y para este nuevo programa, se reitera el algoritmo, teniendo en cuenta la siguiente condición: ninguna variable w_j es introducida en la base hasta que ya no podamos seguir disminuyendo la función objetivo en dirección alguna x_i.

Si ocurre c), se procede como en el caso a) y se sigue aplicando la metodología del simplex al programa inicial.

7) El punto:

$$(x^k_{b_1}(k,\ldots,x^k_{b_p}(k,\overbrace{0,\ldots,0}^{n-p},w^k_{b_1}(k,\ldots,w^k_{b_h}(k,\overbrace{0,\ldots,0}^{k-h}) \in Y^{(k}$$

es el óptimo para el programa $(PP^{(k})$ si y sólo si:

$$\left(\frac{\partial P^{(k}}{\partial x_{f_j}(k}\right)_{(0)} \geq 0 \quad y \quad \left(\frac{\partial P^{(k}}{\partial w_{f_t}(k}\right)_{(0)} = 0 \quad \forall j \quad y \quad \forall t$$

Y el punto:

$$\left(x^k_{b_1}(k,\ldots,x^k_{b_p}(k,\ \overbrace{0,\ldots,0}^{n-p}\right) \in Y^{(o}$$

es el óptimo para el programa inicial.

Cuando se llega a un punto que verifica las condiciones anteriores, se finaliza el algoritmo.

2.4.2. Propiedades del algoritmo

2.4.2.1. Notaciones utilizadas

Los programas serán notados así: $(PP^{(o)})$ el programa de partida o inicial, $(PP^{(1)})$, $(PP^{(2)})$, etc. los programas posteriores según el orden natural. Análogamente, los sucesivos conjuntos de restricciones, se denotarán: $Y^{(o}$, $Y^{(1}$, etc.. Obviamente, $(PP^{(k)})$ es el programa en el que hemos introducido k ecuaciones lineales con k variables w_i. Además, si $Y^{(o} \subset R^n$, entonces $Y^{(k} \subset R^{n+k}$.

El programa $(PP^{(o)})$ será:

Min. P(x)

bajo las restricciones:

$$Y^{(o}\begin{cases} A\ x = b \\ x \geq 0 \\ x \in R^n \end{cases}$$

La última función objetivo del programa $(PP^{(o)})$, expresada a través de las variables libres, la denotaremos por $P^{(o}(x_{f_1},\ldots\ldots,x_{f_{n-m}})$.

El programa $(PP^{(1)})$, será:

$$
\begin{array}{l}
\text{Min. } P(x) \\
\text{bajo las restricciones:} \\
Y^{(1} \left\{
\begin{array}{l}
A\,x = b \\
\nabla\left(\dfrac{\partial P^{(o}}{\partial x_{f_1}}\right)^T_{(x_{f_1}=x^1_{f_1},\, x_{f_2}=0,\, \ldots,\, x_{f_{n-m}}=0\,)} \begin{pmatrix} x_{f_1}-x^1_{f_1} \\ x_{f_2} \\ \cdot \\ \cdot \\ x_{f_{n-m}} \end{pmatrix} = w_1 \\
x \geq 0 \\
(x, w_1) \in R^{n+1}
\end{array}
\right.
\end{array}
$$

Que podemos expresar, al involucrar a todas las variables x_i, así:

$$
(PP^{(1)}) \left\{
\begin{array}{l}
\text{Min. } P(x) \\
\text{bajo las restricciones:} \\
Y^{(1} \left\{
\begin{array}{l}
A\,x = b \\
H_1\,x = w_1 \\
x \geq 0 \\
(x, w_1) \in R^{n+1}
\end{array}
\right.
\end{array}
\right.
$$

donde $H_1 \in M(1,n)$ y los elementos de H_1 correspondientes a las variables básicas x_B son ceros.

En general, el programa $(PP^{(k)})$, $\forall k \in N^*$, se expresará:

$$
(PP^{(k)}\left\{
\begin{array}{l}
\text{Min. } P(x) \\
\text{bajo las restricciones:} \\
Y^{(k}\left\{
\begin{array}{l}
A\,x = b \\
H_1\,x = w_1 \\
H_2\begin{pmatrix} x \\ \cdots \\ w_1 \end{pmatrix} = w_2 \\
\cdots\cdots\cdots\cdots \\
H_k\begin{pmatrix} x \\ \cdots \\ w_1 \\ \vdots \\ w_{k-1} \end{pmatrix} = w_k \\
\quad x \geq 0 \\
(x, w_1, \ldots, w_k) \in R^{n+k}
\end{array}
\right.
\end{array}
\right.
$$

donde $H_1 \in M(1,n)$, $H_2 \in M(1,n+1), \ldots, H_k \in M(1,n+k)$ y los elementos de H_i, $i = 1,\ldots,k$, correspondientes a las variables básicas que había al introducir la variable w_i son ceros.

Para el estudio de las propiedades del algoritmo, necesitamos previamente unas definiciones relativas a los diferentes tipos de soluciones del programa.

Definición 2.4.2.1.

Dado el programa:

$$
(PP^{(o)}\left\{
\begin{array}{l}
\text{Min. } P(x) \\
\text{bajo las restricciones:} \\
Y^{(o}\left\{
\begin{array}{l}
A\,x = b \\
\quad x \geq 0 \\
\quad x \in R^n
\end{array}
\right.
\end{array}
\right.
$$

siendo $P(x)$ función polinómica, $A \in M(m,n)$, $rg(A)=m$, $b \in M(m,1)$ y ambas matrices de números reales.

- Se llama base a toda submatriz cuadrada regular de dimensiones $m \times m$ extraída de A (existe al menos una puesto que suponemos que $rg(A) = m$).

  Sea B una base, entonces permutando las columnas de A podemos expresar siempre A en la forma $A = (B \vdots F)$, siendo $F \in M(m,n-m)$ la submatriz de A correspondiente a las columnas de A que no están en la base. Análogamente, podemos poner $x = \begin{pmatrix} x_B \\ \cdots \\ x_F \end{pmatrix}$.

  Si x es solución de $A x = b$, en virtud de lo anterior, se puede expresar:

$$B x_B + F x_F = b \qquad (1)$$

- Se llama solución básica de $A x = b$ (relativa a la base B), a la solución particular de $A x = b$, obtenido haciendo $x_F = 0$ en la expresión (1). Se determina de modo único resolviendo el sistema de Cramer $B x_B = b$, luego $x_B = B^{-1} b$.

Se dice que $x \in R^n$ es:

- Solución factible de $(PP^{(e)})$ si $A x = b$ y $x \geq 0$.
- Solución óptima de $(PP^{(o)})$ si es una solución factible de $(PP^{(o)})$ que minimiza a $P(x)$.

- Solución básica factible de $(PP^{(o)})$ si es una solución básica de $A\ x = b$ y $x_B \geqslant 0$.
  Una base correspondiente a una solución básica factible se llama base factible.
- Solución básica degenerada de $(PP^{(o)})$ si es una solución básica de $A\ x = b$ tal que $x_B = B^{-1}\ b$ tiene componentes nulas.

Teorema 2.4.2.1.

Sean S_o, S_k $\forall k \in N^*$, los conjuntos de soluciones factibles de los programas $(PP^{(o)})$ y $(PP^{(k)})$ respectivamente y SO_o, SO_k $\forall k \in N^*$, los conjuntos de soluciones óptimas de los citados programas. Entonces, se verifica:

a) $\text{Card}(S_o) = \text{Card}(S_k) \qquad \forall k \in N^*$

b) $\left(\dfrac{x^o}{w^o} \right)$ es solución óptima de $(PP^{(k)})$ si y sólo si x^o es solución óptima de $(PP^{(o)})$.

c) $\text{Card}(SO_o) = \text{Card}(SO_k) \quad \forall k \in N^*$

Demostración

a) Sea $\forall x^1 \in S_o \iff \begin{cases} A\ x^1 = b \\ x^1 \geqslant 0 \end{cases} \qquad \Rightarrow$

$$\Rightarrow \quad \left.\begin{array}{l} A\,x^1 = b \\ H_1\,x^1 = w_1^1 \\ H_2\left(\dfrac{x^1}{w_1^1}\right) = w_2^1 \\ \cdots\cdots\cdots\cdots \\ H_k\begin{pmatrix} \dfrac{x^1}{w_1^1} \\ \vdots \\ w_{k-1}^1 \end{pmatrix} = w_k^1 \\ x^1 \geq 0 \end{array}\right\} \quad \Leftrightarrow \quad \frac{x^1}{w^1} \in S_k$$

Luego,

$$\mathrm{Card}(S_o) \leq \mathrm{Card}(S_k) \qquad (1)$$

Recíprocamente:

$$\forall\left(\frac{x^*}{w^*}\right) \in S_k \qquad \Leftrightarrow$$

$$\Leftrightarrow \quad \left.\begin{array}{l} A\,x^* = b \\ H_1 x^* = w_1^* \\ H_2\left(\dfrac{x^*}{w_1^*}\right) = w_2^* \\ \cdots\cdots\cdots\cdots \\ H_k\begin{pmatrix} \dfrac{x^*}{w_1^*} \\ \vdots \\ w_{k-1}^* \end{pmatrix} = w_k^* \\ x^* \geq 0 \end{array}\right\} \quad \Rightarrow$$

$$\Rightarrow \qquad \left.\begin{array}{l} A\,x^* = b \\ x^* \geq 0 \end{array}\right\} \Longleftrightarrow \quad x^* \in S_o$$

Luego,

$$\text{Card}(S_k) \leq \text{Card}(S_o) \qquad (2)$$

De (1) y (2), se concluye que:

$$\text{Card}(S_o) = \text{Card}(S_k) \qquad \forall k \in N^*$$

Es decir, las soluciones factibles de estos programas se corresponden biyectivamente.

b) La demostración de esta proposición la realizaremos por reducción al absurdo.

$\Rightarrow$ Supongamos que $\exists \begin{pmatrix} \dfrac{x^o}{w^o} \end{pmatrix} \in SO_k$ tal que $x^o \notin SO_o$.

Entonces, $\exists\, x^1 \in SO_o$ tal que $P(x^1) < P(x^o)$. Y, en el programa $(PP^{(k)})$, tendríamos:

$$\left.\begin{array}{l} A\,x^1 = b \\ H_1 x^1 = w_1^1 \\ \cdots\cdots\cdots \\ H_K \begin{pmatrix} \dfrac{x^1}{w_1^1} \\ \vdots \\ w_{K-1}^1 \end{pmatrix} = w_K^1 \\ x^1 \geq 0 \end{array}\right\} \quad \text{con} \quad P(x^1) < P(x^o) \qquad \Rightarrow$$

$\Rightarrow \quad \begin{pmatrix} \dfrac{x^o}{w^o} \end{pmatrix}$ no sería óptimo de $(PP^{(k)})$, ya que habríamos encontrado el punto $\begin{pmatrix} \dfrac{x^1}{w^1} \end{pmatrix} \in Y^{(k}$ con

$P(x^1) < P(x^o)$. Contradicción que se origina al suponer que $\exists \left(\frac{x^o}{w^o}\right) \in SO_k$ tal que $x^o \notin SO_o$. Luego, $\forall \left(\frac{x^o}{w^o}\right) \in SO_k$ se concluye que $x^o \in SO_o$.

$\Leftarrow$ Recíprocamente, aplicando un razonamiento análogo al anterior, si $\exists\, x^o \in SO_o$ tal que $\left(\frac{x^o}{w^o}\right) \in SO_k$, ello implicaría que $\exists \left(\frac{x^*}{w^*}\right) \in SO_k$ tal que $P(x^*) < P(x^o)$, verificándose las restricciones:

$$\left.\begin{array}{l} A\,x^* = b \\ H_1 x^* = w_1^* \\ \cdots\cdots\cdots \\ H_k \begin{pmatrix} \frac{x^*}{w_1^*} \\ \vdots \\ w_{k-1}^* \end{pmatrix} = w_k^* \\ \quad x^* \geq 0 \end{array}\right\} \Rightarrow$$

$$\Rightarrow \quad \left.\begin{array}{l} A\,x^* = b \\ x^* \geq 0 \end{array}\right\} \quad y \quad P(x^*) < P(x^o)$$

Ello implicaría, que $x^o \notin SO_o$. Contradicción que nace de suponer que $\exists\, x^o \in SO_o$ tal que $\left(\frac{x^o}{w^o}\right) \notin SO_k$. Luego, $\forall x^o \in SO_o$ se sigue que $\left(\frac{x^o}{w^o}\right) \in SO_k$.

Es decir, existe una biyección entre las soluciones óp timas de los programas $(PP^{(o)})$ y $(PP^{(k)})$.

c) Es una consecuencia de la proposición b), ya que en virtud de la biyección que hay entre las soluciones óptimas de los citados programas, se concluye que $\mathrm{Card}(SO_o) = \mathrm{Card}(SO_k)$.

Termina, así, la demostración del teorema 2.4.2.1.

Consecuencias de este teorema, son los siguientes corolarios.

Corolario 2.4.2.1.

$$\mathrm{Card}(S_k) = \mathrm{Card}(S_r) \qquad \forall k, r \in N^*$$

Demostración

En efecto, de acuerdo con la propiedad a) del teorema 2.4.2.1., es

$$\mathrm{Card}(S_o) = \mathrm{Card}(S_k) \quad \text{y} \quad \mathrm{Card}(S_o) = \mathrm{Card}(S_r)$$

Luego,

$$\mathrm{Card}(S_k) = \mathrm{Card}(S_r) \qquad \forall k, r \in N^*$$

Corolario 2.4.2.2.

$$\mathrm{Card}(SO_k) = \mathrm{Card}(SO_r) \qquad \forall k, r \in N^*$$

Demostración

De acuerdo con la propiedad c) del teorema 2.4.2.1., es:

$$\mathrm{Card}(SO_o) = \mathrm{Card}(SO_k) \quad y \quad \mathrm{Card}(SO_o) = \mathrm{Card}(SO_r)$$

Luego,

$$\mathrm{Card}(SO_k) = \mathrm{Card}(SO_r) \qquad \forall k,r \in N^*$$

Corolario 2.4.2.3.

$$\begin{pmatrix} \frac{x^o}{w_1^o} \\ \vdots \\ w_k^o \end{pmatrix} \in SO_k \iff \begin{pmatrix} \frac{x^o}{w_1^o} \\ \vdots \\ w_k^o \\ \vdots \\ w_r^o \end{pmatrix} \in SO_r \qquad \forall r,k \in N^* \quad \text{con } r > k$$

Demostración

En efecto, de acuerdo con la proposición b) del teorema 2.4.2.1., es:

$$\begin{pmatrix} \frac{x^o}{w_1^o} \\ \vdots \\ w_k^o \end{pmatrix} \in SO_k \iff x^o \in SO_o \qquad (1)$$

Análogamente,

$$\begin{pmatrix} \frac{x^o}{w_1^o} \\ \vdots \\ w_r^o \end{pmatrix} \in SO_r \iff x^o \in SO_o \qquad (2)$$

De (1) y (2), se concluye,

$$\begin{pmatrix} \dfrac{x^o}{w_1^o} \\ \vdots \\ w_k^o \end{pmatrix} \in SO_k \quad <=> \quad \begin{pmatrix} \dfrac{x^o}{w_1^o} \\ \vdots \\ w_k^o \\ \vdots \\ w_r^o \end{pmatrix} \in SO_r$$

Teorema 2.4.2.2.

Sean $S_{BF}^{(o}$ y $S_{BF}^{(k}$ las soluciones básicas factibles, respectivamente, de los programas $(PP^{(o})$ y $(PP^{(k})$ $\forall k \in N^*$. Entonces, se verifica que:

a) El conjunto $S_{BF}^{(o}$ se corresponde biyectivamente con un subconjunto de $S_{BF}^{(k}$ $k \in N^*$, y en las soluciones correspondientes la función objetivo toma el mismo valor.

b) $\forall k, r \in N^*$, el conjunto $S_{BF}^{(k}$ se corresponde biyectivamente con un subconjunto de $S_{BF}^{(k+r}$, y en las soluciones correspondientes, la función objetivo toma el mismo valor.

Demostración

a) Dados $S_{BF}^{(o}$ y $S_{BF}^{(k}$, definimos la siguiente aplicación:

$$S_{BF}^{(o} \xrightarrow{\quad g^{(k} \quad} S_{BF}^{(k}$$

$$x^o = (x_B^o \ \vdots \ x_F^o = 0)^T \longrightarrow g^{(k}(x^o) = (x_B^o \ \vdots \ x_F^o = 0, w_1^o, \ldots, w_k^o)^T$$

donde $w_1^o, \ldots, w_k^o$ son los valores obtenidos en el programa $(PP^{(k})$ según se ha indicado en el subapartado 2.4.2.1.

Veamos que $g^{(k}(x^o) \in S_{BF}^{(k}$.

Sabemos que, dada x solución factible de $(PP^{(o})$ y B base de A, es:

$$B\, x_B + F\, x_F = b$$

Luego,

$$x_B = B^{-1}\, b - B^{-1}\, F\, x_F$$

Para esta solución de $(PP^{(o})$, se obtienen unos valores $w_1, \ldots, w_k$ en $(PP^{(k})$, y tendremos:

$$\begin{pmatrix} B & & & & \bigcirc \\ & -1 & & & \\ & & -1 & & \\ & & & \ddots & \\ \bigcirc & & & & -1 \end{pmatrix} \begin{pmatrix} x_B \\ \hline w_1 \\ \vdots \\ w_k \end{pmatrix} + \begin{pmatrix} F \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix} x_F = \begin{pmatrix} b \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix}$$

Y, en forma simplificada:

$$\bar{B} \begin{pmatrix} x_B \\ \hline w_1 \\ \vdots \\ w_k \end{pmatrix} + \bar{F}\, x_F = \bar{b}$$

siendo, $\bar{B}\in M(m+k,m+k)$, $x_B\in M(m,1)$, $\bar{F}\in M(m+k,n-m)$, $x_F\in M(n-m,1)$ y $\bar{b}\in M(m+k,1)$.

En particular, para $x_F = 0$ obtenemos la solución básica factible de $(PP^{(o)}$: $x_B = x_B^o$, y en $(PP^{(k)}$ se obtiene:

$$\bar{B}\begin{pmatrix} x_B^o \\ \hline w_1^o \\ \vdots \\ w_k^o \end{pmatrix} = \bar{b}$$

Ahora bien, como $\det(B) \neq 0$ se sigue que $\det(\bar{B}) = (-1)^k \det(B) \neq 0$. Luego, $\exists\, \bar{B}^{-1}$, y por tanto:

$$\begin{pmatrix} x_B^o \\ \hline w_1^o \\ \vdots \\ w_k^o \end{pmatrix} = \bar{B}^{-1}\,\bar{b}$$

Luego, según la definición 2.4.2.1., $\begin{pmatrix} x_B^o \\ \hline w_1^o \\ \vdots \\ w_k^o \end{pmatrix}$ es una solución básica factible de $(PP^{(k)}$. Es decir,

$$g^{(k}(x^o) \in S_{BF}^{(k}$$

Por otra parte, $g^{(k}$ es aplicación, ya que a cada elemento de $S_{BF}^{(o}$ le hace corresponder un único elemento de $S_{BF}^{(k}$, que es la única solución básica factible asociada a la base $\bar{B}$.

Probemos, ahora, que $g^{(k}$ es inyectiva.

Sean $x^1, x^2 \in S_{BF}^{(o}$ y $x^1 \neq x^2$

Como,

$$\left.\begin{aligned} g^{(k}(x^1) &= (x^1, w_1^1, \ldots, w_k^1)^T \\ g^{(k}(x^2) &= (x^2, w_1^2, \ldots, w_k^2)^T \end{aligned}\right\} \text{ y al ser } x^1 \neq x^2$$

se sigue que:

$$(x^1, w_1^1, \ldots, w_k^1)^T \neq (x^2, w_1^2, \ldots, w_k^2)^T$$

o lo que es equivalente:

$$g^{(k}(x^1) \neq g^{(k}(x^2)$$

Luego, $g^{(k}$ es inyectiva.

Considerando el conjunto $g^{(k}(S_{BF}^{(o}) \subset S_{BF}^{(k}$, es claro que $g^{(k}$ es una biyección de $S_{BF}^{(o}$ en $g^{(k}(S_{BF}^{(o})$.

Probemos finalmente, que el valor de la función objetivo es el mismo para las soluciones básicas que se corresponden mediante $g^{(k}$.

En efecto, la función objetivo no se modifica en los sucesivos programas, independientemente de que adopte expresiones distintas en función de las variables libres existentes en cada caso. Tomaremos como expresión de la función objetivo la que tiene en el programa inicial: $P(x_1, \ldots, x_n)$.

Sean $x^o = (x_B^o \vdots x_F^o = 0)^T$ y

$g^{(k}(x^o) = (x_B^o, x_F^o = 0, w_1^o, \ldots, w_k^o)^T$ las correspondientes soluciones básicas factibles de $(PP^{(o)})$ y $(PP^{(k)})$, entonces:

$$P(x^o) = P(x_{b_1}^o, \ldots, x_{b_m}^o, \overbrace{0, \ldots, 0}^{n-m})$$

y para $g^{(k}(x^o)$, como $w_1^o, \ldots, w_k^o$ no figuran en $P(x_1, \ldots, x_n)$, resulta el mismo valor de la función objetivo: $P(x^o)$.

b) Pasamos a demostrar la proposición de este apartado.

$\forall z^o = (x^o, w^o)^T \in S_{BF}^{(k}$ tiene exactamente $m+k$ variables básicas, de las cuales a lo sumo k son variables w_j.

Definimos la siguiente aplicación:

$$S_{BF}^{(k} \xrightarrow{\quad g^{(r} \quad} S_{BF}^{(k+r}$$

$$z^o = (x_{b_1}^o, \ldots, x_{b_p}^o, \overbrace{0, \ldots, 0}^{n-p}, w_{b_1}^o, \ldots, w_{b_h}^o, \overbrace{0, \ldots, 0}^{k-h}) \xrightarrow[\substack{p \geqslant m \\ h \leq k \\ p+h=m+k}]{} g^{(k}(z^o) =$$

$$= (x_{b_1}^o, \ldots, x_{b_p}^o, \overbrace{0, \ldots, 0}^{n-p}, w_{b_1}^o, \ldots, w_{b_h}^o, \overbrace{0, \ldots, 0}^{k-h}, w_{k+1}^o, \ldots, w_{k+r}^o)$$

Razonando de modo análogo a como se hizo en el punto a), si z es solución factible de $(PP^{(k)})$, entonces:

$$\bar{B}\, z_B + \bar{F}\, z_F = \bar{b}$$

siendo $\bar{B}$ y $\bar{F}$ las matrices básica y de variables libres relativas al programa $(PP^{(k)})$, con $\bar{B} \in M(m+k, m+k)$,

$\bar{F} \in M(m+k,n-m)$ y $\bar{b} \in M(m+k,1)$.

Para esta solución factible de $(PP^{(k)})$, se obtienen unos valores de $w_{k+1},\ldots,w_{k+r}$ en $(PP^{(k+r)})$, y tendremos:

$$\begin{pmatrix} \bar{B} & & & & 0 \\ & -1 & & & \\ & & -1 & & \\ & & & \ddots & \\ 0 & & & & -1 \end{pmatrix} \begin{pmatrix} z_B \\ \cdots \\ w_{k+1} \\ \vdots \\ w_{k+r} \end{pmatrix} + \begin{pmatrix} \bar{F} \\ \cdots \\ 0 \end{pmatrix} z_F = \begin{pmatrix} \bar{b} \\ \cdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Y, simplificadamente:

$$B^* \begin{pmatrix} z_B \\ \cdots \\ w_{k+1} \\ \vdots \\ w_{k+r} \end{pmatrix} + F^* z_F = b^*$$

siendo $B^* \in M(m+k+r,m+k+r)$, $F^* \in M(m+k+r,n-m)$ y $b^* \in M(m+k+r,1)$.

En particular, para $z_F = 0$, obtenemos la solución básica factible de $(PP^{(k)})$: $z_B = z_B^o$, y en $(PP^{(k+r)})$ se obtiene:

$$B^* \begin{pmatrix} z_B^o \\ \cdots \\ w_{k+1}^o \\ \vdots \\ w_{k+r}^o \end{pmatrix} = b^*$$

Ahora bien, al ser $\det(\bar{B}) \neq 0$, se sigue que:

$\det(B^*) = (-1)^r \det(\bar{B}) \neq 0 \quad \Rightarrow \quad \exists (B^*)^{-1}.$

Luego,

$$\begin{pmatrix} z_B^o \\ \cdots \\ w_{k+1}^o \\ \vdots \\ w_{k+r}^o \end{pmatrix} = (B^*)^{-1} b^*$$

Por tanto, según la definición 2.4.2.1.,

$(z_B^o \vdots w_{k+1}^o, \ldots, w_{k+r}^o)^T$ es una solución básica factible de $(PP^{(k+r)})$.

Por otra parte, $g^{(r}$ es aplicación, ya que a cada elemento de $S_{BF}^{(k}$ le hace corresponder un único elemento de $S_{BF}^{(k+r}$ que es la solución básica factible única de $(PP^{(k+r)})$ asociada a la base B^*.

Probemos, seguidamente, que $g^{(r}$ es inyectiva.

En efecto, sean $z^1, z^2 \in S_{BF}^{(k+r}$ y $z^1 \neq z^2$.

Al ser:

$$\left.\begin{aligned} g^{(r}(z^1) &= (z^1, w_{k+1}^1, \ldots, w_{k+r}^1)^T \\ g^{(r}(z^2) &= (z^2, w_{k+1}^2, \ldots, w_{k+r}^2)^T \end{aligned}\right\} \text{ y } \quad z^1 \neq z^2$$

se sigue que:

$$g^{(r}(z^1) \neq g^{(r}(z^2)$$

Luego, $g^{(r}$ es inyectiva.

Considerando el conjunto $g^{(k}(S_{BF}^{(k}) \subset S_{BF}^{(k+r}$, es evidente que $g^{(k}$ es una biyección de $S_{BF}^{(k}$ en $g^{(k}(S_{BF}^{(k+r})$.

Finalmente, el valor de la función objetivo es el mismo para las soluciones básicas factibles que se corresponden mediante $g^{(k}$, debido a que tal función objetivo es invariante en los sucesivos programas, aunque adopte distintas expresiones al explicitarla en función de las variables libres que existan en cada caso.

Si tomamos para la función objetivo la expresión inicial $P(x_1,\dots,x_n)$, que depende sólo de las variables x_i, vemos que el valor de la misma para $z^o = (x^o,w^o)^T$ y $g^{(k}(z^o)$ es $P(x^o)$.

Termina, así, la demostración del teorema 2.4.2.2..

El teorema y el corolario siguientes, son válidos cuando las soluciones básicas factibles de los sucesivos programas del algoritmo son no degeneradas.

<u>Teorema 2.4.2.3.</u>

Sean $(PP^{(k})$ y $(PP^{(k+r})$ $\forall k \in N$, $\forall r \in N^*$ dos programas cualesquiera, $g^{(r}$ la aplicación inyectiva definida, en el teorema 2.4.2.2., entre $S_{BF}^{(k}$ y $S_{BF}^{(k+r}$. Supongamos que las soluciones básicas factibles de los programas del algoritmo son no degeneradas, entonces se verifica que:

$$g^{(k}(S_{BF}^{(k}) = \{z=(x,w)^T \in S_{BF}^{(k+r} / w_{k+1} \neq 0, w_{k+2} \neq 0, \dots, w_{k+r} \neq 0\}$$

Demostración

$\forall u = (x, w_1, \ldots, w_k)^T \in S_{BF}^{(k}$ se cumple que u tiene exactamente m+k variables básicas no nulas, por hipótesis.

Por otra parte, según el teorema 2.4.2.2., es

$$g^{(r}(u) = (x, w_1, \ldots, w_k, w_{k+1}, \ldots, w_{k+r})^T \in S_{BF}^{(k+r}$$

por tanto, y de acuerdo con la hipótesis formulada, $g^{(r}(u)$ tiene exactamente m+k+r variables básicas no nulas.

Si alguna w_j, $j = k+1, \ldots, k+r$, fuese nula, entonces $(x, w_1, \ldots, w_k)^T = u$ tendría más de m+k variables no nulas y, por tanto, no sería una solución básica factible de $(PP^{(k})$. Contradicción, que tiene su origen en suponer que alguna w_j, $j=k+1, \ldots, k+r$, puede ser nula. Luego, todas ellas son no nulas. Por tanto,

$$g^{(r}(S_{BF}^{(k}) = \{ z=(x,w)^T \in S_{BF}^{(k+r} \ / \ w_{k+1} \neq 0, \ldots, w_{k+r} \neq 0 \}$$

Con lo que se termina la demostración de este teorema.

Corolario 2.4.2.4.

Sean $(PP^{(k})$ y $(PP^{(k+r})$ $\forall k \in N$, $\forall r \in N^*$ dos programas cualesquiera del algoritmo, $g^{(r}$ la aplicación inyectiva, definida en el teorema 2.4.2.2., entre $S_{BF}^{(k}$ y $S_{BF}^{(k+r}$.

Supongamos que las soluciones básicas factibles de los programas del algoritmo son no degeneradas, entonces se verifica que las soluciones básicas factibles de $(PP^{(k+r)})$ que no se corresponden, mediante $g^{(r}$, con las soluciones básicas factibles de $(PP^{(k})$ son:

$$\{z=(x,w)^T \in S_{BF}^{(k+r} \ / \ w_{k+1}=0 \ \text{ó} \ w_{k+2}=0 \ \text{ó} \ldots \text{ó} \ w_{k+r}=0\}$$

Demostración

Por el teorema 2.4.2.3., las soluciones básicas factibles de $(PP^{(k+r)})$ que se corresponden, mediante $g^{(k}$, con las soluciones básicas factibles de $(PP^{(k})$ son aquellas para las que $w_{k+1} \neq 0, w_{k+2} \neq 0, \ldots, w_{k+r} \neq 0$; luego, aquellas que no se corresponden serán las que cumplan que:

$$w_{k+1} = 0 \ \text{ó} \ w_{k+2} = 0 \ \text{ó} \ldots \text{ó} \ w_{k+r} = 0$$

Se termina, así, la demostración de este corolario.

Nota 2.4.2.1.

Si existieran soluciones básicas factibles degeneradas, podría haber alguna en $(PP^{(k+r)})$ verificando $w_{k+1}=0$ ó $w_{k+2}=0$ ó.....ó $w_{k+r}=0$ y, sin embargo, corresponderse con alguna solución básica factible de $(PP^{(k)})$, bastaría que una solución básica factible de $(PP^{(k+r)})$ tuviera, de sus $m+k+r$ variables básicas, las r variables w_{k+1}, $w_{k+2}, \ldots, w_{k+r}$ nulas, entonces las $m+k$ variables restantes se corresponderían con una solución básica factible de $(PP^{(k)})$.

Teorema 2.4.2.4.

Sean S_k y S_{k+r} $\forall k \in N$, $\forall r \in N^*$, los conjuntos de soluciones factibles de $(PP^{(k)})$ y $(PP^{(k+r)})$, respectivamente. Entonces, se verifica que:

S_k está acotado si y sólo si S_{k+r} está acotado

Demostración

Sabemos, por el corolario 2.4.2.1., que existe una correspondencia biyectiva entre S_k y S_{k+r}. Por tanto, vamos a probar que las soluciones correspondientes por tal biyección están acotadas.

$\Rightarrow$] S_k acotado $\Leftrightarrow$

$$\Leftrightarrow \quad \forall u^o = (x^o; w_1^o, \ldots, w_k^o)^T \in S_k \text{ es}$$

$$\|u^o\| = \sqrt{\sum_{i=1}^{n} (x_i^o)^2 + \sum_{j=1}^{k} (w_j^o)^2} < +\infty$$

La solución correspondiente en S_{k+r} es:

$$z^o = (u^o; w_{k+1}^o, \ldots, w_{k+r}^o)^T \quad \text{con} \quad w_{k+1}^o = H_{k+1} u^o,$$

$$w_{k+2}^o = H_{k+2}(u^o; w_{k+1}^o)^T, \ldots, w_{k+r}^o = H_{k+r}(u^o; w_{k+1}^o, \ldots, w_{k+r-1}^o)^T$$

Pero,

$$|w^{o}_{k+1}| = |H_{k+1}\ u^{o}| \underset{\uparrow}{\leq} \|H_{k+1}\|\ \|u^{o}\| =$$

Desigualdad de
Cauchy-Schwarz

$$= \sqrt{\sum_{i=1}^{n+k} h^{2}_{k+1,i}}\ \|u^{o}\| \leq \sqrt{(n+k) \underset{i=1,\dots,n+k}{\text{máx.}}\ h^{2}_{k+1,i}}\ \|u^{o}\| =$$

$$= \sqrt{n+k}\ \sqrt{\underset{i=1,\dots,n+k}{\text{máx.}}\ h^{2}_{k+1,i}}\ \|u^{o}\| = \sqrt{n+k}\ h_{k+1}\ \|u^{o}\|$$

Luego,

$$|w^{o}_{k+1}| \leq \sqrt{n+k}\ h_{k+1}\ \|u^{o}\| < +\infty$$

ya que $\|u^{o}\| < +\infty$, $\sqrt{n+k} < +\infty$, $h_{k+1} < +\infty$

Análogamente,

$$|w^{o}_{k+2}| = |H_{k+2}(u^{o};w^{o}_{k+1})^{T}| \underset{\uparrow}{\leq} \|H_{k+2}\|\ \|(u^{o};w^{o}_{k+1})^{T}\| =$$

Desigualdad de
Cauchy-Schwarz

$$= \sqrt{\sum_{i=1}^{n+k+1} h^{2}_{k+2,i}}\ \sqrt{\sum_{j=1}^{n+k}(u^{o}_{j})^{2} + (w^{o}_{k+1})^{2}} \leq$$

$$\leq \sqrt{\sum_{i=1}^{n+k+1} h^{2}_{k+2,i}}\ \left(\sqrt{\sum_{j=1}^{n+k}(u^{o}_{j})^{2}} + |w^{o}_{k+1}|\right) \leq$$

$$\leq \sqrt{(n+k+1) \underset{i=1,\dots,n+k+1}{\text{máx.}}\ h^{2}_{k+2,i}}\ \left(\|u^{o}\| + |w^{o}_{k+1}|\right) =$$

$$= \sqrt{n+k+1}\ h_{k+2}\ \left(\|u^{o}\| + |w^{o}_{k+1}|\right)$$

Luego,

$$\left|w_{k+2}^{o}\right| \leq \sqrt{n+k+1}\; h_{k+2}\left(\|u^{o}\| + \left|w_{k+1}^{o}\right|\right) < +\infty$$

ya que $\|u^{o}\| < +\infty$, $\left|w_{k+1}^{o}\right| < +\infty$, $\sqrt{n+k+1} < +\infty$ y $h_{k+2} < +\infty$.

Y así sucesivamente proseguiríamos, llegando, por último, a w_{k+r}^{o}:

$$\left|w_{k+r}^{o}\right| = \left|H_{k+r}\,(u^{o};w_{k+1}^{o},\ldots,w_{k+r-1}^{o})^{T}\right| \underset{\substack{\uparrow\\ \text{Desigualdad de}\\ \text{Cauchy-Schwarz}}}{\leq}$$

$$\leq \|H_{k+r}\|\;\|(u^{o};w_{k+1}^{o},\ldots,w_{k+r-1}^{o})^{T}\| =$$

$$= \sqrt{\sum_{i=1}^{n+k+r-1} h_{k+r,i}^{2}}\;\sqrt{\sum_{j=1}^{n+k}(u_j^{o})^{2} + \sum_{t=1}^{r-1}(w_{k+t}^{o})^{2}} \leq$$

$$\leq \sqrt{(n+k+r-1)\;\max_{i=1,\ldots,n+k+r-1}\{h_{k+r,i}^{2}\}}\;\left(\|u^{o}\| + \left|w_{k+1}^{o}\right| + \ldots + \left|w_{k+r-1}^{o}\right|\right) =$$

$$= \sqrt{n+k+r-1}\; h_{k+r}\left(\|u^{o}\| + \left|w_{k+1}^{o}\right| + \ldots + \left|w_{k+r-1}^{o}\right|\right)$$

Luego,

$$\left|w_{k+r}^{o}\right| \leq \sqrt{n+k+r-1}\; h_{k+r}\left(\|u^{o}\| + \left|w_{k+1}^{o}\right| + \ldots + \left|w_{k+r-1}^{o}\right|\right) < +\infty$$

puesto que $\|u^{o}\| < +\infty$, $\left|w_{k+1}^{o}\right| < +\infty, \ldots, \left|w_{k+r-1}^{o}\right| < +\infty$, $\sqrt{n+k+r-1} < +\infty$ y $h_{k+r} < +\infty$

Por tanto,

si $\|u^o\| < +\infty \Rightarrow \|z^o\| = \|(u^o; w^o_{k+1}, \ldots, w^o_{k+r})^T\| < +\infty$

Luego, S_{k+r} está acotado.

$\Leftarrow$ Recíprocamente,

S_{k+r} acotado $\Leftrightarrow \forall z^o = (x^o; w^o_1, \ldots, w^o_k, w^o_{k+1}, \ldots, w^o_{k+r})^T \in S_{k+r}$

es $\|z^o\| = \|(u^o; w^o_{k+1}, \ldots, w^o_{k+r})^T\| < +\infty$

La correspondiente solución en S_k es u^o. Pero,

$$\|u^o\| = \sqrt{\sum_{i=1}^{n+k} (u^o_i)^2} \leqslant \sqrt{\sum_{i=1}^{n+k} (u^o_i)^2 + \sum_{j=1}^{r} (w^o_{k+j})^2} =$$

$$= \|(u^o; w^o_{k+1}, \ldots, w^o_{k+r})\| = \|z^o\|$$

Luego,

$$\|u^o\| \leqslant \|z^o\| < +\infty \Rightarrow \|u^o\| < +\infty$$

Por tanto, S_k está acotado.

Termina, así, la demostración del teorema 2.4.2.4.

Nota 2.4.2.2.

De este teorema se concluye que si el conjunto de soluciones factibles del programa inicial está acotado, entonces los conjuntos de soluciones factibles de los restantes programas $(PP^{(k)})$ $\forall k \in N^*$, están acotados. Como sabemos, una condición suficiente para que las soluciones factibles del programa inicial estén acotadas es que el conjunto de las restricciones lineales del citado programa sea compacto; lo cual se consigue exigiendo que las restricciones lineales formen un poliedro convexo.

Teorema 2.4.2.5.

Dado el programa:

$$
(PP^{(k)}) \quad \left\{ \begin{array}{l} \text{Min. } P^{(k} (x_F;\ w_F) \\ \text{bajo las restricciones:} \\ \quad Y^{(k} \left\{ \begin{array}{l} x_B = x_B^o + Cx_F + Dw_F \\ w_B = w_B^o + C'x_F + D'w_F \\ x_B \geqslant 0 \\ x_F \geqslant 0 \end{array} \right. \end{array} \right.
$$

en el que tanto la función objetivo como las variables básicas están expresadas a través de las variables libres y $k \in N$. Entonces,

$$(x^o;\ w^o)^T = (x_B^o; x_F=0;\ w_B^o; w_F=0)^T \in Y^{(k}$$

es solución óptima de $(PP^{(k)})$ si y sólo si:

$$\left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0,\,w_F=0)} \geqslant 0 \quad y \quad \left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0,\,w_F=0)} = 0$$

Demostración

$\Longrightarrow$ La haremos por reducción al absurdo.

Supongamos que siendo óptimo $(x^o;\ w^o)^T$ ocurriera que:

$$\exists j \text{ con } \left(\frac{\partial p^{(k}}{\partial x_{f_j}}\right)_{(0)} < 0 \quad \text{ó} \quad \exists t \text{ con } \left(\frac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} \neq 0$$

a) Si $\left(\frac{\partial p^{(k}}{\partial x_{f_j}}\right)_{(0)} < 0$, podríamos avanzar en la dirección x_{f_j}, haciéndose $x_{f_j} > 0$, sin perderse la factibilidad. Llegaríamos a un punto $(x^*; w^*)^T \in Y^{(k}$, que mejora la solución anterior, con lo cual $(x^o; w^o)^T \in Y^{(k}$ no sería óptimo. Contradicción, que tiene su origen en suponer que $\exists j$ tal que $\left(\frac{\partial p^{(k}}{\partial x_{f_j}}\right)_{(0)} < 0$. Luego, $\forall j$ es $\left(\frac{\partial p^{(k}}{\partial x_{f_j}}\right)_{(0)} \geq 0$.

b) Si $\exists t$ con $\left(\frac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} \neq 0$, podría ocurrir:

1) $\left(\frac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} > 0$, en cuyo caso avanzaríamos en la dirección w_{f_t}, con $w_{f_t} < 0$, obteniendo una solución factible que mejora a $(x^o; w^o)^T$ y, por tanto, esta solución no sería óptima. La contradicción se origina al suponer que $\exists t$ con $\left(\frac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} > 0$

2) $\left(\frac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} < 0$, avanzaríamos en la dirección w_{f_t}, con $w_{f_t} > 0$, sin deshacer la factibilidad. La nueva solución mejoraría a $(x^o; w^o)^T$ y, por tanto, ésta no sería óptima. Contradicción que se origina al suponer que $\exists t$ con $\left(\frac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} > 0$.

Luego, $\forall t$ es $\left(\dfrac{\partial P^{(k}}{\partial w_{f_t}}\right)_{(0)} = 0$

Así, pues, si $(x^{o};\ w^{o})^{T} \in Y^{(k}$ es solución óptima, entonces:

$$\left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0;\,w_F=0)} \geqslant 0 \quad \text{y} \quad \left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0;\,w_F=0)} = 0 \quad (*)$$

$\Longleftarrow$ Recíprocamente, si para $(x^{o};\ w^{o})^{T} \in Y^{(k}$ es:

$$\left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0;\,w_F=0)} \geqslant 0 \quad \text{y} \quad \left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0;\,w_F=0)} = 0$$

probemos que se verifican en dicho punto las condiciones de Kuhn y Tucker, y, por tanto, $(x^{o};\ w^{o})^{T} \in Y^{(k}$ es una solución óptima de $(PP^{(k})$.

En efecto, la función Lagrangiana asociada a $(PP^{(k})$ es:

$$L(x;w;\lambda;\mu;\nu;\tau) = P^{(k}(x_F;w_F) + \lambda^{T}(x_B-x_B^{o}-Cx_F-Dw_F) +$$
$$+ \mu^{T}(w_B-w_B^{o}-C'x_F-D'w_F) - \nu^{T}x_B - \tau^{T}x_F$$

En el punto $(x^{o};\ w^{o})^{T} \in Y^{(k}$, se verificarán las condiciones de Kuhn y Tucker si y sólo si:

$\exists\lambda^{o}$ arbitrario en el signo, $\lambda^{o} \in R^{p}$ con $p \leqslant \text{mín}\{m+k,n\}$, $\exists\mu^{o}$ arbitrario en el signo, $\mu^{o} \in R^{h}$ con $h \leqslant k$, $\nu^{o} \geqslant 0$, $\nu^{o} \in R^{p}$ y

---

(*) La obtención, en los supuestos a) y b) anteriores, de una solución factible que mejora la solución $(x^{o};\ w^{o})^{T}$ es inmediata, mas no evidente. Se demuestra en la Nota 2.4.2.3.

$\tau^{o} \geq 0$, $\tau^{o} \in R^{n-p}$, tales que:

1°) $$\left(\frac{\partial L}{\partial x_B}\right)_{(x^o;w^o)} = \lambda^o - \nu^o = 0 \qquad (1)$$

$$\left(\frac{\partial L}{\partial w_B}\right)_{(x^o;w^o)} = \mu^o = 0 \qquad (2)$$

$$\left(\frac{\partial L}{\partial x_F}\right)_{(x^o;w^o)} =$$

$$= \left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0;w_F=0)} - C^T \lambda^o - (C')^T \mu^o - \tau^o = 0 \qquad (3)$$

$$\left(\frac{\partial L}{\partial w_F}\right)_{(x^o;w^o)} =$$

$$= \left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0;w_F=0)} - D^T \lambda^o - (D')^T \mu^o = 0 \qquad (4)$$

2°) $\lambda_i^o \, (x_{b_i} - x_{b_i}^o) = 0 \qquad i = 1,\dots,p$

$\mu_j^o \, (w_{b_j} - w_{b_j}^o) = 0 \qquad j = 1,\dots,h$

$\nu_i^o \, x_{b_i}^o = 0 \qquad i = 1,\dots,p$

$\tau_k^o \, (x_{f_k} = 0) = 0 \qquad k = 1,\dots,n-p$

De (1) $\qquad \lambda^o = \nu^o$

De (2) $\qquad \mu^o = 0$

De (3) y de $\mu^o = 0 \qquad \Rightarrow$

$$\Rightarrow \quad \left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0;w_F=0)} = C^T \lambda^o + \tau^o \qquad (5)$$

De (4) y de $\mu^{o} = 0 \quad \Longrightarrow$

$$\Longrightarrow \left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0;\, w_F=0)} = D^T \lambda^{o} \qquad (6)$$

Pero, en (5) y (6), por cumplirse, por hipótesis,

$$\left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0;\, w_F=0)} \geqslant 0 \quad y \quad \left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0;\, w_F=0)} = 0$$

ha de ser:

$$C^T \lambda^{o} + \tau^{o} \geqslant 0 \qquad (7)$$

$$D^T \lambda^{o} = 0 \qquad (8)$$

Tomando $\lambda^{o} = 0$ en (8), se obtiene en (7): $\tau^{o} \geqslant 0$.

Se comprueba, además, que para tales valores también se verifica la segunda condición de Kuhn y Tucker.

Luego, si $\left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0;\, w_F=0)} \geqslant 0$ y $\left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0;\, w_F=0)} = 0$, hemos encontrado que se verifican en el punto $(x^{o};\, w^{o})^T$ las condiciones de Kuhn y Tucker, siendo $\lambda^{o} = \nu^{o} = \mu^{o} = 0$ y $\tau^{o} \geqslant 0$. Por tanto, concluímos que $(x^{o};\, w^{o})^T \in Y^{(k}$ es una solución óptima para el programa $(PP^{(k})$.

Queda, así, demostrado el teorema 2.4.2.5.

Nota 2.4.2.3.

Según se ha indicado anteriormente, formulamos la si guiente :

Proposición 2.4.2.1.

Dado el programa del teorema 2.4.2.5. $(PP^{(k)})$, sea $(x^o;\ w^o)^T \in S_k$ si:

$\exists j \in \{1,..,n-p\}$ tal que $\left(\dfrac{\partial p^{(k}}{\partial x_{f_j}}\right)_{(x_F=0;\,w_F=0)} < 0$ ó

$\exists t \in \{1,..,k-h\}$ tal que $\left(\dfrac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(x_F=0;\,w_F=0)} \neq 0$, entonces la solución se puede mejorar.

Demostración

1) Supongamos que $\left(\dfrac{\partial p^{(k}}{\partial x_{f_j}}\right)_{(0)} < 0$ para algún $j \in \{1,..,n-p\}$

Avanzaríamos en la dirección x_{f_j} con $x_{f_j} > 0$.

Sea $x_{f_j} = \varepsilon > 0$

Como:

$$x_{b_i} = x^o_{b_i} + \sum_{k=1}^{n-p} c_{ik}\, x_{f_k} + \sum_{t+h+1}^{q} d_{it} w_{f_t} \qquad i = 1,..,p$$

$$w_{b_s} = w^o_{b_s} + \sum_{k=1}^{n-p} c'_{sk}\, x_{f_k} + \sum_{t=1}^{q} d'_{st} w_{f_t} \qquad s = 1,..,h$$

Llegaríamos a la nueva solución:

$$x_{f_j} = \varepsilon > 0$$

$$x_{b_i} = x^o_{b_i} + c_{ij}\,\varepsilon \qquad i = 1,...,p$$

$$w_{b_s} = w^o_{b_s} + c'_{sj}\,\varepsilon \qquad s = 1,...,h$$

Y puede ocurrir:

a) Que todos los $c_{ij} \geq 0$. Entonces, la solución obtenida es factible y mejora la solución $(x^o; w^o)^T$.

b) Que algún $c_{ij} < 0$.

Como $x_{b_i} = x^o_{b_i} + c_{ij}\,\varepsilon$ ha de ser positivo, bastaría tomar $\varepsilon < \min_i \left\{ - \dfrac{x^o_{b_i}}{c_{ij}} \right\}$ para que la solución obtenida fuese factible y se mejoraría la solución de partida.

2) Supongamos que $\left(\dfrac{\partial p^{(k}}{\partial x_F}\right)_{(0)} \geq 0$ y que $\exists t \in \{1,..,k-h\}$ tal que $\left(\dfrac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} \neq 0$.

Puede suceder:

a) $\left(\dfrac{\partial p^{(k}}{\partial w_{f_t}}\right)_{(0)} > 0$ para algún $t \in \{1,...,k-h\}$

Avanzaríamos en la dirección w_{f_t}, con $w_{f_t} < 0$.

La nueva solución sería:

$$x_{b_i} = x^o_{b_i} + d_{it} w_{f_t} \qquad i = 1,...,p$$

$$w_{b_s} = w^o_{b_s} + d'_{st} w_{f_t} \qquad s = 1,...,h$$

$$w_{f_t} = \varepsilon < 0$$

Para que tal solución sea factible, habría de ocurrir que $x_{b_i} = x^o_{b_i} + d_{it}\,\varepsilon \geq 0 \qquad i = 1,...,p$

- Si $d_{it} < 0 \quad \forall i$, como $\varepsilon < 0$, concluímos que la solución obtenida es factible y mejora la solución inicial.

- Si $d_{it} > 0$ para algún i, bastaría tomar

$$\underset{d_{it}>0}{\text{máx.}} \left\{ - \frac{x^o_{b_i}}{d_{it}} \right\}$$ para que la nueva solución obtenida sea factible y mejore la solución inicial.

b) Por último, si $\left(\frac{\partial P^{(k}}{\partial w_{f_t}} \right)_{(0)} < 0$ para algún $t \in \{1,\ldots,k-h\}$, razonando de manera análoga a como se ha hecho en 2a), concluiríamos que la solución nueva obtenida es factible y mejora la solución $(x^o; w^o)^T$.

Queda probada la proposición 2.4.2.1.

Convergencia del algoritmo

Vamos a demostrar que el algoritmo formulado converge al óptimo.

Pueden presentarse dos posibilidades:

a) Se llega al óptimo en un número finito de pasos, es decir, $\exists\, k \in N$ tal que $P(x^k) = P(x^*)$, siendo x^* una solución óptima de (PP^o).

b) No se llega en un número finito de casos.

En este último caso, vamos a demostrar el siguiente:

Teorema 2.4.2.6.

$$\lim_{k \to +\infty} P(x^k) = P(x^*)$$

siendo P(x) la función polinómica objetivo, $\{x^k\}_{k\in N}$ la sucesión de puntos, en las variables iniciales x_i, obtenida aplicando el algoritmo y x^* una solución óptima de $(PP^{(o)})$.

Demostración

a) $\{P(x^k)\}_{k\in N}$ es monótona decreciente.

En efecto, por la forma de obtener los x^k, al aplicar el algoritmo, se disminuye, en cada iteración, el valor de la función objetivo. Luego,

$$\forall k \in N \quad \text{es} \quad P(x^{k+1}) \leqslant P(x^k)$$

b) $\{P(x^k)\}_{k\in N}$ está acotada inferiormente por $P(x^*)$.

En efecto, como $P(x^*)$ es el valor mínimo de la función objetivo en todos los programas del algoritmo, se concluye que:

$$\forall k \in N \quad \text{es} \quad P(x^k) \geqslant P(x^*)$$

De las propiedades a) y b), se concluye que $\{P(x^k)\}_{k\in N}$ es una sucesión de números reales monótona decreciente y acotada inferiormente por $P(x^*)$. Luego, $\exists \lim\limits_{k\to+\infty} P(x^k)$

Probemos, finalmente, que $\lim\limits_{k\to+\infty} P(x^k) = P(x^*)$

De acuerdo con el algoritmo formulado, se prosigue la iteración hasta que se llega a un punto $(\bar{x}; \bar{w})$ que verifique el teorema 2.4.2.5., y, por tanto, $\bar{x} \in Y^{(o}$ es una solución óptima del programa inicial. Luego,

$$\lim_{k \to +\infty} P(x^k) = P(\bar{x}) = P(x^o)$$

Finaliza, así, la demostración del teorema.

2.4.3. Propiedad adicional del algoritmo para el caso cuadrático

Al aplicar el algoritmo establecido por nosotros al caso cuadrático, se obtiene el formulado por Beale [7] .

En efecto, supongamos que la función objetivo es polinómica de segundo grado, que pasamos del programa $(PP^{(k})$ al $(PP^{(k+1})$, que la variable añadida es z_{f_i} $(z_{f_i} = x_{f_i}$ ó $z_{f_i} = w_{f_i})$ y que la función objetivo, expresada a través de las variables libres, es: $P^{(k}(z_{f_1},\ldots,z_{f_i},\ldots,z_{f_t})$.

Entonces, de acuerdo con las notaciones introducidas en 2.4.2.1., la ecuación a añadir es:

$$\nabla\left(\frac{\partial P^{(k}}{\partial z_{f_i}}\right)^T_{(z_{f_1}=0,\ldots,z_{f_i}=z_{f_i}^{(k+1},\ldots,z_{f_t}=0)} \begin{pmatrix} z_{f_1} \\ \vdots \\ z_{f_i} - z_{f_i}^{(k+1} \\ \vdots \\ z_{f_t} \end{pmatrix} = w_{k+1} \quad (1$$

Ahora bien, por ser $P^{(k}$ cuadrática, se sigue que $\frac{\partial P^{(k}}{\partial z_{f_i}}$ es lineal y, como $\left(\frac{\partial P^{(k}}{\partial z_{f_i}}\right)_{(z_{f_i}^{(k+1};\ 0)} = 0$, resulta:

$$\frac{\partial P^{(k}}{\partial z_{f_i}} = h_{k+1,f_1}\, z_{f_1} + \ldots + h_{k+1,f_i}(z_{f_i} - z_{f_i}^{(k+1)}) + \ldots + h_{k+1,f_t}\, z_{f_t} \qquad (2)$$

Luego,

$$\nabla\left(\frac{\partial P^{(k}}{\partial z_{f_i}}\right)^T = (h_{k+1,f_1}, \ldots, h_{k+1,f_i}, \ldots, h_{k+1,f_t})$$

Y la ecuación (1) quedará:

$$h_{k+1,f_1}\, z_{f_i} + \ldots + h_{k+1,f_1}(z_{f_i} - z_{f_i}^{(k+1)}) + \ldots + h_{k+1,f_t}\, z_{f_t} = w_{k+1} \qquad (3)$$

Y, comparando (2) y (3), resulta:

$$\frac{\partial P^{(k}}{\partial z_{f_i}} = w_{k+1}$$

Es decir,

$$\nabla\left(\frac{\partial P^{(k}}{\partial z_{f_i}}\right)^T_{(z_{f_1}=0, \ldots, z_{f_i}=z_{f_i}^{(k+1}, \ldots, z_{f_t}=0)} \begin{pmatrix} z_{f_1} \\ \vdots \\ z_{f_i} - z_{f_i}^{(k+1} \\ \vdots \\ z_{f_t} \end{pmatrix} = \frac{\partial P^{(k}}{\partial z_{f_i}}$$

El algoritmo para el caso cuadrático tiene una propiedad muy importante, que ha sido formulada por Beale, en la referencia antes citada, y que podemos expresar mediante el siguiente:

Teorema 2.4.3.1.

Dado el programa:

$$(P^{(o)}) \begin{cases} \text{Mín. } P(x) \\ \text{bajo las restricciones:} \\ Y^{(o} \begin{cases} A\,x = b \\ x \geq 0 \\ x \in R^n \end{cases} \end{cases}$$

siendo $P(x)$ función polinómica de segundo grado, convexa en $Y^{(o}$, $A \in M(m,n)$, $b \in M(m,1)$ y ambas matrices de números reales. Entonces, el algoritmo formulado por Beale aplicado al programa $(P^{(o)})$ es finito.

Nota 2.4.3.1.

Debido a la finitud del algoritmo para el caso cuadrático, todos los teoremas vistos en el subapartado 2.4.2. son de aplicación aquí, a los programas inicial $(P^{(o)})$ y final $(P^{(E)})$, o programa en el que se alcanza la solución óptima, ya que en este caso $\exists k \in N$ tal que $(P^{(k)}) = (P^{(E)})$.

2.4.4. Resolución de un ejemplo

Consideramos, a continuación, un ejemplo con función objetivo polinómica de tercer grado.

Sea el programa convexo:

$$
(PP^{(o)})\begin{cases}
\text{Min. } P(x_1,x_2,x_3,x_4) = x_1^3 + x_2^3 + 2x_3^2 + 2x_4^2 \\
\text{bajo las restricciones:} \\
Y^{(o}\begin{cases}
x_1 + x_2 + x_3 + x_4 = 1 \\
\qquad\quad x_3 + x_4 = \frac{1}{2} \\
\qquad\quad x_i \geqslant 0 \qquad i = 1, 2, 3, 4 \\
(x_1, x_2, x_3, x_4) \in R^4
\end{cases}
\end{cases}
$$

Elegimos como primera solución básica factible: $(x_1 = \frac{1}{2}, x_2 = 0, x_3 = \frac{1}{2}, x_4 = 0) = x^{(o}$. El valor de la función objetivo en tal punto es $P(x^{(o)}) = \frac{5}{8}$.

Las variables básicas en función de las libres, son:

$$
\begin{aligned}
x_1 &= \tfrac{1}{2} - x_2 \\
x_3 &= \tfrac{1}{2} - x_4
\end{aligned} \qquad (1)
$$

La función objetivo, expresada a través de las libres, queda:

$$P^{(o}(x_2, x_4) = (\tfrac{1}{2} - x_2)^3 + x_2^3 + 2(\tfrac{1}{2} - x_4)^2 + 2x_4^2$$

Veamos si es posible avanzar en alguna dirección libre:

$$\frac{\partial P^{(o}(x_2,x_4)}{\partial x_2} = -3(\tfrac{1}{2} - x_2)^2 + 3x_2^2 \qquad y$$

$$\frac{\partial P^{(o}(x_2,x_4)}{\partial x_4} = -4(\tfrac{1}{2} - x_4) + 4x_4$$

Particularizando en el punto $(x_2=0,\ x_4=0)$,

$$\left(\frac{\partial P^{(o}}{\partial x_2}\right)_{(0,0)} = -\frac{3}{4} < 0 \quad y \quad \left(\frac{\partial P^{(o}}{\partial x_4}\right)_{(0,0)} = -2 < 0$$

Por tanto, el descenso más rápido se produce en la dirección x_4.

Obtengamos el valor de x_4 hasta el que se puede avanzar:

$$\left(\frac{\partial P^{(o}}{\partial x_4}\right)_{(x_2=0,x_4)} = -2 + 8x_4 = 0 \;\Rightarrow\; x_4 = \frac{1}{4}\ .\ \text{Como}$$

$x_2 = 0$, sustituyendo estos valores en (1):

$$x_1 = \frac{1}{2} \quad , \quad x_3 = \frac{1}{4}$$

Se llega, así, a la solución factible no básica de $(PP^{(o)})$: $(x_1 = \frac{1}{2},\ x_2 = 0,\ x_3 = \frac{1}{4},\ x_4 = \frac{1}{4})$, con valor: $P\ (\frac{1}{2},\ 0,\ \frac{1}{4},\ \frac{1}{4}) = \frac{3}{8}$.

Para conseguir que la solución obtenida se convierta en vértice, añadimos la nueva ecuación con la variable w_1 sin restricción en el signo:

$$\nabla\left(\frac{\partial P^{(o}}{\partial x_4}\right)^T_{(x_2=0,\ x_4=\frac{1}{4})} \begin{pmatrix} x_2 \\ x_4 - \frac{1}{4} \end{pmatrix} = w_1 \quad \Leftrightarrow$$

$$\Leftrightarrow \quad (0 \quad 8)\begin{pmatrix} x_2 \\ x_4 - \frac{1}{4} \end{pmatrix} = w_1 \Leftrightarrow 8(x_4 - \frac{1}{4}) = w_1$$

Resulta el nuevo programa:

$$
(PP^{(1)}\left\{\begin{array}{l}
\text{Min. } P(x_1,x_2,x_3,x_4) \\
\text{bajo las restricciones:} \\
Y^{(1}\left\{\begin{array}{ll}
x_1+x_2+x_3+x_4 & = 1 \\
x_3+x_4 & = \frac{1}{2} \\
8x_4-w_1 & = 2 \\
x_i \geqslant 0 & i = 1,2,3,4 \\
(x;\ w_1) \in R^5 &
\end{array}\right.
\end{array}\right.
$$

El punto $(x_1= \frac{1}{2},\ x_2=0,\ x_3= \frac{1}{4},\ x_4= \frac{1}{4},\ w_1=0)$ es vértice de $Y^{(1}$, con $P(\frac{1}{2},\ 0,\ \frac{1}{4},\ \frac{1}{4}) = \frac{3}{8}$.

En este programa, reiteramos la metodología del algoritmo.

Las variables básicas en función de las libres, quedan:

$$
\begin{aligned}
x_1 &= \frac{1}{2} - x_2 \\
x_3 &= \frac{1}{4} - \frac{w_1}{8} \qquad\qquad (2) \\
x_4 &= \frac{1}{4} + \frac{w_1}{8}
\end{aligned}
$$

La función objetivo, expresada a través de las libre queda:

$$
P^{(1}(x_2,w_1) = (\frac{1}{2} - x_2)^3 + x_2^3 + 2(\frac{1}{4} - \frac{w_1}{8})^2 + 2(\frac{1}{4} + \frac{w_1}{8})^2
$$

Comprobemos si se puede avanzar en alguna de las direcciones libres:

$$\frac{\partial P^{(1}(x_2,w_1)}{\partial x_2} = -3\left(\frac{1}{2}-x_2\right)^2 + 3x_2^2 \qquad y$$

$$\frac{\partial P^{(1}(x_2,w_1)}{\partial w_1} = \frac{1}{8}\,w_1$$

Luego,

$$\left(\frac{\partial P^{(1}}{\partial x_2}\right)_{(x_2=0,w_1=0)} = -\frac{3}{4} \quad y \quad \left(\frac{\partial P^{(1}}{\partial w_1}\right)_{(x_2=0,w_1=0)} = 0$$

Por tanto, avanzamos en la dirección x_2. Veamos hasta qué valor:

$$\left(\frac{\partial P^{(1}}{\partial x_2}\right)_{(x_2,w_1=0)} = -\frac{3}{4} + 3x_2 = 0 \implies x_2 = \frac{1}{4}$$

Como $w_1 = 0$, en (2) resulta:

$$x_1 = \frac{1}{4}, \quad x_3 = \frac{1}{4}, \quad x_4 = \frac{1}{4}$$

Se obtiene la solución factible no básica de $(PP^{(1})$:

$$(x_1 = \frac{1}{4},\ x_2 = \frac{1}{4},\ x_3 = \frac{1}{4},\ x_4 = \frac{1}{4},\ w_1 = 0) \qquad \text{con}$$

$$P\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) = \frac{9}{32}$$

Como la solución obtenida no es vértice de $Y^{(1}$, añadimos la ecuación con la nueva variable w_2 sin restricción en el signo:

$$\nabla\left(\frac{\partial P^{(1}}{\partial x_2}\right)^T_{(x_2=\frac{1}{4},w_1=0)}\begin{pmatrix} x_2-\frac{1}{4}\\ w_1\end{pmatrix}=w_2$$

Operando:

$$(3 \quad 0)\begin{pmatrix} x_2-\frac{1}{4}\\ w_1\end{pmatrix}=w_2 \quad \Longleftrightarrow \quad 3(x_2-\tfrac{1}{4})=w_2$$

Llegamos, así, al programa:

$$(PP^{(2)})\begin{cases} \text{Min. } P(x_1,x_2,x_3,x_4)\\ \text{bajo las restricciones:}\\ Y^{(2}\begin{cases} x_1+x_2+x_3+x_4 = 1\\ x_3+x_4 = \frac{1}{2}\\ 8x_4-w_1 = 2\\ 3x_2 \qquad -w_2 = \frac{3}{4}\\ x_i \geqslant 0 \quad i=1,2,3,4\\ (x;w_1,w_2)\in R^6\end{cases}\end{cases}$$

El punto $(x_1=\frac{1}{4},\ x_2=\frac{1}{4},\ x_3=\frac{1}{4},\ x_4=\frac{1}{4},\ w_1=0,\ w_2=0)$ es una solución básica factible de $(PP^{(2)})$, con

$P(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4})=\frac{9}{32}$

Reiteramos la metodología del algoritmo para el programa $(PP^{(2)})$.

Expresamos las variables básicas en función de las libres, quedando:

$$x_1 = \frac{1}{4} - \frac{1}{3} w_2$$

$$x_2 = \frac{1}{4} + \frac{1}{3} w_2$$

$$x_3 = \frac{1}{4} - \frac{1}{8} w_1$$

$$x_4 = \frac{1}{4} + \frac{1}{8} w_1$$

Análogamente, la función objetivo a través de las li bres, queda:

$$P^{(2}(w_1,w_2) = \left(\frac{1}{4} - \frac{1}{3} w_2\right)^3 + \left(\frac{1}{4} + \frac{1}{3} w_2\right)^3 + $$
$$+ 2\left(\frac{1}{4} - \frac{1}{8} w_1\right)^2 + 2\left(\frac{1}{4} + \frac{1}{8} w_1\right)^2$$

Comprobemos si es posible avanzar en alguna dirección libre:

$$\frac{\partial P^{(2}(w_1,w_2)}{\partial w_1} = \frac{1}{8} w_1 \quad \text{y} \quad \frac{P^{(2}(w_1,w_2)}{w_2} = \frac{1}{3} w_2$$

$$\text{Como:} \left(\frac{\partial P^{(2}}{\partial w_1}\right)_{(w_1=0,w_2=0)} = 0 \quad \text{y} \quad \left(\frac{\partial P^{(2}}{\partial w_2}\right)_{(w_1=0,w_2=0)} = 0$$

concluímos, por el teorema 2.4.2.5., que hemos llegado al óptimo en $(PP^{(2})$, con punto de óptimo:

$$(x_1 = \frac{1}{4},\ x_2 = \frac{1}{4},\ x_3 = \frac{1}{4},\ x_4 = \frac{1}{4},\ w_1 = 0,\ w_2 = 0)$$

Por tanto, el punto de óptimo en el programa inicial $(PP^{(o})$ es: $(x_1 = \frac{1}{4},\ x_2 = \frac{1}{4},\ x_3 = \frac{1}{4},\ x_4 = \frac{1}{4})$ y el valor óptimo:

$$P\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) = \frac{9}{32}$$

Nota 2.4.4.1.

Al pasar de un programa a otro añadiendo una nueva ecuación lineal con una nueva variable sin restricción en el signo, dicha variable nueva no puede entrar en la base en el paso siguiente.

En efecto, supongamos que se pasa del programa $(PP^{(k})$ al $(PP^{(k+1})$, $k \in N$, añadiendo la ecuación lineal:

$$\nabla \left(\frac{\partial P^{(k}}{\partial z_{f_1}} \right)^T_{(z_{f_1} = z_{f_1}^{(k+1}; 0)} \begin{pmatrix} z_{f_1} - z_{f_1}^{(k+1} \\ \cdots\cdots\cdots \\ \bar{z}_F \end{pmatrix} = w_{k+1}$$

Y, por tanto, será:

$$\left(\frac{\partial P^{(k}}{\partial z_{f_1}} \right)_{(z_{f_1} = z_{f_1}^{(k+1}\ ;\ 0)} = 0 \qquad (1)$$

donde la variable z_j representa a la variable x_j o a la w_j.

Supongamos, además, que pasamos de la solución básica factible de $(PP^{(k})$: $(z_B^{(k})$ a la solución básica factible de $(PP^{(k+1})$: $(z_B^{(k+1}\ ;\ z_{f_1}^{(k+1})$. Sea la nueva función objetivo, expresada a través de las variables libres, $P^{(k+1}(\bar{z}_F; w_{k+1})$ donde $\bar{z}_F$ representa el conjunto de variables libres que había en la función $P^{(k}$ excluída la variable z_{f_1} que ha pasado a formar parte de la nueva base.

Calculemos el valor de $\left(\dfrac{\partial p^{(k+1}}{\partial w_{k+1}}\right)_{(\bar{z}_F=0;\,w_{k+1}=0)}$

$$\frac{\partial p^{(k+1}}{\partial w_{k+1}} = \frac{\partial p^{(k}}{\partial z_{f_1}} \quad \frac{\partial z_{f_1}}{\partial w_{k+1}} \qquad (2)$$

Y, al particularizar $\dfrac{\partial p^{(k+1}}{\partial w_{k+1}}$ en el punto $(\bar{z}_F=0;w_{k+1}=0)$ el punto correspondiente a $\dfrac{\partial p^{(k}}{\partial z_{f_1}}$ es $(\bar{z}_F=0;z_{f_1}=z_{f_1}^{(k+1})$. Luego, en (2):

$$\left(\frac{\partial p^{(k+1}}{\partial w_{k+1}}\right)_{(\bar{z}_F=0;\,w_{k+1}=0)} =$$

$$= \left(\frac{\partial p^{(k}}{\partial z_{f_1}}\right)_{(\bar{z}_F=0;\,z_{f_1}=z_{f_1}^{(k+1})} \left(\frac{\partial z_{f_1}}{\partial w_{k+1}}\right)_{(\bar{z}_F=0;\,z_{f_1}=z_{f_1}^{(k+1})} \overset{(1)}{=}$$

$$= 0 \left(\frac{\partial z_{f_1}}{\partial w_{k+1}}\right)_{(\bar{z}_F=0;\,z_{f_1}=z_{f_1}^{(k+1})} = 0$$

Es decir, la nueva variable w_{k+1} no puede entrar en la base en el paso siguiente.

El fundamento de la finitud del algoritmo para el caso cuadrático se basa en la propiedad anterior y, además, en que:

$$\nabla\left(\frac{\partial p^{(k}}{\partial z_{f_1}}\right)^T_{(z_{f_1}=z_{f_1}^{(k+1};\,\bar{z}_F=0)} \begin{pmatrix} z_{f_1} - z_{f_1}^{(k+1} \\ \cdots\cdots\cdots\cdots \\ \bar{z}_F \end{pmatrix} = \frac{\partial p^{(k}}{\partial z_{f_1}} \qquad (3)$$

El razonamiento no es extensible al caso no cuadrático debido a que la igualdad (3) no se verifica, en general, más que en el punto $(z_{f_1} = z_{f_1}^{(k+1} \; ; \; \bar{z}_F = 0)$.

2.5. EXTENSION DEL ALGORITMO A LA PROGRAMACION POLINOMICA EN GENERAL

El algoritmo formulado por nosotros, se puede extender a la Programación Polinómica no Convexa.

Dado el programa:

$$(P) \begin{cases} \text{Min } P(x) \\ \text{bajo las restricciones:} \\ A\,x = b \\ \quad x \geqslant 0 \\ \quad x \in R^n \end{cases}$$

siendo $P(x)$ función polinómica, $A \in M(m,n)$, $rg(A) = m$, $b \in M(m,1)$ y ambas matrices en R.

Al aplicar el citado algoritmo a este programa nos conduce, en el caso finito, a un punto x^o que verifica:

$$\left(\frac{\partial P^{(k}}{\partial x_F}\right)_{(x_F=0;w_F=0)} \geqslant 0 \quad \text{y} \quad \left(\frac{\partial P^{(k}}{\partial w_F}\right)_{(x_F=0;w_F=0)} = 0$$

(es decir, las condiciones de Kuhn y Tucker según hemos demostrado en el teorema 2.4.2.5.).

Si además, en tal punto x^o existe un entorno de x^o en el cual $P(x)$ es convexa, entonces x^o es un óptimo local, de acuerdo con el siguiente teorema tomado de Minoux, M. [40], Tome 1, p. 189.

<u>Teorema 2.5.1.</u>

Sea el programa:

$$(P')\begin{cases} \text{Min } F(x) \\ \text{bajo las restricciones:} \\ g_i(x) \leqslant 0 \qquad i \in I = \{1,\ldots,m\} \\ \quad x \in R^n \end{cases}$$

donde $F(x)$ y $g_i(x)$ $i \in I$ son funciones diferenciables. Entonces, una condición suficiente para que $\bar{x}$ sea un óptimo local de (P') es que:

a) Exista un entorno de $\bar{x}$ en el cual las funciones $F(x)$ y $g_i(x)$ $i \in I$ sean convexas.

b) Se verifiquen en $\bar{x}$ las condiciones de Kuhn y Tucker.

C A P I T U L O 3

LA PROGRAMACION POLINOMICA COMO UN PRO-BLEMA DE PROGRAMACION MULTIOBJETIVO

3.1. INTRODUCCION

En este capítulo, se presenta un enfoque nuevo de la Programación polinómica, transformándola en un problema equivalente de Programación multiobjetivo. Para ello, en el apartado 3.2., establecemos cuáles han de ser las funciones multiobjetivo y cuál la relación de preferencia ($\succ$ "más preferido que") en el conjunto factible Z del programa multiobjetivo ($Z \subset R^k$, $k \in N^*$, es el transformado del conjunto Y de las restricciones del programa inicial, mediante las k funciones multiobjetivo definidas). Se comprueba que hallar el óptimo en el programa inicial es equivalente a encontrar los elementos más preferidos mediante la relación $\succ$ en Z (tales elementos, como veremos posteriormente, se denominan eficientes).

Se establecen, en la sección 3.3. (teorema 3.3.1.), las propiedades de la relación $\succ$.

Asociadas a la relación anterior, se encuentran otras dos: $\sim$ ($\sim$ "indiferente a") y $\succsim$ ($\succsim$ "más preferido o indiferente que") que son tratadas en 3.4., dándose su definición y estableciéndose sus propiedades.

En el apartado 3.5., se definen el conjunto eficiente y la estructura de dominación relativos a la relación $\succ$.

Se tratan seguidamente, sección 3.6., las propiedades de la estructura de dominación D, probándose en el teorema 3.6.1. qué elementos de R^k forman el conjunto D, en el teorema 3.6.4. se demuestra que D es acíclica y en el teore ma 3.6.5. se establece que D es un cono convexo y apuntado.

Finalmente, en el epígrafe 3.7., se prueba la existencia de soluciones eficientes y se da la caracterización de las mismas. Se demuestra, en el teorema 3.7.1., que el conjunto de elementos eficientes es el transformado de las soluciones óptimas del programa polinómico inicial; en el teorema 3.7.2., se prueba la existencia de soluciones óptimas para el programa inicial, y como consecuencia de los teoremas 3.7.1. y 3.7.2. se establece, en el teorema 3.7.3., la existencia de soluciones eficientes para el problema de programación multiobjetivo formulado. A continuación, damos la caracterización de las soluciones eficientes a través del cono polar positivo estricto D^{so} de la estructura de dominación D (previamente probamos cuál es la expresión de los conos polar positivo D^{o} y polar positivo estricto D^{so}, teoremas 3.7.4. y 3.7.5.). Se demuestra, en el teorema 3.7.7., que los elementos eficientes de Z son aquellos tales que su producto escalar con los de D^{so} es mínimo. Debido a la sencilla expresión de D^{so}, se profundiza aún más sobre el conjunto de elementos eficientes $\mathcal{E}(Z;\ D)$ y sobre su ubicación (teorema 3.7.8.). Por último, en el teorema 3.7.9., se establece que los elementos eficientes están contenidos en la frontera de Z.

El estudio realizado, en este capítulo, se ha formulado para el problema de mínimo, un tratamiento análogo se haría para el caso de máximo.

3.2. PLANTEAMIENTO DEL PROBLEMA

Sea el programa:

$$(I) \begin{cases} \text{Min } P(x) \\ \text{bajo las restricciones:} \\ x \in Y \subset R^n \end{cases}$$

siendo $P(x)$ función polinómica en Y e Y conjunto compacto de R^n para la topología usual.

La función polinómica $P(x)$, la podemos expresar como suma de k funciones polinómicas:

$$P(x) = P_1(x) + \ldots + P_k(x)$$

en las que cada $P_i(x)$, $i = 1,\ldots,k$, no depende, en general, más que de algunas de las variables $x_1,\ldots,x_n$.

Consideremos la aplicación $T : Y \rightarrow R^k$ definida así:

$$T(x) = (P_1(x), \ldots, P_k(x)) \qquad \forall x \in Y$$

Sea $\quad Z = Im(T) = T(Y) =$

$$= \{(z_1,\ldots,z_k) \in R^k \;/\; \exists x \in Y \text{ tal que } T(x) = (P_1(x),\ldots,P_k(x)) = (z_1,\ldots,z_k)$$

Definición 2.3.1.

En Z establecemos la siguiente relación binaria ($\succ$ "más preferido que"):

Dados $(z_1,\dots,z_k)$, $(z_1',\dots,z_k') \in Z$:

$$(z_1,\dots,z_k) \succ (z_1',\dots,z_k') \Longleftrightarrow z_1+\dots+z_k < z_1'+\dots+z_k'$$

Nota 3.2.1.

Dado el programa (I), si llamamos S_o al conjunto de soluciones poptimas del mismo, se verifica que:

$$x^o \in S_o \Longleftrightarrow \forall x \in Y-S_o: \quad P(x^o) < P(x) \Longleftrightarrow$$

$$\Longleftrightarrow \forall x \in Y-S_o: \quad P_1(x^o)+\dots+P_k(x^o) < P_1(x)+\dots+P_k(x)$$

$$\underset{\uparrow}{\Longleftrightarrow} \quad (P_1(x^o),\dots,P_k(x^o)) \succ (P_1(x),\dots,P_k(x) \quad \forall x \in Y-S_o$$

por definición de $\succ$

$$\Longleftrightarrow \quad (z_1^o,\dots,z_k^o) \succ (z_1,\dots,z_k) \quad \forall (z_1,\dots,z_k) \in Z-T(S_o)$$

$$\text{con } (z_1^o,\dots,z_k^o) = T(x^o)$$

$$\Longleftrightarrow \nexists (z_1,\dots,z_k) \in Z \text{ tal que } (z_1,\dots,z_k) \succ (z_1^o,\dots,z_k^o)$$

$$\text{con } (z_1^o,\dots,z_k^o) = T(x^o)$$

Luego, hallar el óptimo de (I) es equivalente a obtener los elementos más preferidos para la relación $\succ$ en el conjunto Z o equivalentemente obtener aquellos elementos tales que no existan otros que sean más preferidos que ellos.

Por tanto, el programa (I) es equivalente al siguiente problema de programación multiobjetivo:

$$\text{(II)} \begin{cases} \text{Min } (P_1(x),\dots,\ P_k(x)) \\ \text{bajo las restricciones:} \\ x \in Y \subset R^n \end{cases}$$

entendiendo por $\text{Min}(P_1(x),\dots,P_k(x))$ el encontrar aquellos elementos de Z para los cuales no exista elemento alguno de

Z que sea más preferido que ellos respecto de la relación $\succ$, siendo $P_i(x)$ $i = 1,..,k$ e Y las funciones y el conjunto anteriormente especificados.

3.3. PROPIEDADES DE LA RELACION $\succ$

Definiciones previas 3.3.1. (Fishburn [18] y [19])

Una relación binaria $\succ$ definida en $C \subset R^p$ se dice que es:

i) Reflexiva si $\forall u \in C$ es $u \succ u$

ii) Irreflexiva si $\forall u \in C$ es $u \not\succ u$

iii) Simétrica si $u \succ v \Rightarrow v \succ u$ $\quad u,v \in C$

iv) Asimétrica si $u \succ v \Rightarrow v \not\succ u$ $\quad u,v \in C$

v) Antisimétrica si $u \succ v$ y $v \succ u$ $\quad u = v$ $\quad u,v \in C$

vi) Transitiva si $u \succ v$ y $v \succ w \Rightarrow u \succ w$ $\quad u,v,w \in C$

vii) Negativamente transitiva si $u \not\succ v$ y $v \not\succ w \Rightarrow u \not\succ w$ $\quad u,v,w \in C$

viii) Completa si $\forall u,v \in C$ es $u \succ v$ ó $v \succ u$

ix) Un orden parcial estricto si es irreflexiva y transitiva.

x) Un orden débil si es asimétrica y negativamente transitiva.

Teorema 3.3.1.

La relación $\succ$ definida en Z es irreflexiva, asimétrica, transitiva y negativamente transitiva. Por tanto, $\succ$ es un or den parcial estricto y un orden débil.

Demostración

1°) $\succ$ es irreflexiva

En efecto, $\forall (z_1, \ldots, z_k) \in Z$ como

$$z_1 + \ldots + z_k \nless z_1 + \ldots + z_k$$

ello equivale, por definición de $\succ$

$$(z_1, \ldots, z_k) \not\succ (z_1, \ldots, z_k)$$

2°) $\succ$ es asimétrica

En efecto, dados $(z_1, \ldots, z_k)$, $(z'_1, \ldots, z'_k) \in Z$ si

$$(z_1, \ldots, z_k) \succ (z'_1, \ldots, z'_k)$$

ello equivale, por definición de

$$z_1 + \ldots + z_k < z'_1 + \ldots + z'_k$$

lo que implica que:

$$z'_1 + \ldots + z'_k \nless z_1 + \ldots + z_k$$

y ello equivale, por definición de $\succ$:

$$(z'_1, \ldots, z'_k) \not\succ (z_1, \ldots, z_k)$$

3°) $\succ$ es transitiva

Efectivamente,

sean $(z_1, \ldots, z_k)$, $(z'_1, \ldots, z'_k)$, $(z''_1, \ldots, z''_k) \in Z$ tales que $(z_1,\dots,z_k) \succ (z'_1,\dots,z'_k)$ y $(z'_1,\dots,z'_k) \succ (z''_1,\dots,$

Ahora bien,

$$\left.\begin{array}{l} (z_1,\dots,z_k) \succ (z'_1,\dots,z'_k) \overset{\text{por definición de } \succ}{\Longleftrightarrow} z_1+\dots+z_k < z'_1+\dots+z'_k \\ (z'_1,\dots,z'_k) \succ (z''_1,\dots,z''_k) \underset{\text{por definición de } \succ}{\Longleftrightarrow} z'_1+\dots+z'_k < z''_1+\dots+z''_k \end{array}\right\} \Rightarrow$$

$$\Rightarrow \quad z_1+\dots+z_k < z''_1+\dots+z''_k$$

ello equivale, por definición de $\succ$ que:

$$(z_1,\dots,z_k) \succ (z''_1,\dots,z''_k)$$

4°) <u>$\succ$ es negativamente transitiva</u>

Sean $(z_1,\dots,z_k),\ (z'_1,\dots,z'_k),\ (z''_1,\dots,z''_k) \in Z$

tales que:

$(z_1,\dots,z_k) \not\succ (z'_1,\dots,z'_k)$ y $(z'_1,\dots,z'_k) \not\succ (z''_1,\dots,z''_k)$

Entonces:

$(z_1,\dots,z_k) \not\succ (z'_1,\dots,z'_k)$ ello equivale, por definición de

$z_1+\dots+z_k \nless z'_1+\dots+z'_k$ lo que es equivalente a:

$$\begin{cases} z_1+\dots+z_k = z'_1+\dots+z'_k & (1) \\ \text{ó} & \\ z_1+\dots+z_k > z'_1+\dots+z'_k & (2) \end{cases}$$

Análogamente,

$(z'_1,..,z'_k) \not\succ (z''_1,..,z''_k)$ equivale por definición de $\succ$

$z'_1+..+z'_k \nless z''_1+..+z''_k$ lo que es equivalente a:

$$\begin{cases} z'_1+..+z'_k = z''_1+..+z''_k & (3) \\ \text{ó} & \\ z'_1+..+z'_k > z''_1+..+z''_k & (4) \end{cases}$$

Si (1) y (3) $\Rightarrow$ $z_1+..+z_k \doteq z''_1+..+z''_k$ $\Rightarrow$

$\Rightarrow$ $z_1+..+z_k < z''_1+..+z''_k$ por definición de $\succ$ $\Leftrightarrow$

$\Leftrightarrow$ $(z_1,..,z_k) \not\succ (z''_1,..,z''_k)$

Si (1) y (4) $\Rightarrow$ $z_1+..+z_k > z''_1+..+z''_k$ $\Rightarrow$

$\Rightarrow$ $z_1+..+z_k \nless z''_1+..+z''_k$ por definición de $\succ$ $\Leftrightarrow$

$\Leftrightarrow$ $(z_1,..,z_k) \not\succ (z''_1,..,z''_k)$

Si (2) y (3) $\Rightarrow$ $z_1+..+z_k > z''_1+..+z''_k$ $\Rightarrow$

$\Rightarrow$ $z_1+..+z_k < z''_1+..+z''_k$ por definicidón de $\succ$ $\Leftrightarrow$

$\Leftrightarrow$ $(z_1,..,z_k) \not\succ (z''_1,..,z''_k)$

Si (2) y (4) $\Rightarrow$ $z_1+..+z_k > z''_1+..+z''_k$ $\Rightarrow$

$z_1+..+z_k \nless z''_1+..+z''_k$ por definición de $\succ$ $\Leftrightarrow$

$\Leftrightarrow$ $(z_1,..,z_k) \not\succ (z''_1,..,z''_k)$

Luego, $\succ$ es negativamente transitiva.

Según las definiciones ix) y x) anteriores, como $\succ$ es irreflexiva y transitiva, se sigue que $\succ$ es un orden parcial estricto y por ser $\succ$ asimétrica y transitiva, se concluye que $\succ$ es un orden débil.

3.4. DEFINICION Y PROPIEDADES DE $\sim$ Y $\succsim$

A partir de la relación de preferencia $\succ$ definida anteriormente en Z podemos establecer otras dos relaciones binarias en Z: $\sim$ y $\succsim$ ($\sim$ "indiferente a" y $\succsim$ "más preferido o indiferente que").

Definición 3.4.1.

Dados $(z_1,\ldots, z_k)$, $(z_1',\ldots, z_k') \in Z$, se dice que:

i) $(z_1,\ldots,z_k) \sim (z_1',\ldots,z_k')$ si y sólo si

$$\begin{cases} (z_1,\ldots,z_k) \not\succ (z_1',\ldots,z_k') & \text{y} \\ (z_1',\ldots,z_k') \not\succ (z_1,\ldots,z_k) \end{cases}$$

ii) $(z_1,\ldots,z_k) \succsim (z_1',\ldots,z_k')$ si y sólo si

$$\begin{cases} (z_1,\ldots,z_k) \succ (z_1',\ldots,z_k') & \text{ó} \\ (z_1,\ldots,z_k) \sim (z_1',\ldots,z_k') \end{cases}$$

Teorema 3.4.1.

La relación binaria $\sim$ definida en Z verifica las siguientes propiedades:

1°) Es de equivalencia

2°) $\forall (a_1,..,a_k) \in Z$:

$$[(a_1,..,a_k)] = \{(x_1,..,x_k) \in Z/\ x_1+..+x_k = c,\quad c = \sum_{i}^{k} a$$

$$= \{(x_1,..,x_k) \in R^k / \sum_{i=1}^{k} x_i = c,\quad c = \sum_{i=1}^{k} a_i\} \cap Z$$

siendo $[(a_1,..,a_k)]$ la clase de equivalencia de representante $(a_1,..,a_k) \in Z$.

3°) $Z_{/\sim} = \{ H \cap Z\ /\ H$ es un hiperplano de R^k de ecuación

$$\sum_{i=1}^{k} x_i = c,\quad c = \sum_{i=1}^{k} a_i \qquad (a_1,..,a_k) \in Z \}$$

Demostración

Veamos, previamente, que dados

$(x_1,..,x_k)$, $(x_1',..,x_k') \in Z$, entonces

$$(x_1,..,x_k) \sim (x_1',..,x_k') \Longleftrightarrow x_1+..+x_k = x_1'+..+x_k' \qquad (1)$$

En efecto, por definición de $\sim$:

$(x_1,..,x_k) \sim (x_1',..,x_k')$ si y sólo si

$(x_1,..,x_k) \not\succ (x_1',..,x_k')$ y $(x_1',..,x_k') \not\succ (x_1,..,x_k)$

Pero lo anterior, según la definición de $\succ$:

$$\left.\begin{array}{l}(x_1,..,x_k)\not\succ(x_1',..,x_k') \Leftrightarrow x_1+..+x_k \nless x_1'+..+x_k' \\ (x_1',..,x_k')\not\succ(x_1,..,x_k) \Leftrightarrow x_1'+..+x_k' \nless x_1+..+x_k\end{array}\right\} \Leftrightarrow$$

$$\Leftrightarrow \qquad x_1+....+x_k = x_1'+...+ x_k'$$

Luego,

$$(x_1,..,x_k) \sim (x_1',..,x_k') \Leftrightarrow x_1+...+x_k = x_1'+...+x_k'$$

Probemos, ahora, la primera proposición: $\succ$ es una relación de equivalencia.

a) $\sim$ <u>es reflexiva</u>

En efecto, $\forall (x_1,..,x_k) \in Z$ es

$(x_1,..,x_k) \sim (x_1,..,x_k)$ en virtud de la propiedad (1)

b) $\sim$ <u>es simétrica</u>

En efecto, dados $(x_1,..,x_k)$, $(x_1',..,x_k') \in Z$ si

$$(x_1,..,x_k) \sim (x_1',..,x_k') \overset{(1)}{\Leftrightarrow} x_1+..+x_k = x_1'+..+x_k' \Leftrightarrow$$

$$\Leftrightarrow \qquad (x_1',..,x_k') \sim (x_1,..,x_k)$$

c) $\sim$ <u>es transitiva</u>

Efectivamente, sean $(x_1,..,x_k)$, $(x_1',..,x_k')$, $(x_1'',..,x_k'') \in Z$ tales que:

$(x_1,..,x_k) \sim (x_1',..,x_k')$ y $(x_1',..,x_k') \sim (x_1'',..,x_k'')$

De acuerdo con la propiedad (1)

$$\left.\begin{array}{l}(x_1,..,x_k) \sim (x_1',..,x_k') \iff x_1+..+x_k = x_1'+..+x_k' \\ (x_1',..,x_k') \sim (x_1'',..,x_k'') \iff x_1'+..+x_k' = x_1''+..+x_k''\end{array}\right\} \Rightarrow$$

$$\Rightarrow x_1+..+x_k = x_1''+..+x_k'' \overset{(1)}{\iff} (x_1,..,x_k) \sim (x_1'',..,x_k'')$$

Luego, $\sim$ es una relación de equivalencia en Z.

Probemos, ahora, que $\forall(a_1,..,a_k) \in Z$:

$$[(a_1,..,a_k)] = \{(x_1,..,x_k) \in Z \ / \ x_1+..+x_k = c, \text{ con } c=\sum_{i=1}^{k} a_i\}$$

En efecto, por definición de clase de equivalencia:

$$[(a_1,..,a_k)] = \{(x_1,..,x_k) \in Z \ / \ (x_1,..,x_k) \sim (a_1,..,a_k)\}$$

Y por la propiedad (1):

$$(x_1,..,x_k) \sim (a_1,..,a_k) \iff x_1+..+x_k = a_1+..+a_k$$

Haciendo $c = \sum_{i=1}^{k} a_i$, $x_1+..+x_k = c$

Luego:

$$[(a_1,..,a_k)] = \{(x_1,..,x_k) \in Z \ / \ \sum_{i=1}^{k} x_i = c, \text{ con } c = \sum_{i=1}^{k} a_i\} =$$

$$= \{(x_1,..,x_k) \in R^k / \sum_{i=1}^{k} x_i = c, \text{ con } c = \sum_{i=1}^{k} a_i\} \cap Z$$

Demostremos, finalmente, que:

$$Z/_{\sim} = \left\{H \cap Z / \ H \text{ es un hiperplano de } R^k \text{ de ecuación} \sum_{i=1}^{k} x_i = c, \text{ con } c = \sum_{i=1}^{k} a_i \text{ y } (a_1,..,a_k) \in Z\right\}$$

Es una consecuencia de la proposición anterior, ya que $\forall(a_1,..,a_k) \in Z$ su clase de equivalencia es la intersección de Z con el hiperplano $\sum_{i=1}^{k} x_i = c$, siendo $c = \sum_{i=1}^{k} a_i$.

Teorema 3.4.2.

La relación binaria $\succsim$ definida en Z es reflexiva, transitiva, negativamente transitiva y completa.

Demostración

Probemos antes que dados

$(x_1,\ldots,x_k)$, $(x_1',\ldots,x_k') \in Z$, entonces

$$(x_1,\ldots,x_k) \succsim (x_1',\ldots,x_k') \Longleftrightarrow x_1+\ldots+x_k \leq x_1'+\ldots+x_k' \qquad (1)$$

En efecto, por definición de $\succsim$:

$$(x_1,\ldots,x_k) \succsim (x_1',\ldots,x_k') \Longleftrightarrow \left\{\begin{array}{c}(x_1,\ldots,x_k) \succ (x_1',\ldots,x_k') \\ \text{ó} \\ (x_1,\ldots,x_k) \sim (x_1',\ldots,x_k')\end{array}\right\} \Longleftrightarrow$$

definiciones de $\succ$
y de $\sim$

$$\Longleftrightarrow \left\{\begin{array}{c} x_1+\ldots+x_k < x_1'+\ldots+x_k' \\ \text{ó} \\ x_1+\ldots+x_k = x_1'+\ldots+x_k' \end{array}\right\} \Longleftrightarrow x_1+\ldots+x_k \leq x_1'+\ldots+x_k'$$

Pasamos a probar las propiedades enunciadas en el teorema para la relación $\succsim$.

a) $\succsim$ es reflexiva

En efecto, $\forall (x_1,\ldots,x_k) \in Z$, como

$x_1+\ldots+x_k = x_1+\ldots+x_k$, ello equivale, por una propiedad de la relación $\sim$

$$(x_1,..,x_k) \sim (x_1,..,x_k), \text{ de donde concluímos que:}$$

$$(x_1,..,x_k) \succsim (x_1,..,x_k)$$

b) <u>$\succsim$ es transitiva</u>

Sean

$(x_1,..,x_k),(x_1',..,x_k'),(x_1'',..,x_k'') \in Z$, tales que:

$(x_1,..,x_k) \succsim (x_1',..,x_k')$ y $(x_1',..,x_k') \succsim (x_1'',..,x_k'')$

Ahora bien,

$$\left.\begin{array}{l}(x_1,..,x_k) \succsim (x_1',..,x_k') \overset{(1)}{\Longleftrightarrow} x_1+..+x_k \leq x_1'+..+x_k' \\ (x_1',..,x_k') \succsim (x_1'',..,x_k'') \overset{(1)}{\Longleftrightarrow} x_1'+..+x_k' \leq x_1''+..+x_k''\end{array}\right\} \Rightarrow$$

$$\Rightarrow \quad x_1+..+x_k \leq x_1''+..+x_k'' \overset{(1)}{\Longleftrightarrow} (x_1,..,x_k) \succsim (x_1'',..,x_k'')$$

c) <u>$\succsim$ es negativamente transitiva</u>

En efecto,

sean $(x_1,..,x_k),(x_1',..,x_k'),(x_1'',..,x_k'') \in Z$, tales que

$(x_1,..,x_k) \not\succsim (x_1',..,x_k')$ y $(x_1',..,x_k') \not\succsim (x_1'',..,x_k'')$

Ahora bien:

$$\left.\begin{array}{l}(x_1,..,x_k) \not\succsim (x_1',..,x_k') \overset{(1)}{\Longleftrightarrow} x_1+..+x_k \not\leq x_1'+..+x_k' \\ (x_1',..,x_k') \not\succsim (x_1'',..,x_k'') \overset{(1)}{\Longleftrightarrow} x_1'+..+x_k' \not\leq x_1''+..+x_k''\end{array}\right\} \Longleftrightarrow$$

$$\Longleftrightarrow \left.\begin{array}{l}x_1+..+x_k > x_1'+..+x_k' \\ x_1'+..+x_k' > x_1''+..+x_k''\end{array}\right\} \Rightarrow x_1+..+x_k > x_1''+..+x_k'' \Longleftrightarrow$$

$$x_1+..+x_k \nleq x_1''+..+x_k'' \Longleftrightarrow (x_1,..,x_k) \not\succ (x_1'',..,x_k'')$$

d) $\succsim$ es completa

En efecto,

sean $\forall (x_1,\ldots,x_k)$, $(x_1',\ldots,x_k') \in Z$, entonces

$$x_1+..+x_k \leq x_1'+..+x_k' \quad \text{ó} \quad x_1'+..+x_k' \leq x_1+..+x_k \Longleftrightarrow$$

$$\Longleftrightarrow (x_1,..,x_k) \succsim (x_1',..,x_k') \quad \text{ó} \quad (x_1',..,x_k') \succsim (x_1,..,x_k)$$

3.5. ELEMENTOS EFICIENTES Y ESTRUCTURA DE DOMINACION RELATIVOS A $\succ$

Definición 3.5.1.

Sea la estructura de orden $(Z, \succ)$, un elemento $(\bar{x}_1,..,\bar{x}_k) \in Z$ se dice que es eficiente respecto de la relación $\succ$ si $\nexists (x_1,..,x_k) \in Z$ tal que $(x_1,..,x_k) \succ (\bar{x}_1,..,\bar{x}_k$

El conjunto de elementos eficientes será denotado así:

$$E(Z,\succ)=\{(\bar{x}_1,..,\bar{x}_k)\in Z/ \nexists (x_1,..,x_k)\in Z \text{ tal que } (x_1,..,x_k)\succ(\bar{x}_-$$

Nota 3.5.1.

Según la nota 3.2.1. y la definición 3.5.1., obtener el óptimo en el programa inicial (I) es equivalente a determinar los elementos eficientes de la estructura de orden $(Z, \succ)$.

Nota 3.5.2.

Si $\succ$ hubiera sido un orden en Z, entonces los elementos eficientes serían los elementos maximales de Z respecto de la relación $\succ$.

Nota 3.5.3.

La relación de preferencia $\succ$, puede expresarse a través del concepto de estructura de dominación dado por Yu [51].

Por otra parte, a efectos de simplificación de las demostraciones, definiremos la estructura de dominación en un conjunto que contenga a Z. Concretamente tomaremos R^k.

Además, consideraremos extendidas las definiciones de $\succ$, $\sim$, $\succsim$ a R^k tomando como definiciones de las mismas las dadas en 3.2. y 3.4.

Definición 3.5.2.

$\forall x=(x_1,\ldots,x_k)\in R^k$ llamaremos factores de dominación relativos a **x** al siguiente conjunto:

$$D(x) = \{d\in R^k / \ x \succ x+d\}\cup\{0\}$$

siendo $d = (d_1,\ldots,d_k)$ y $0 = (0,\ldots,0)\in R^k$

La aplicación de punto a conjunto:

$$\begin{aligned} D:\quad & R^k \longrightarrow \mathcal{P}(R^k) \\ & (x_1,\ldots,x_k) \longrightarrow D(x_1,\ldots,x_k) \end{aligned}$$

se denomina estructura de dominación asociada a la relación $\succ$, siendo $\mathcal{P}(R^k)$ el conjunto de las partes de R^k.

En lo que sigue, identificaremos la estructura D con su imagen.

Nota 3.5.4.

Podemos redefinir los elementos eficientes de $(Z, \succ)$, mediante la estructura de dominación D.

Por definición, es:

$$E(Z,\succ) = \{ \bar{x} \in Z \ / \ \nexists x \in Z \ \text{ tal que } \ x \succ \bar{x} \}$$

Pero,

$$x \succ \bar{x} \iff \exists\, d \in D(x) - \{0\} \ \text{ tal que } \ \bar{x} = x + d \iff$$

$$\iff \bar{x} \in x + (D(x) - \{0\})$$

Luego,

$$E(Z,\succ) = \{ \bar{x} \in Z \ / \ \nexists x \in Z \ \text{ tal que } \ \bar{x} \in x + (D(x)-\{0\}) \}$$

Y, cuando consideremos la estructura D asociada a $\succ$, escribiremos E(Z, D). Por tanto,

$$E(Z, D) = \{ \bar{x} \in Z \ / \ \nexists x \in Z \ \text{ tal que } \ \bar{x} \in x + (D(x)-\{0\}) \}$$

3.6. PROPIEDADES DE LA ESTRUCTURA DE DOMINACION D

Se demuestra, en este epígrafe, un conjunto de propiedades de la estructura D que nos servirá para probar la existencia de soluciones eficientes y precisar con gran detalle la ubicación de las mismas.

Teorema 3.6.1.

La estructura D viene dada por la expresión:

$$D = \{(x_1,..,x_k) \in R^k \ / \ x_1+..+x_k > 0\} \cup \{(0,..,0)\}$$

Y, por tanto, D es una aplicación constante.

Demostración

$$\forall (u_1,..,u_k) \in R^k \quad \text{es:}$$

$$D(u_1,..,u_k) = \{(x_1,..,x_k) \in R^k / \ (u_1,..,u_k) \succ (u_1,..,u_k)+(x_1,..,x_k)\} \cup \{0\} =$$

$$= \{(x_1,..,x_k) \in R^k \ / \ u_1+..+u_k < u_1+..+u_k+x_1+..+x_k\} \cup \{(0,..,0)\} =$$

$$= \{(x_1,..,x_k) \in R^k \ / \ x_1+..+x_k > 0\} \cup \{(0,..,0)\}$$

Luego, la imagen de cualquier elemento de R^k, mediante D, es siempre el mismo conjunto. Por tanto, como hemos convenido en representar Im(D) por D, resulta:

$$D = \{(x_1,..,x_k) \in R^k \ / \ x_1+...+x_k > 0\} \cup \{(0,..,0)\}$$

Por otra parte, según hemos visto antes, la imagen de cualquier elemento de R^k es siempre el mismo conjunto de R^k. Luego, D es una aplicación constante.

Nota 3.6.1.

En virtud del teorema 3.6.1. y de la forma en que se ha redefinido $E(Z, \succ)$ mediante D (Nota 3.5.1.), podemos expresar de modo más simplificado el conjunto $E(Z, D)$, quedando así:

$$E(Z, D) = \{ \bar{x} \in R^k \ / \ \nexists x \in Z \text{ tal que } \bar{x} \in x + (D - \{0\}) \}$$

Definición 3.6.1.

Se dice que una estructura de dominación $D \subset R^p$ es:

i) Asimétrica si

$$d \in D(y),\ d \neq 0 \Rightarrow -d \notin D(y+d) \qquad \forall y \in R^p$$

ii) Transitiva si

$$d \in D(y),\ d' \in D(y+d) \Rightarrow d+d' \in D(y) \qquad \forall y \in R^p$$

iii) Acíclica si $\forall n \in N^*$ nunca ocurre que:

$$y^1 \in y^2 + (D(y^2)-\{0\}),\ y^2 \in y^3 + (D(y^3)-\{0\}),\ldots,$$

$$y^n \in y^1 + (D(y^1)-\{0\})$$

con $y^1,\ldots\ldots,y^n \in R^p$

Teorema 3.6.2.

La estructura $D \subset R^k$ es asimétrica y transitiva.

Demostración

Según hemos visto en el teorema 3.6.1., D es constante.

Por otra parte, de la definición 3.6.1.i), D es asimétrica si $(d_1,..,d_k) \in D$, $(d_1,..,d_k) \neq (0,..,0) \Rightarrow$

$$\Rightarrow \quad -(d_1,..,d_k) \notin D$$

Pero,

$$\left.\begin{array}{l}(d_1,..,d_k) \in D \\ (d_1,..,d_k) \neq (0,..,0)\end{array}\right\} \overset{\text{Teorema 3.6.1.}}{\Longleftrightarrow} d_1+...+d_k > 0 \quad \Rightarrow$$

$$\Rightarrow \quad -(d_1+...+d_k) < 0 \quad \Rightarrow \quad -(d_1,...,d_k) \notin D$$

Luego, D es asimétrica.

Por otro lado, definición 3.6.1.ii), D es transitiva si:

$$(d_1,..,d_k) \in D,\ (d_1',..,d_k') \in D \Rightarrow (d_1,..,d_k)+(d_1',..,d_k') \in D$$

Ahora bien,

$$\left.\begin{array}{l}(d_1,..,d_k) \in D \Longleftrightarrow d_1+...+d_k > 0 \ \text{ó} \ (d_1,..,d_k)=(0,..,0) \\ \qquad\qquad \text{Teorema 3.6.1.} \\ (d_1',..,d_k') \in D \Longleftrightarrow d_1'+...+d_k' > 0 \ \text{ó} \ (d_1',..,d_k')=(0,..,0)\end{array}\right\} \Rightarrow$$

$$\Rightarrow \left\{\begin{array}{c}(d_1+..+d_k)+(d_1'+..+d_k') > 0 \\ \text{ó} \\ (d_1,..,d_k)+(d_1',..,d_k') = (0,..,0)\end{array}\right\} \Rightarrow$$

$\Rightarrow (d_1,\dots,d_k) + (d'_1,\dots,d'_k) \in D$

Luego, D es transitiva.

Queda demostrado el teorema 3.6.2.

Enunciamos a continuación, el siguiente resultado tomado del libro : "Theory of Multiobjective Optimization" de Sawaragi,Y., Nakayama,H. y Tanino, T.S. [42]

Teorema 3.6.3.

Si una estructura de dominación $D \subset R^p$ es asimétrica y transitiva, entonces es acíclica. Por otra parte, si una estructura de dominación $D \subset R^p$ es acíclica, entonces es asimétrica.

Como consecuencia de los dos teoremas anteriores, establecemos el siguiente:

Teorema 3.6.4.

La estructura de dominación $D \subset R^k$ es acíclica.

Demostración

Por el teorema 3.6.2., D es asimétrica y transitiva. Y por el teorema 3.6.3., se sigue que D es acíclica.

Quedando, así, demostrado el teorema.

En el teorema que daremos seguidamente, se prueba en primer lugar, que D es un cono, y por ser D constante (según se ha visto en el teorema 3.6.1.) las propiedades de transitividad y asimetría de la estructura $\mathcal{D}$ son equivalentes a otras dos propiedades de los conos.

Para la demostración del mencionado teorema, necesitamos previamente algunas definiciones.

Definición 3.6.2.

Un subconjunto K de R^p se dice que es:

i) Cono si $\forall x \in K,\ \alpha > 0,\ \alpha \in R \Rightarrow \alpha x \in K$

ii) Cono convexo si:

a) K es cono y b) K es convexo

iii) Cono apuntado si:

a) K es cono y b) $\forall x \in K,\ x \neq 0 \Rightarrow -x \notin K$

Se demuestra fácilmente, que:

$$K \subset R^p \text{ es un cono convexo} \Leftrightarrow \begin{cases} 1^\circ)\ K \text{ es un cono} \\ 2^\circ)\ \forall x,y \in K \Rightarrow x+y \in K \end{cases}$$

Teorema 3.6.5.

$D \subset R^k$ es un cono convexo apuntado.

Demostración

Probemos, en primer lugar, que D es un cono.

Por el teorema 3.6.1., sabemos que:

$D = \{(x_1,..,x_k) \in R^k \ / \ x_1+..+x_k > 0\} \cup \{(0,..,0)\}$

Sean $\forall (x_1,..,x_k) \in D$ y $\forall \alpha \in R, \ \alpha > 0$

Si $(x_1,..,x_k) \neq (0,..,0)$, se sigue por el teorema 3.6.1.

$$x_1+..+x_k > 0$$

Y por ser $\alpha > 0$

$$\alpha x_1+....+\alpha x_k > 0$$

Ello implica, por el teorema 3.6.1.:

$$(\alpha x_1,...,\alpha x_k) \in D \iff \alpha(x_1,..,x_k) \in D$$

Si $(x_1,...,x_k) = (0,..,0)$, concluímos que:

$$\alpha(x_1,..,x_k) = (0,..,0) \quad y \quad (0,..,0) \in D$$

Luego,

$$\forall (x_1,..,x_k) \in D, \ \forall \alpha \in R, \ \alpha > 0 \implies \alpha(x_1,..,x_k) \in D$$

Por tanto, D es un cono.

Demostremos, ahora, que D es un cono convexo.

Sabemos, por lo probado anteriormente, que D es un cono y, por el teorema 3.6.2., que D es transitiva.

Ahora bien,

$$\left.\begin{array}{l}\text{D cono}\\ \quad\text{y}\\ \text{D transitiva}\end{array}\right\} \Leftrightarrow \left\{\begin{array}{l}\text{D cono}\\ \quad\text{y}\\ \forall d,d' \in D \Rightarrow d+d' \in D\end{array}\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{\begin{array}{l}\text{D cono}\\ \quad\text{y}\\ \text{D convexo}\end{array}\right.$$

Luego, D es un cono convexo.

Probemos, finalmente, que D es un cono apuntado.

Hemos demostrado ya que D es un cono y, por el teorema 3.6.2., D es asimétrica.

Ahora bien,

$$\text{D cono y D asimétrica} \Leftrightarrow \left\{\begin{array}{l}\text{D cono}\\ \quad\text{y}\\ \forall d \in D,\ d \neq 0 \Rightarrow -d \notin D\end{array}\right\} \Leftrightarrow$$

$\Leftrightarrow$ D es un cono apuntado.

Termina, así, la demostración.

3.7. EXISTENCIA Y CARACTERIZACION DE LAS SOLUCIONES EFICIENTES

En este epígrafe, pasamos a estudiar el núcleo central de este capítulo, probando la existencia de soluciones eficientes del programa multiobjetivo formulado y caracterizando dichas soluciones eficientes. La caracteriza ción se realiza de dos formas distintas: la primera como conjunto transformado, mediante las funciones multiobjetivo definidas, de las soluciones óptimas del programa inicial y la segunda a través del cono polar positivo estricto de la estructura D y posteriormente, debido a la sencilla expresión de D^{so}, se precisa aún más la ubicación del conjunto $E(Z\ ,\ D)$.

Definición 3.7.1.

Un elemento $\bar{x} \in Y$ se dice que es una solución eficiente del problema de programación multiobjetivo (II) con respecto a la estructura de dominación D si

$$T(\bar{x}) = (P_1(\bar{x}),\dots,P_k(\bar{x})) \in E(Z\ ,\ D)$$

Teorema 3.7.1.

$$E(Z\ ,\ D) = T\ (S_o)$$

siendo S_o el conjunto de soluciones óptimas del programa (I). Y, por tanto, S_o es el conjunto de soluciones eficientes del programa (II).

Demostración

Probemos, en primer lugar, que $T(S_o) \subset E(Z\ ,\ D)$.

$$(\bar{x}_1,\dots,\bar{x}_k) \in T(S_o) \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \quad \exists\, x^o \in S_o \quad \text{tal que} \quad T(x^o) = (\bar{x}_1,\dots,\bar{x}_k)$$

$\Rightarrow \exists x^o \in Y$, con $T(x^o) = (\bar{x}_1, \ldots, \bar{x}_k)$, tal que

$P(x^o) \leqslant P(x) \qquad \forall x \in Y \qquad \Leftrightarrow$

$\Leftrightarrow \exists x^o \in Y$, con $T(x^o) = (\bar{x}_1, \ldots, \bar{x}_k)$, tal que

$\nexists x \in Y$ verificando que $P(x) < P(x^o) \quad \Leftrightarrow$

$\Leftrightarrow \exists x^o \in Y$, con $T(x^o) = (\bar{x}_1, \ldots, \bar{x}_k)$, tal que

Def. 3.2.1. $\quad \nexists x \in Y$ verificando que $T(x) \succ T(x^o)$

$\Leftrightarrow \exists x^o \in Y$, con $T(x^o) = (\bar{x}_1, \ldots, \bar{x}_k)$, tal que $\nexists x \in Y$

verificando que $(P_1(x), \ldots, P_k(x)) \succ (P_1(x^o), \ldots, P_k(x^o))$

Def.3.7.1.

$\Leftrightarrow \exists x^o \in Y$, con $T(x^o) = (\bar{x}_1, \ldots, \bar{x}_k)$, tal que x^o es una solución eficiente al problema de programación multiobjetivo (II) $\Leftrightarrow$

$\Leftrightarrow \exists x^o \in Y$ tal que $T(x^o) = (\bar{x}_1, \ldots, \bar{x}_k) \in E(Z, D)$

Def.3.7.1.

Luego,

$$\forall (\bar{x}_1, \ldots, \bar{x}_k) \in T(S_o) \Rightarrow (\bar{x}_1, \ldots, \bar{x}_k) \in E(Z, D)$$

Por tanto,

$$T(S_o) \subset E(Z, D) \qquad (1)$$

Recíprocamente, demostremos que $E(Z, D) \subset T(S_o)$.

En efecto:

$\forall (\bar{x}_1, \ldots, \bar{x}_k) \in E(Z, D) \Rightarrow \exists x^o \in Y$ tal que

$T(x^o) = (\bar{x}_1, \ldots, \bar{x}_k) \in E(Z, D) \quad \Leftrightarrow$

Def. 3.7.1.

$\overset{\downarrow}{<=>}$ $\exists\, x^o \in Y$, con $T(x^o) = (\bar{x}_1,\ldots,\bar{x}_k)$, tal que x^o es una solución eficiente del programa (II) $<=>$

$<=>$ $\exists\, x^o \in Y$, con $T(x^o) = (\bar{x}_1,\ldots,\bar{x}_k)$, tal que $T(x^o) \succsim T(x) \quad \forall x \in Y$ $<=>$

$<=>$ $\exists\, x^o \in Y$, con $T(x^o) = (\bar{x}_1,\ldots,\bar{x}_k)$, tal que $\nexists\, x \in Y$ verificando que $T(x) \succ T(x^o)$ $<=>$

$<=>$ $\exists\, x^o \in Y$, con $T(x^o) = (\bar{x}_1,\ldots,\bar{x}_k)$; tal que $\nexists\, x \in Y$ verificando que $P(x) < P(x^o)$ $<=>$

$<=>$ $\exists\, x^o \in Y$, con $T(x^o) = (\bar{x}_1,\ldots,\bar{x}_k)$, tal que $\forall x \in Y$ es $P(x^o) \leq P(x)$ $=>$

$\exists\, x^o \in S_o$, con $T(x^o) = (\bar{x}_1,\ldots,\bar{x}_k)$

Por tanto,

$$(\bar{x}_1,\ldots,\bar{x}_k) \in T(S_o)$$

Luego,

$$\forall (\bar{x}_1,\ldots,\bar{x}_k) \in E(Z,\ D) => (\bar{x}_1,\ldots,\bar{x}_k) \in T(S_o)$$

Concluímos, por ello, que:

$$E(Z,\ D) \subset T(S_o) \qquad (2)$$

De (1) y (2), se sigue que:

$$E(Z,\ D) = T(S_o)$$

Y, de acuerdo con la definición 3.7.1., concluímos que S_o es el conjunto de soluciones eficientes del programa (II).

Quedando, así, probado el teorema.

Teorema 3.7.2.

$S_o \neq \emptyset$, siendo S_o el conjunto de soluciones óptimas del programa (I) (que, como se ha visto en el teorema anterior, coincide con las soluciones eficientes del programa (II)).

Demostración

El programa inicial es:

$$(I) \begin{cases} \text{Min } P(x) \\ \text{bajo las restricciones:} \\ x \in Y \subset R^n \end{cases}$$

siendo P(x) función polinómica e Y compacto en R^n con la topología usual.

Tenemos, por tanto, que:

P(x) es una función contínua (por ser P(x) polinómica) en el conjunto compacto Y, luego, de acuerdo con el teorema de Weierstrass, P(x) alcanza su máximo y su mínimo absolutos en Y. Por consiguiente,

S_o = { $x \in Y$ / x es solución óptima del programa (I)} =

= { $x \in Y$ / x es mínimo absoluto de $P(x)$ en Y } =

= { $x \in Y$ / x es solución eficiente del programa (II)} =

Con lo que se termina la demostración.

Teorema 3.7.3.

$$E(Z, D) \neq \emptyset$$

siendo $E(Z, D)$ el conjunto de elementos eficientes de Z respecto de la estructura de dominación D.

Demostración

Por el teorema 3.7.2., $S_o \neq \emptyset \iff T(S_o) \neq \emptyset$

Ahora bien, por el teorema 3.7.1.,

$$T(S_o) = E(Z, D)$$

Luego,

$$E(Z, D) \neq \emptyset$$

Queda, así, probado el teorema.

Nota 3.7.1.

La demostración de que $E(Z, D) \neq \emptyset$, podíamos también haberla realizado utilizando un teorema dado por

Hazen y Morín [31], que se basa en la aciclicidad de una estructura de dominación bajo ciertas hipótesis. Hipótesis que verifica la estructura D obtenida por nosotros.

Para dar otra caracterización más precisa del conjunto E(Z, D), vamos a considerar los conos polar positivo y polar positivo estricto de la estructura de dominación D.

Definición 3.7.2.

Dado $X \subset R^p$, $X \neq \emptyset$, se denomina conjunto polar positivo de X y lo denotaremos X^o, a:

$$X^o = \{x^* \in R^p \ / <x,x^*> \geq 0, \ \forall x \in X\}$$

Análogamente, se denomina conjunto polar positivo estricto de X y se denotará X^{so}, a:

$$X^{so} = \{x^* \in R^p \ / <x,x^*> > 0, \quad \forall x \in X, \ x \neq 0\}$$

siendo <. , .> el producto escalar habitual de R^p respecto de la base canónica, es decir,

$$<x , y> = \sum_{i=1}^{p} x_i y_i$$

Se demuestra de forma inmediata que X^o es un cono convexo cerrado y que X^{so} es un cono convexo.

Pasamos a estudiar las expresiones de los conos D^{o} y D^{so}.

Teorema 3.7.4.

El cono polar positivo D^{o} está dado por la semirrecta:

$$D^{o} = \{ (x_1,..,x_k) \in R^k / x_1=x_2=..=x_k \ \text{ y } \ x_i \geq 0 \quad i= 1,..,k \}$$

Demostración

Sea:

$$r^{o}=\{ (x_1,...,x_k) \in R^k / x_1 =...= x_k \ \text{ y } \ x_i \geq 0 \quad i = 1,..,k \}$$

Demostremos que $r^{o} = D^{o}$

1°) $r^{o} \subset D^{o}$

En efecto, $\forall x=(x_1,..,x_k) \in r^{o}$ es $x_1=...= x_k \geq 0$

Luego,

$\forall x^*=(x_1^*,..,x_k^*) \in D$ es, por la forma analítica de D,

$$<x , x^* > = \sum_{i=1}^{k} x_i x_i^* = x_1 (x_1^*+...+x_k^*) \geq 0$$

Luego, $x \in D^{o}$. Por tanto, $\forall x \in r^{o} \Rightarrow x \in D^{o}$. Es decir,

$$r^{o} \subset D^{o} \qquad (1)$$

2°) Recíprocamente, probemos que $D^{o} \subset \Gamma^{o}$

Para ello, vamos a demostrar tres proposiciones:

a) $D^{o} \subset D$

b) $\forall x = (x_1,..,x_k) \in D^{o} \Rightarrow x_1 \geq 0,...,x_k \geq 0$

c) $\forall x = (x_1,...,x_k) \in D^{o} \Rightarrow x_1 = x_2 =...= x_k$

Demostración de a)

La haremos por reducción al absurdo.

Si $D^{o} \not\subset D \Rightarrow \exists x = (x_1,..,x_k) \in D^{o}$ y $x \notin D$

Ello implica que:

$\exists x \in D^{o}$ con $x_1+..+x_k \leq 0$ y $(x_1,..,x_k) \neq (0,..,0)$

De donde se sigue que:

$\exists x \in D^{o}$ con $x_i < 0$ para algún $i=1,..,k$

Tomando, entonces, $x^{*} \in D$ tal que $x_i^{*} > 0$ y

$x_j^{*} = 0 \quad j \neq i$, tendríamos:

$$< x \, , \, x^{*} > = x_i \, x_i^{*} < 0 \overset{\text{Def. de } D^{o}}{\Rightarrow} x \notin D^{o}$$

Contradicción, que nace de suponer que $D^{o} \not\subset D$.

Luego, $D^{o} \subset D$.

Demostración de b)

Si $\exists x = (x_1,..,x_k) \in D^{o}$ con $x_i < 0$ para algún $i = 1,..,k$. Tomando $x^{*} \in D$ con $x_i^{*} > 0$ y $x_j = 0 \quad \forall j \neq i$, se tendría:

$$\langle x\, , x^* \rangle = x_i\, x_i^* < 0 \quad \overset{\text{Def. de } D^\circ}{\Longrightarrow} \quad x \notin D^\circ$$

Contradicción, que se origina al suponer que $\exists\, x = (x_1,..,x_k) \in D^\circ$ con $x_i < 0$ para algún $i = 1,..,k$.

Luego,

$$\forall x = (x_1,..,x_k) \in D^\circ \Rightarrow x_1 \geq 0,\ x_2 \geq 0,..,x_k \geq 0$$

Demostración de c)

Si $\exists x = (x_1,..,x_k) \in D^\circ$ con no todas sus componentes iguales, entonces:

$\exists\, i \in \{1,..,k\}$ con $x_i = \max\{ x_1,..,x_k\} > 0$ y

$\exists j \neq i$ con $x_j < x_i$

Tomando $x^* = (0,..,\underset{i}{-1},..,\underset{j}{1},...,0) \in D$ (pues $-1 + 1 = 0$), resultaría:

$$\langle x\, , x^* \rangle = - x_i + x_j < 0 \quad \overset{\text{Def. de } D^\circ}{\Longrightarrow} \quad x \notin D^\circ$$

Contradicción, que tiene su origen en suponer que $\exists x \in D^\circ$ con no todas sus componentes iguales. Luego,

$$\forall x = (x_1,..,x_k) \in D^\circ \Rightarrow x_1 = x_2 = ... = x_k$$

De las proposiciones b) y c) anteriores, se concluye que:

$$D^\circ \subset r^\circ \qquad (2)$$

De (1) y (2), deducimos finalmente que:

$D^{o}=\{(x_1,..,x_k)\in R^k/\ x_1=...=x_k\ \ y\ \ x_i\geqslant 0\ \ i=1,..,k\}$

Se termina, así, la demostración del teorema 3.7.4.

Teorema 3.7.5.

El cono polar positivo estricto de D es:

$D^{so}=D^{o}-\{0\}=\{(x_1,..,x_k)\in R^k\ /\ x_1=...=x_k\ y\ x_i>0\ \ i=1,..,k\}$

Demostración

Se realizará, demostrando tres propiedades:

1ª) $D^{so}\subset D$

Es una consecuencia de las definiciones de los conos polar positivo y polar positivo estricto.

2ª) $(0,...,0)\notin D^{so}$

En efecto, $\forall x^*\in D-\{(0,..,0)\}$ es:

$$<x^*,\ \theta>=0$$

De donde se sigue, por definición de cono polar positivo estricto, que:

$$(0,...,0)\notin D^{so}$$

3ª) $D^{o}-\{(0,..,0)\}\subset D^{so}$

En efecto,

$$\forall(x_1,..,x_k)\in D^{o}-\{(0,..,0)\}\xrightarrow{\text{Teorema 3.7.4.}}x_1=...=x_k>0$$

Entonces,

$\forall x^* = (x_1^*,\dots,x_k^*) \in D - \{(0,\dots,0)\}$, como

$x_1^* + \dots + x_k^* > 0$ (teorema 3.6.1.), se verifica

que:

$$< x \, , \, x^* > = \sum_{i=1}^{k} x_i x_i^* = x_1 (\sum_{i=1}^{k} x_i^*) > 0$$

ya que $x_1 > 0$ y $\sum_{i=1}^{k} x_i^* > 0$

Luego, por definición de cono polar positivo estricto, $x = (x_1,\dots,x_k) \in D^{so}$.

Es decir,

$$D^o - \{(0,\dots,0)\} \subset D^{so}$$

De las propiedades anteriormente demostradas, concluímos, finalmente, que:

$$D^{so} = D^o - \{(0,\dots,0)\}$$

Así termina la demostración de este teorema.

Para la caracterización de E(Z, D), que vamos a establecer a través de D^{so}, necesitamos introducir un nuevo conjunto.

Dado $\alpha \in R^k$, sea

$$S(\alpha, Z) = \{\bar{x} = (\bar{x}_1,\dots,\bar{x}_k) \in Z \,/ < \alpha, \bar{x} > = \inf_{x \in Z}.\{< \alpha, x >\}\} =$$

$$=\{\bar{x}=(\bar{x}_1,\ldots,\bar{x}_k)\in Z/\sum_{i=1}^{k}\alpha_i\bar{x}_i=\inf_{x=(x_1,\ldots,x_k)\in Z}\{\sum_{i=1}^{k}\alpha_i x_i\}\}$$

Es claro que $S(\lambda\alpha\ ,\ Z)=S(\alpha\ ,\ Z)\quad \forall\lambda>0\ ,\ \lambda\in R$

Sea $F(Z\ ,\ D)=\bigcup_{\alpha\in D^{so}} S(\alpha,\ Z)$

De acuerdo con el teorema 3.7.5., los elementos de D^{so} son de la forma $\lambda(1,1,\ldots,1)\quad\forall\lambda>0,\quad\lambda\in R.$

Por tanto,

$$\forall\alpha\in D^{so}\quad\exists\lambda>0,\quad\lambda\in R\ \text{ tal que }\ \alpha=\lambda\ (1,\ldots,1)$$

Luego, según la observación anterior, $\forall\alpha\in D^{so}$, se verifica que

$$S(\alpha,\ Z)=S\ (\lambda(1,\ldots,1),\ Z)=S\ ((1,\ldots,1),\ Z)$$

Por tanto,

$$F\ (Z\ ,\ D)=\bigcup_{\alpha\in D^{so}} S(\alpha\ ,\ Z)=$$

$$=\bigcup_{\substack{\lambda(1,\ldots,1)\in D^{so}\\ \lambda>0}} S\ (\lambda\ (1,\ldots,1),\ Z)=S\ ((1,\ldots,1),\ Z)$$

Luego,

$$F(Z,D)=\{\bar{x}=(\bar{x}_1,\ldots,\bar{x}_k)\in Z/<(1,\ldots,1),\bar{x}>=\inf_{x\in Z}\{<(1,\ldots,1),x>\}\}=$$

$$=\{\bar{x}=(\bar{x}_1,\ldots,\bar{x}_k)\in Z\ /\sum_{i=1}^{k}\bar{x}_i=\inf_{x\in Z}\{\sum_{i=1}^{k}x_i\}\}$$

Teorema 3.7.6.

$$F\ (Z,\ D)\ \neq\ \emptyset$$

Demostración

Por definición de T, se verifica que $T(Y) = Z$. Ahora bien, Y es compacto en R^n (por hipótesis) y $T = (P_1,\ldots,P_k)$ es una aplicación contínua; de ello se sigue (por un teorema de continuidad en conjuntos compactos) que Z es un conjunto compacto en R^k.

Por otra parte, $f(.) = <(1,\ldots,1),\ .>$ es una aplicación contínua de Z en R, pues:

$$\forall x^o = (x_1^o,\ldots,x_k^o) \in Z: \lim_{\substack{x \to x^o \\ x \in Z}} f(x) = \lim_{\substack{x \to x^o \\ x \in Z}} <(1,\ldots,1),\ x> =$$

$$= \lim_{\substack{x \to x^o \\ x \in Z}} \left(\sum_{i=1}^{k} x_i\right) = \sum_{i=1}^{k} x_i^o = <(1,\ldots,1),\ x^o> = f(x^o) \iff$$

$$\iff \quad f = <(1,\ldots,1),\ .> \text{ es contínua en } Z.$$

De la continuidad de $f(.)$ en el compacto Z se sigue, teorema de Weierstrass, que f alcanza su máximo y su mínimo absolutos en Z. Por tanto,

$$\exists\, \bar{x} = (\bar{x}_1, \ldots, \bar{x}_k) \in Z \quad \text{tal que:}$$

$$\sum_{i=1}^{k} \bar{x}_i = \inf_{x \in Z} . \{ \sum_{i=1}^{k} x_i \} = \min_{x \in Z} \{ \sum_{i=1}^{k} x_i \}$$

Ello es equivalente a:

$$\exists\, \bar{x} \in Z \quad \text{tal que} \quad \bar{x} \in F(Z, D)$$

Luego,

$$F\,(Z,\, D) \neq \emptyset$$

Con lo que finaliza la demostración.

Teorema 3.7.7.

$$E(Z,\, D) = F(Z,\, D)$$

Demostración

Por definición de conjunto eficiente:

$$E(Z,\, D) = \{ \bar{x} \in Z \;/\; \nexists\, x \in Z \quad \text{tal que} \quad x \succ \bar{x} \}$$

Pero, teniendo en cuenta la relación $\succsim$

$$\{ \bar{x} \in Z \;/\; \nexists\, x \in Z \;\; \text{tal que} \;\; x \succ \bar{x} \} = \{ \bar{x} \in Z \;/\; \forall x \in Z \;\; \text{es} \;\; \bar{x} \succsim x \} =$$

$$= \{ \bar{x} \in Z \;/\; \forall x \in Z \;\; \text{es} \; \sum_{i=1}^{k} \bar{x}_i \leq \sum_{i=1}^{k} x_i \} =$$

$$= \{ \bar{x} \in Z \;/ \sum_{i=1}^{k} \bar{x}_i = \inf_{x \in Y} . \{ \sum_{i=1}^{k} x_i \} \} \underset{\text{Def. de } F}{=} F(Z,\, D)$$

Luego,

$$E(Z,\ D) = F(Z,\ D)$$

Quedando, así, demostrado el teorema.

Nota 3.7.2.

También se podía haber demostrado el teorema 3.7.7. antes de probar que $F(Z,\ D) \neq \emptyset$, y entonces, teniendo en cuenta que $F(Z,\ D) = E(Z,\ D)$ y el teorema 3.7.3., se habría concluído que $F(Z,\ D) \neq \emptyset$.

Aún podemos afinar más sobre los elementos del conjunto $E(Z,\ D)$ y sobre la ubicación del mismo, precisión que nos viene dada por el siguiente:

Teorema 3.7.8.

El conjunto $E(Z,\ D)$ verifica las siguientes propiedades:

1ª) $\forall x, x' \in E(Z,\ D) \Rightarrow x \sim x'$

2ª) $E(Z,\ D) \subset D - \{(0,\ldots,0)\}$ ó $E(Z,\ D) \subset (D - \{(0,\ldots,0$

3ª) $E(Z,\ D) \subset D - \{(0,\ldots,0)\}$ si y sólo si el valor mínimo del programa (I) es mayor que cero.

4ª) $E(Z, D) \subset (D - \{(0,\dots,0)\})^c$ si y sólo si el valor mínimo del programa (I) es menor o igual que cero.

5ª) $E(Z, D) \subset D - \{(0,\dots,0)\}$ si y sólo si $Z \subset D - \{(0,\dots,0)\}$

6ª) $Z \subset (D - \{(0,\dots,0)\})^c$ implica que $E(Z, D) \subset (D - \{(0,\dots,0)\})^c$

7ª) Si $Z \subset D - \{(0,\dots,0)\}$, entonces

$$d(E, Fr(D - \{(0,\dots,0)\})) = \min_{x \in Z} d(x, Fr(D - \{(0,..,0)\})) =$$

$$= d(Z, Fr(D - \{(0,\dots,0)\}))$$

8ª) Si $Z \not\subset D - \{(0,\dots,0)\}$, entonces

$$d(E(Z, D), Fr(D - \{(0,\dots,0)\})) =$$

$$= \max_{x \in Z \cap (D - \{(0,\dots,0)\})^c} \{d(x, Fr(D - \{(0,\dots,0)\}))\}$$

Demostración

1ª) Sean $\forall x, x' \in E(Z, D)$

Por definición de $E(Z, D)$:

$x \in E(Z, D) \iff x \in Z$ y $\forall x^* \in Z$ es $x \succsim x^*$

En particular, para $x^* = x' \in Z$ se verifica que:

$$x \succsim x' \qquad (1)$$

Análogamente,

$$x' \in E(Z, D) \iff x' \in Z \quad y \quad \forall x^* \in Z \quad es \quad x' \succsim x^*$$

En particular, para $x^* = x \in Z$ se verifica que:

$$x' \succsim x \qquad (2)$$

De (1) y (2) se sigue que:

$$\forall x, x' \in E(Z, D) \implies x \sim x'$$

2ª) De acuerdo con el teorema 3.7.7.

$$E(Z, D) = \{x \in Z \,/\, \sum_{i=1}^{k} x_i = \min_{x^* \in Z} \{ \sum_{i=1}^{k} x_i^* \} \} =$$

$$= \{x \in Z \,/\, \sum_{i=1}^{k} x_i = c, \text{ siendo } c = \min_{x^* \in Z} \{ \sum_{i=1}^{k} x_i^* \} \} \quad (3)$$

Por otra parte, del teorema 3.6.1. se sigue que:

$$Fr(D - \{(0,\ldots,0)\}) = \{x \in R^k \,/\, \sum_{i=1}^{k} x_i = 0 \} \qquad (4)$$

Comparando (3) y (4), se observa que:

$$E(Z, D) \subset \text{Hiperplano paralelo al } \{ x \in R^k \,/\, \sum_{i=1}^{k} x_i = 0\} =$$

$$= Fr(D - \{(0,\ldots,0)\})$$

Por tanto,

$$\text{si} \quad \min_{x \in Z} \{\sum_{i=1}^{k} x_i\} = c > 0 \implies E(Z, D) \subset \{x \in R^k / \sum_{i=1}^{k} x_i = c,$$

De donde concluímos que:

$$E(Z,\ D) \subset D - \{(0,\dots,0)\}$$

Análogamente,

si $\min_{x \in Z} \{\sum_{i=1}^{k} x_i\} = c \leqslant 0 \quad \Rightarrow$

$\Rightarrow \quad E(Z,\ D) \subset \{ x \subset R^k \ / \sum_{i=1}^{k} x_i \leqslant c, \quad c \leqslant 0 \}$

De donde se concluye que:

$$E(Z,\ D) \subset (D - \{(0,\dots,0)\})^c$$

3ª) $E(Z,\ D) \subset D - \{(0,\dots,0)\} \underset{\text{Teórema 3.6.1.}}{\overset{}{\Leftrightarrow}} \forall x \in E(Z,\ D)$ es $\sum_{i=1}^{k} x_i > 0$

Teorema 3.7.1.

$\Leftrightarrow \quad \forall u \in S_o$ (soluciones óptimas del programa (I))

es $\sum_{i=1}^{k} P_i(u) = P(u) = \sum_{i=1}^{k} x_i > 0 \quad \Leftrightarrow$

$\Leftrightarrow$ El valor mínimo del programa (I) es mayor que cero.

4ª) $E(Z,\ D) \subset (D - \{(0,\dots,0)\})^c \quad \overset{\text{Teorema 3.6.1.}}{\Leftrightarrow}$

$\Leftrightarrow \quad \forall x \in E(Z,\ D)$ es $\sum x_i \leqslant 0 \quad \overset{\text{Teorema 3.7.1.}}{\Leftrightarrow}$

$\Leftrightarrow \quad \forall u \in S_o$ es $\sum_{i=1}^{k} P_i(u) = P(u) = \sum_{i=1}^{k} x_i \leqslant 0 \quad \Leftrightarrow$

$\Leftrightarrow$ El valor mínimo del programa (I) es menor o igual que cero.

5ª)

$\Rightarrow$ $E(Z,\ D) \subset D - \{(0,\dots,0)\} \Leftrightarrow \forall x \in E(Z,\ D)$ es $\sum_{i=1}^{k} x_i > 0$

Pero, por el teorema 3.7.7.,

$$E(Z,\ D) = \{\ x \in Z\ /\sum_{i=1}^{k} x_i = \min_{x^* \in Z} \{\sum_{i=1}^{k} x_i^*\} > 0\ \} \Rightarrow$$

$$\Rightarrow \quad \forall x^* \in Z \quad \text{es} \quad \sum_{i=1}^{k} x_i^* > 0$$

Luego,

$$Z \subset D - \{(0,..,0)\}$$

$\Leftarrow$ Recíprocamente,

Si $\quad Z \subset D - \{(0,...,0)\}$, como $\quad E(Z,\ D) \subset Z$

concluímos que,

$$E(Z,\ D) \subset D - \{(0,...,0)\}$$

6ª) Esta propiedad es inmediata, ya que:

Si $\quad Z \subset (D - \{(0,...,0)\}\ )^c$, como $\quad E(Z,\ D) \subset Z$

se concluye que:

$$E(Z,\ D) \subset (D - \{(0,...,0)\}\ )^c$$

7ª) $Z \subset D - \{(0,...,0)\} \underset{\text{Propiedad } 5^a}{\Leftrightarrow} E(Z,\ D) \subset D - \{(0,...,0)\} \Rightarrow$

$$\Rightarrow \quad \forall x^* \in E(Z,\ D) \quad \text{es} \quad \sum_{i=1}^{k} x_i^* > 0 \qquad (1)$$

Por otra parte, teorema 3.7.7., se verifica que:

$$E(Z,\ D) = \{x^* \in Z\ /\sum_{i=1}^{k} x_i^* = \min_{x \in Z} \{\sum_{i=1}^{k} x_i\}\ \} \qquad (2)$$

Por ser $\sum_{i=1}^{k} x_i^* > 0 \qquad \forall x^* \in E(Z,\ D)$, se sigue que:

$$\frac{\left|\sum_{i=1}^{k} x_i^*\right|}{\sqrt{n}} = \frac{\sum_{i=1}^{k} x_i^*}{\sqrt{n}} = d(x^*, \{x \in R^k / \sum_{i=1}^{k} x_i = 0\}) =$$

$$= d(x^*, Fr(D - \{(0,\ldots,0)\})) \underset{\substack{\uparrow \\ \text{Propiedad } 1^a}}{=} d(E(Z, D), Fr(D - \{(0,\ldots,0)\})) \quad (3)$$

Asimismo,

$$\min_{x \in Z} \left\{ \frac{\sum_{i=1}^{k} x_i}{\sqrt{n}} \right\} \overset{\substack{Z \subset D - \{(0,\ldots,0)\} \\ \downarrow}}{=} \min_{x \in Z} \left\{ \frac{\left|\sum_{i=1}^{k} x_i\right|}{\sqrt{n}} \right\} =$$

$$= \min_{x \in Z} \{d(x, \{x' \in R^k / \sum_{i=1}^{k} x'_i = 0\})\} =$$

$$= \min_{x \in Z} \{d(x, Fr(D - \{(0,\ldots,0)\}))\} =$$

$$= d(Z, Fr(D - \{(0,\ldots,0)\})) \quad (4)$$

De (2), (3) y (4), se sigue que:

$$d(E(Z, D), Fr(D - \{(0,\ldots,0)\})) = d(Z, Fr(D - \{(0,\ldots,0)\}))$$

Luego, si $Z \subset D - \{(0,\ldots,0)\}$ se concluye que:

$$d(E(Z, D), Fr(D - \{(0,\ldots,0)\})) = \min_{x \in Z} d(x, Fr(D - \{(0,\ldots,0)\})) =$$

$$= d(Z, Fr(D - \{(0,\ldots,0)\}))$$

8ª) Si $Z \not\subset D - \{(0,\ldots,0)\}$ =>

=> $\exists x^* \in Z$ tal que $x^* \in (D - \{(0,\ldots,0)\})^c$

$$\exists x^* \in Z \qquad \text{tal que} \quad \sum_{i=1}^{k} x_i^* \leq 0$$

Como, por el teorema 3.7.7., es:

$$E(Z, D) = \{ \bar{x} \in Z \,/ \sum_{i=1}^{k} \bar{x}_i = \min_{x \in Z} \{ \sum_{i=1}^{k} x_i \} \}$$

resulta claro que tal mínimo se ha de alcanzar en puntos de Z tales que la suma de sus coordenadas sea menor o igual que cero. Luego,

$$E(Z, D) \subset (D - \{ (0,\ldots,0) \})^c$$

Y, por consiguiente, para la obtención de $\min_{x \in Z} \{ \sum_{i=1}^{k} x_i \}$ basta elegir aquellos elementos de Z contenidos en $(D - \{(0,\ldots,0)\})^c$. Es decir, si $Z \not\subset D - \{(0,\ldots,0)\}$, entonces:

$$\min_{x \in Z} \{ \sum_{i=1}^{k} x_i \} = \min_{x \in Z \cap (D - \{0\})^c} \{ \sum_{i=1}^{k} x_i \}$$

Luego,

$$E(Z, D) = \{ \bar{x} \in Z \,/ \sum_{i=1}^{k} \bar{x}_i = \min_{x \in Z \cap (D - \{0\})^c} \{ \sum_{i=1}^{k} x_i \} \}$$

Pero,

$$\{ \bar{x} \in Z \,/ \sum_{i=1}^{k} \bar{x}_i = \min_{x \in Z \cap (D - \{0\})^c} \{ \sum_{i=1}^{k} x_i \} \} =$$

$$= \{ \bar{x} \in Z \,/ -\sum_{i=1}^{k} \bar{x}_i = -\min_{x \in Z \cap (D - \{0\})^c} \{ \sum_{i=1}^{k} x_i \} \} =$$

$$=\{\ \bar{x}\in Z\ /\ -\sum_{i=1}^{k}\bar{x}_i = \max_{x\in Z\cap(D-\{0\})^c}\{-\sum_{i=1}^{k}x_i\}\}$$

Ahora bien, $\forall\bar{x}\in Z$:

$$\frac{-\sum_{i=1}^{k}\bar{x}_i}{\sqrt{n}} = \frac{\left|\sum_{i=1}^{k}x_i\right|}{\sqrt{n}} = d(\bar{x},\{\ x\in R^k/\sum_{i=1}^{k}x_i = 0\}\ ) =$$

$$= d(\bar{x},Fr(D-\{(0,\ldots,0)\})) \underset{\substack{\uparrow\\ \text{Propiedad } 1^a}}{=} d(E(Z,\ D),Fr(D-\{(0,\ldots,0)\}))$$

Análogamente,

$$\max_{x\in Z\cap(D-\{0\})^c}\left\{\frac{-\sum_{i=1}^{k}x_i}{\sqrt{n}}\right\} = \max_{x\in Z\cap(D-\{0\})^c}\left\{\frac{\left|\sum_{i=1}^{k}x_i\right|}{\sqrt{n}}\right\} =$$

$$= \max_{x\in Z\cap(D-\{0\})^c}\{\ d(x,\{x^*\in R^k/\sum_{i=1}^{k}x_i^* = 0\}\ )\} =$$

$$= \max_{x\in Z\cap(D-\{0\})^c}\{\ d(x,\ Fr(D-\{(0,\ldots,0)\}\ ))\}$$

Queda, así, demostrado completamente el teorema 3.7.8.

Nota 3.7.3.

Las propiedades séptima y octava del teorema 3.7.8., también se pueden enunciar como sigue:

7ª) Si $Z \subset D - \{(0,\ldots,0)\}$, entonces

$\bar{x} \in E(Z, D)$ si y sólo si

$$d(\bar{x}, Fr(D - \{(0,\ldots,0)\})) = \min_{x \in Z} \{d(x, Fr(D - \{0,\ldots,0)\})\}$$

8ª) Si $Z \not\subset D - \{(0,\ldots,0)\}$, entonces

$\bar{x} \in E(Z, D)$ si y sólo si

$$d(\bar{x}, Fr(D - \{(0,\ldots,0)\})) = \max_{x \in Z \cap (D - \{0\})^c} \{d(x, Fr(D - \{(0,\ldots,0)\})\}$$

El teorema que formulamos, finalmente, nos indica que cualquiera que sea la ubicación del óptimo en el programa (I), se verifica que los elementos eficientes del equivalente programa multiobjetivo (II), están en la frontera de Z.

Teorema 3.7.9.

$$E(Z, D) \subset Fr(Z)$$

Demostración

Pueden presentarse dos posibilidades:

a) $Int(Z) = \emptyset$

En este caso, evidentemente, $Z = Fr(Z)$ y, como $E(Z, D) \subset Z$ (por definición de $E(Z, D)$), se concluye que:

$$E(Z, D) \subset Fr(Z)$$

Y, para este caso, queda probado el teorema.

b) $Int(Z) \neq \emptyset$

Vamos a realizar la demostración por reducción al absurdo.

Supongamos que $E(Z, D) \not\subset Fr(Z)$

Ello equivale a que:

$$\exists x \in E(Z, D) \quad y \quad x \notin Fr(Z) \quad \Rightarrow$$

$$\Rightarrow \quad \exists x \in E(Z, D) \quad y \quad x \in Int(Z) \quad \Rightarrow$$

$$\Rightarrow \quad \exists x \in E(Z, D) \quad y \quad \exists \mathring{B}(x;r) \subset Int(Z) \quad \Rightarrow$$

$$\Rightarrow \quad \exists x \in E(Z, D) \quad y \quad \exists \bar{x} \in \mathring{B}(x;r) \subset Int(Z)$$

tales que $\sum_{i=1}^{k} \bar{x}_i < \sum_{i=1}^{k} x_i$

Como los elementos de $E(Z, D)$, son indiferentes entre sí (propiedad 1ª del teorema 3.7.8.), se sigue que:

$\exists\ \bar{x} \in Z$ tal que $\sum_{i=1}^{k} \bar{x}_i < \sum_{i=1}^{k} x_i^* \quad \forall x_i^* \in E(Z\ ,\ D)$ <=>

$\underset{\text{Def. de } \succ}{<=>}$ $\exists\ \bar{x} \in Z$ tal que $\bar{x} \succ x^* \quad \forall x^* \in E(Z\ ,\ D)$

Contradicción con la definición de $E(Z\ ,\ D)$, que nace de suponer que $E(Z\ ,\ D) \not\subset Fr(Z)$. Luego,

$$E(Z\ ,\ D) \subset Fr(Z)$$

Queda, así, demostrado el teorema.

A P E N D I C E

A.1. INTRODUCCION

De los resultados conocidos de Programación Convexa, se enuncian en este Apéndice aquellos que nos han servido de apoyo en algunos teoremas de los que aparecen en esta Memoria. Tales resultados pueden encontrarse en libros que traten de la Programación Convexa, como por ejemplo, Mangasarian, O.L. [37] y Minoux, M. [40]. Asimismo, se incluye una demostración original de un resultado topológico conocido que es necesario para establecer un teorema fundamental del Capítulo 1.

Teorema A.1. (Convexidad del conjunto de restricciones)

Dado el conjunto:

$$Y = \{x \in R^n \ / \ g_i(x) \leq 0 \qquad i \in I = \{1,\dots,m\}\},$$

si $Y \neq \emptyset$ y $g_i(x)$ $i \in I$ son convexas en R^n, entonces Y es convexo.

Teorema A.2. (Convexidad del conjunto de soluciones óptimas)

Dado el programa:

$$(P_1) \left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{Min } F(x) \\ \text{Max } F(x) \end{array} \right\} \\ \text{bajo las restricciones:} \\ \quad Y \left\{ \begin{array}{ll} g_i(x) \leq 0 & i \in I = \{1,\dots,m\} \\ \quad x \in R^n & \end{array} \right. \end{array} \right.$$

siendo $F(x)$ función $\left\{ \begin{array}{l} \text{convexa} \\ \text{cóncava} \end{array} \right\}$ en Y, $g_i(x)$ $i \in I$ funciones convexas en R^n. Entonces, se verifica que el conjunto de soluciones óptimas de (P_1) es convexo en R^n.

Teorema A.3. (Unicidad de la solución óptima)

Dado el programa:

$$(P_2)\begin{cases}\begin{Bmatrix}\text{Mín. } F(x)\\ \text{Máx. } F(x)\end{Bmatrix}\\ \text{bajo las restricciones:}\\ Y\quad\begin{cases}g_i(x)\leq 0 & i\in I=\{1,..,m\}\\ x\in R^n\end{cases}\end{cases}$$

siendo $F(x)$ función estrictamente $\begin{Bmatrix}\text{convexa}\\ \text{cóncava}\end{Bmatrix}$ en Y y $g_i(x)$ $i\in I$ funciones convexas en R^n. Si x^o es solución óptima de (P_2), entonces x^o es única.

Teorema A.4. (Optimo en la frontera de las restricciones)

Dado el programa:

$$(P_3)\begin{cases}\begin{Bmatrix}\text{Mín. } F(x)\\ \text{Máx. } F(x)\end{Bmatrix}\\ \text{bajo las restricciones:}\\ Y\quad\begin{cases}g_i(x)\leq 0 & i\in I=\{1,\dots,m\}\\ x\in R^n\end{cases}\end{cases}$$

siendo $F(x)$ función $\begin{Bmatrix}\text{cóncava}\\ \text{convexa}\end{Bmatrix}$ no constante en Y y $g_i(x)$ $i\in I$ funciones convexas en R^n. Si x^o es solución óptima de (P_3), entonces $x^o\in Fr(Y)$.

Teorema A.5. (Kuhn y Tucker)

Dado el programa:

$$(P_4)\begin{cases} \text{Min } F(x) \\ \text{bajo las restricciones:} \\ g_i(x) \leqslant 0 \quad i \in I = \{1,\dots,m\} \\ \quad x \in R^n \end{cases}$$

siendo $F(x)$ y $g_i(x)$ $i \in I$ funciones convexas diferenciables, además $\exists \bar{x} \in R^n$ tal que $g_i(\bar{x}) < 0$ $\forall i \in I$. Entonces, x^o es un óptimo global de (P_4) si y sólo si:

$$\exists \lambda^o \geqslant 0, \quad \lambda^o = (\lambda_1^o,\dots,\lambda_m^o)^T \in R^m \quad \text{tal que:}$$

1°) $\nabla_x L(x^o, \lambda^o) = 0$

2°) $\lambda_i^o g_i(x^o) = 0 \quad i \in I$

donde $L(x, \lambda) = F(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x)$ es la función de Lagrange asociada al programa (P_4) y $\nabla_x L$ el vector gradiente de la Lagrangiana respecto de la variable x.

Teorema A.6.

Si Y es cerrado en R^n con la topología usual, $Y \subsetneq R^n$, $\text{Int}(Y) \neq \emptyset$, entonces $\forall x^1 \in \text{Int}(Y)$, $\forall x^2 \in \text{Ext}(Y)$ se verifica que $L\left[x^1, x^2\right] \cap \text{Fr}(Y) \neq \emptyset$ (siendo $L\left[x^1, x^2\right] = \{x \in R^n / x = \lambda x^1 + (1-\lambda)x^2, \ 0 \leqslant \lambda \leqslant 1\}$

Demostración

La haremos por reducción al absurdo.

Supongamos que $\exists x^1 \in \mathrm{Int}(Y)$ y $\exists x^2 \in \mathrm{Ext}(Y)$ tales que $L[x^1, x^2] \cap \mathrm{Fr}(Y) = \emptyset$.

Sabemos, por una parte, que $L[x^1, x^2]$ es un conjunto conexo respecto de la topología relativa.

Vamos a probar que, bajo el supuesto anterior, $L[x^1, x^2]$ no es conexo.

Sean:

$$A = Y \cap L[x^1, x^2]$$

$$B = \mathrm{Ext}(Y) \cap L[x^1, x^2]$$

Se comprueba que:

$$A \cup B = (Y \cap L[x^1, x^2]) \cup (\mathrm{Ext}(Y) \cap L[x^1, x^2]) =$$

$$= (Y \cup \mathrm{Ext}(Y)) \cap L[x^1, x^2] = R^n \cap L[x^1, x^2] =$$

$$= L[x^1, x^2]$$

$$A \cap B = (Y \cap L[x^1, x^2]) \cap (\mathrm{Ext}(Y) \cap L[x^1, x^2]) =$$

$$= (Y \cap \mathrm{Ext}(Y)) \cap L[x^1, x^2] =$$

$$= \emptyset \cap L[x^1, x^2] = \emptyset$$

Probemos que A y B son abiertos en $L\left[x^1, x^2\right]$ con la topología relativa.

Demostremos, en primer lugar, que $\forall x^* \in A$, se sigue que $x^* \in \mathrm{Int}(Y)$.

En efecto,

si $\exists x^* \in A$ tal que $x^* \in \mathrm{Fr}(Y) \Rightarrow$

$\Rightarrow x^* \in A \cap \mathrm{Fr}(Y)$.

Pero,

$$A \cap \mathrm{Fr}(Y) = (Y \cap L\left[x^1,x^2\right]) \cap \mathrm{Fr}(Y) =$$

$$= L\left[x^1,x^2\right] \cap \mathrm{Fr}(Y)$$

Luego,

$$x^* \in L\left[x^1,x^2\right] \cap \mathrm{Fr}(Y) \Rightarrow$$

$$\Rightarrow L\left[x^1,x^2\right] \cap \mathrm{Fr}(Y) \neq \emptyset$$

en contra de lo supuesto. Contradicción, que nace de suponer que:

$\exists x^* \in A$ tal que $x^* \in \mathrm{Fr}(Y)$.

Luego,

$$\forall x^* \in A \Rightarrow x^* \in \mathrm{Int}(Y)$$

De lo anterior se concluye que:

$$A = Y \cap L\left[x^1,x^2\right] = \mathrm{Int}(Y) \cap L\left[x^1,x^2\right]$$

Demostremos, ahora, que A y B son abiertos en $L\left[x^1,x^2\right]$.

En efecto,

$$A = \mathrm{Int}(Y) \cap L\left[x^1,x^2\right]$$

Pero,

$$\mathrm{Int}(Y) \text{ es abierto en } R^n$$

luego, por definición de abiertos para la topología relativa, se concluye que:

$$A \text{ es abierto en } L\left[x^1,x^2\right]$$

Análogamente,

$$B = \mathrm{Ext}(Y) \cap L\left[x^1,x^2\right]$$

Pero,

$$\mathrm{Ext}(Y) \text{ es abierto en } R^n$$

luego, por el mismo razonamiento anterior, se sigue que:

$$B \text{ es abierto en } L\left[x^1,x^2\right]$$

Hemos llegado a la siguiente situación:

$$L\left[x^1,x^2\right] = A \cup B$$
$$\emptyset = A \cap B$$
$$A \text{ y } B \text{ abiertos en } L\left[x^1,x^2\right]$$

de la cual concluímos, por definición de conjunto conexo, que $L\left[x^1,x^2\right]$ no es conexo. Contradicción, que nace de suponer que:

$\exists x^1 \in \text{Int}(Y)$ y $\exists x^2 \in \text{Ext}(Y)$ tales que:

$$L\left[x^1,\ x^2\right] \cap \text{Fr}(Y) = \emptyset$$

Por tanto,

$\forall x^1 \in \text{Int}(Y)$ y $\forall x^2 \in \text{Ext}(Y)$ se verifica que:

$$L\left[x^1,\ x^2\right] \cap \text{Fr}(Y) \neq \emptyset$$

Queda así, demostrado el teorema A.6.

B I B L I O G R A F I A

[1] ABADIE, J. (1967): "Nonlinear Programming" North-Holland, Amsterdam.

[2] ABADIE, J. (1970): "Integer and Nonlinear Programming" North-Holland, Amsterdam.

[3] ADAMS, W.P. and SHERALI, H.D. (1986):"A Tight Linearization and Algorithm for Zero-One Quadratic Programming Problems", Management Science, Vol. 32, N°10, pp. 1274-1290.

[4] BALAS, E. and MAZZOLA, J.B. (1984): "Nonlinear 0-1 Programming: I. Linearization Techniques", Mathematical Programming, N°30, pp. 1-21.

[5] BALAS, E. and MAZZOLA, J.B. (1984): "Nonlinear 0-1 Programming: II. Dominance Relations and Algorithms", Mathematical Programming, N°30, pp. 22-45.

[6] BAZARAA, M.S. and SHETTY, C.M. (1979):"Nonlinear Programming. Theory and Algorithms", John Wiley, New York.

[7] BEALE, E.M.L. (1967): "Numerical Methods", Cap. 7 in Abadie, J. [1].

[8] BEJAR ALAMO, J. (1967): "Teoría Matemática de la Programación", Memorias de la Real Academia de Ciencias Exactas, Físicas y Naturales de Madrid. Serie de Ciencias Exactas, Tomo VI, Memoria N°2.

[9] BITRAN, G.R. (1979): "Theory and Algorithms for Linear Multiple Objective Programs with Zero One Variables", Mathematical Programming, N°17, pp. 362-390.

[10] BOLTIANSKI, V. (1976): "Commande Optimale des Systèmes Discrets", Editions Mir, Moscou, Traduction française.

[11] BORWEIN, J.M. and NIEUWENHUIS, J.W. (1984): "Two Kinds of Normality in Vector Optimization", Mathematical Programming, N°28, pp. 185-191.

[12] CAMERON, N. (1985): "Introduction to Linear and Convex Programming", Australian Mathematical Society, Lecture Series 1, Cambridge University Press, Cambridge.

[13] CHANKONG, V. and HAIMES, Y.Y. (1983): "Multiobjective Decision Making. Theory and Methodology", North-Holland, Amsterdam.

[14] DANTZING, G.B (1963): "Linear Programming and Extensions", Princeton University Press, Princeton, New Jersey.

[15] DYER, M.E. and PROLL, L.G. (1977): "An Algorithm for Determining All Extreme Points of a Convex Polytope", Mathematical Programming, N°12, pp. 81-96.

[16] ECH-CHERIF, A. and ECKER, J.G. (1984): A Class of Rank-Two Ellipsoid Algorithms for Convex Programming", Mathematical Programming, N°29, pp. 187-202.

[17] ECKER, J.G. and KOUADA, I.A. (1978): "Finding All Efficient Extreme Points for Multiple Objective Linear Programs", Mathematical Programming, N°14, pp. 249-261.

[18] FISHBURN, P.C. (1970): "Utility Theory for Decision Making", John Wiley, New York.

[19] FISHBURN, P.C. (1973): "The Theory of Social Choice", Princeton University Press, Princeton, New Jersey.

[20] FISHER, M.L. (1981): "The Lagrangian Relaxation Method for Solving Integer Programming Problems", Management Science, Vol. 27, N°1, pp. 1-18.

[21] GARFINKEL, R.S. and NEMHAUSER, G.L. (1972): "Integer Programming", John Wiley, New York.

[22] GLOVER, F. (1975): "New Results on Equivalent Integer Programming Formulations", Mathematical Programming, N°8, pp. 84-90.

[23] GLOVER, F. and WOOLSEY, E. (1973): "Further Reduction of Zero-One Polynomial Programming Problems to Zero-One Linear Programming Problems", Operations Research, Vol. 21, pp. 141-161.

[24] GLOVER, F. and WOOLSEY, E. (1974): "Converting the 0-1 Polynomial Programming Problem to a 0-1 Linear Program", Operations Research, Vol. 22, N° 1, pp. 180-182.

[25] GRANOT, D. and GRANOT, F. (1980): "Generalized Covering Relaxation for 0-1 Programs", Operations Research, Vol. 28, N°6, pp. 1442-1450.

[26] GRANOT, D., GRANOT, F. and KALLBER, J. (1979): "Covering Relaxation for Positive 0-1 Polynomial Programs", Management Science, Vol. 25, N°3, pp. 264-272.

[27] GRANOT, D., GRANOT, F. and VAESSEN, W. (1982): "An Accelerated Covering Relaxation Algorithm for Solving 0-1 Positive Polynomial Programs", Mathematical Programming, Vol. 22, N°3.

[28] GREENBERG, H. (1971): "Integer Programming", Academic Press, New York.

[29] GRUNBAUM, B. (1967): "Convex Polytopes", John Wiley, New York.

[30] HAMMER, P.L., HANSEN, P. and SIMEONE, B. (1984): "Roof Duality, Complementation and Persistency in Quadratic 0-1 Optimization", Mathematical Programming, N°28, pp. 121-155.

[31] HAZEN, G.B. and MORIN, T.L. (1983): "Optimality Conditions in Nonconical Multiple-Objective Programming", Journal of Optimization Theory and Applications, Vol. 40, N°1, pp. 25-60.

[32] HENIG, M.I. (1982): "Existence and Characterization of Efficient Decisions with Respect to Cones", Mathematical Programming, N°23, pp. 111-116.

[33] KEENEY, R.L. and RAIFFA, H. (1976): "Decisions with Multiple Objectives: Preferences and Value Trade Offs", John Wiley, New York.

[34] KUNZI, H.P., KRELLE, W. et OETTLI, W. (1969): "La Progammation non linéaire", Gauthier-Villars, Paris.

[35] LU, S.H. and WILLIAMS, A.C. (1987): "Roof Duality for Polynomial 0-1 Optimization", Mathematical Programming, N°37, pp. 357-360.

[36] LUENBERGER, D.G. (1973): "Introduction to Linear and Non Linear Programming", Addison-Wesley, Massachusetts.

[37] MANGASARIAN, O.L. (1969): "Nonlinear Programming", Mc Graw-Hill, New York.

[38] MARCOTTE, O. and SOLAND, R.M. (1986): "An Interactive Branch-and-Bound Algorithm for Multiple Criteria Optimization", Management Science, Vol. 32, N°1, pp. 61-75.

[39] MC CORMICK, G.P. (1983): "Nonlinear Programming. Theory and Applications", John Wiley, New York.

[40] MINOUX, M. (1983): "Programmation Mathématique. Théorie et Algorithmes". Tomes 1 et 2, Dunot Decision, Paris.

[41] MURTY, K.G. (1976): "Linear and Combinatorial Programming", John Wiley, New York.

[42] NAKAYAMA, H., SAWARAGI, Y. and TANINO, T.S. (1985): "Theory of Multiobjective Optimization", Academic Press, Vol. 176, Colec. Mathematics in Science and Engeneering, New York.

[43] RAO, S.S. (1978): "Optimization. Theory and Applications", Wiley Eastern Limited, New Delhi.

[44] ROCKAFELLAR, R.T. (1970): "Convex Analysis", Princeton University Press, Princeton, New Jersey.

[45] SHAPIRO, J.F. (1979): "Mathematical Programming: Structures and Algorithms", John Wiley, New York.

[46] SIMONNARD, M. (1962): "Programmation linéaire", Dunot, Paris.

[47] SOLAN, R.M. (1979): "Multicriteria Optimization: A General Characterization of Efficient Solutions", Decision Sciences, Vol. 10, pp. 26-38.

[48] TAHA, H.A. (1975): "Integer Programming, Theory, Applications, Computations", Academic Press, New York.

[49] WENDELL, R.E. and LEE, D.N. (1977): "Efficienty in Multiple Objective Optimization Problems", Mathematical Programming, N°12, pp. 406-414.

[50] WHITE, D.J. (1982): "Optimality and Efficiency", John Wiley, New York.

[51] YU, P.L. (1974):"Cone Convexity, Cone Extreme Points, and Nondominated Solutions in Decision Problems with Multiple Objectives", Journal of Optimization Theory and Applications, Vol. 14, pp. 319-377.

[52] ZELENY, M. (1982):"Multiple Criteria Decision Making", Mc Graw-Hill, New York.