

Resolución de problemas cuadráticos, separables y fraccionales basada en modificaciones del algoritmo Simplex.



María Martín González
Trabajo de fin de grado en Matemáticas
Universidad de Zaragoza

Director del trabajo: Pedro Mateo Collazos.
Diciembre de 2019

Prólogo

La investigación operativa es una rama de las matemáticas cuyo nacimiento formal puede situarse, aproximadamente, a mitad del siglo XX, una de sus herramientas más utilizadas es, sin duda, el algoritmo del Simplex, desarrollado por George Dantzing en 1947, herramienta básica para la resolución de problemas de programación lineal (PL). Desde entonces, esta rama no ha parado de ampliarse y desarrollar diferentes resultados y algoritmos para la resolución de problemas muy diversos, sobretodo con la aparición y mejora de los recursos informáticos, que hace disminuir significativamente el tiempo de cálculo de los problemas que esta rama propone gracias a los avances computacionales.

En la asignatura de Investigación Operativa del tercer curso del Grado de Matemáticas se ha estudiado en profundidad dicha herramienta para la resolución de problemas de programación lineal generales y algunas versiones para los problemas de transporte, transbordo y asignación (casos particulares de los PL). El algoritmo Simplex además de utilizarse como herramienta de resolución, también aparece como herramienta básica en el diseño de algoritmos para otros tipos de problemas relacionados con los PL, como es el caso del algoritmo de Ramificación y Acotación para problemas de programación lineal entera, algoritmo que se presentó brevemente en la asignatura.

Sin embargo, existen otras aplicaciones del método Simplex a otros problemas que se quedan fuera de la asignatura y que resultan interesantes. En este trabajo fin de grado se han seleccionado tres de estos problemas, los cuales se presentan y se desarrollan algunos de los algoritmos que pueden ser usados para resolverlos, mostrando su funcionamiento y su corrección. Los tres tipos de problema seleccionados tienen en común que su resolución se realiza mediante algoritmos que son modificaciones, en general, de los algoritmos Simplex estudiados en el grado.

En cada capítulo se plantea un problema no lineal diferente y se exponen las modificaciones y resultados necesarios para que, con técnicas de la programación lineal, se pueda dar la solución al problema y justificar su consistencia o inconsistencia bajo determinadas hipótesis. Además, en el anexo se pueden encontrar los conceptos y teoremas básicos sobre programación lineal y funciones convexas necesarios para abordar el trabajo y demostrar los resultados que presentamos a lo largo del documento.

El primer capítulo trata la resolución del problema de programación cuadrática (PPC), problema en el que la función objetivo es una función cuadrática y las restricciones son lineales. Para ello se define el problema lineal complementario (PLC) y el método de Lemke para obtener su solución ya que aplicando las condiciones de Kuhn-Tucker al problema cuadrático, este se reduce al PLC. Además, se demuestran los resultados necesarios para saber si el problema es consistente y, por lo tanto, tiene solución factible básica complementaria o, si por el contrario, el método de Lemke produce una terminación en rayo cuando el sistema sea inconsistente.

El problema de programación separable, otro de los tres seleccionados, se explica y desarrolla en el segundo capítulo. En este caso, no encontraremos la solución exacta, si no que a través de otro problema auxiliar construiremos una aproximación lineal al original. El problema de aproximación construido es lineal, luego podremos utilizar el algoritmo Simplex para resolverlo. Al final del capítulo, se establece también un algoritmo para la selección de los puntos que debemos tomar en la partición para que la

aproximación tomada sea óptima.

En el tercer capítulo se exponen los resultados relacionados con la programación fraccional. El problema fraccional es aquel en el que tratamos de minimizar el cociente de dos funciones afines sujeto a restricciones lineales. En esta parte del documento mostraremos dos algoritmos para la resolución de este problema bajo la condición de convexidad sobre S que nos darán la solución del sistema: el primero de ellos, el algoritmo de Gilmore y Gomory, es una modificación del método Simplex para problemas convexos, el segundo es el propio algoritmo Simplex, pues en este caso reformularemos el problema fraccional hasta alcanzar el formato de un problema de PL.

Para finalizar, a lo largo del cuarto capítulo se exponen las aplicaciones de los problemas de programación no lineal desarrollados y ejemplos de resolución mediante los algoritmos explicados en cada capítulo anterior para que sea más sencillo e ilustrativo comprenderlos.

Abstract

Through this work, we are going to talk about non linear programming. In particular, we are going to address three problems which have in common that their solutions can be produced by methods related to Simplex algorithm.

Through the first chapter, we study Lemke's method and its complementary linear programming convergence, finishing with the properties in our matrix associated to the system that allows us to ensure a basic feasible complementary solution. Afterwards, we apply the required modifications to the quadratic programming problem to transform it into a complementary linear problem and solve it with Lemke's method. Additionally, we adapt the convergence results about the complementary linear problem to the quadratic programming problem.

Throughout second chapter we are going to define the separable programming problem and build up a linear approximation to this problem that we solve with Simplex method. Subsequently, we see an algorithm that allows us to decide how many points we should take to make an optimal approximation to the linear functions.

Third chapter is dedicated to develop fractional programming. In this chapter, two different algorithms are shown in order to solve this problem under the convex assumption for the feasibility. The first one is called Gilmore and Gomory's algorithm which is a generalization of Simplex algorithm for convex problems, and the second one is the typical Simplex algorithm because we only reformulate the fractional problem to transform it into a linear programming problem.

Finally, fourth chapter is dedicated to show the application of every nonlinear programming problems shown during the research work. Moreover, we display an example of each one in order to easily understand the algorithms shown along this work.

Also we can find some useful results to remind some basic concepts about linear programming and others convex functions theorems' which are key to address the subject in the appendix.

The study of all this algorithms allows us to spread all the knowledge got with Operations Research and Stochastic Optimization subjects in third and fourth years of Maths degree.

Índice general

Prólogo	III
Abstract	V
1. Programación cuadrática.	1
1.1. El Problema lineal complementario.	1
1.1.1. Resolución del problema complementario lineal.	2
1.1.2. Convergencia del método de Lemke.	3
1.2. El problema de programación cuadrática.	5
1.3. Análisis de convergencia del método de Lemke.	5
2. Programación separable.	9
2.1. Aproximación del problema separable.	9
2.2. Resolución del problema lineal aproximado.	11
2.3. Selección o generación de los puntos de la partición.	12
2.3.1. Algoritmo de selección de puntos para la partición.	14
3. Programación lineal fraccional.	17
3.1. Algoritmo simplex para problemas convexos.	17
3.2. Resumen del algoritmo de programación fraccional de Gilmore y Gomory.	19
3.3. Método de Charnes y Cooper.	20
4. Ejemplos y aplicaciones.	21
4.1. Programación cuadrática.	21
4.2. Programación separable.	23
4.3. Programación lineal fraccional.	25
Bibliografía	29
Anexos	29
A. Anexo	31
A.1. El problema de programación lineal.	31
A.2. El algoritmo Simplex.	32
A.3. Resultados útiles para la programación cuadrática y fraccional.	33
A.4. Resultados útiles para la programación no lineal separable.	34

Capítulo 1

Programación cuadrática.

Los problemas de programación cuadrática son una familia muy importante de problemas que aparecen en diversos ámbitos. A lo largo de este capítulo vamos a definir el problema lineal complementario y el algoritmo del pivote complementario o método de Lemke que nos servirá de base para dar solución al problema de programación cuadrática. Además, enunciaremos y demostraremos resultados que, bajo ciertas hipótesis, nos garanticen la convergencia del algoritmo.¹

Consideramos el siguiente problema de programación cuadrática (PPC) en la forma:

$$\begin{aligned} \text{mín } & \mathbf{c}^t \mathbf{x} + \frac{1}{2} \mathbf{x}^t H \mathbf{x} \\ \text{suje } & \text{to a } \mathbf{A} \mathbf{x} \leq \mathbf{0} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

donde \mathbf{c} es un vector n -dimensional, \mathbf{b} es un vector m -dimensional, A es una matriz $m \times n$ y H es una matriz $n \times n$ simétrica. Ahora, veamos los resultados necesarios para llegar a su resolución.

1.1. El Problema lineal complementario.

La herramienta clave para la resolución del problema de programación cuadrática es el problema lineal complementario. A lo largo de esta sección, se define este problema y se da un algoritmo de resolución junto a los elementos necesarios para justificar su convergencia. El algoritmo se denomina método de Lemke o algoritmo del pivote complementario. Dicho algoritmo sigue un esquema de funcionamiento muy relacionado con el método Simplex de programación lineal. La unión entre este problema y el de programación cuadrática aparece cuando se imponen las condiciones de Kuhn-Tucker al problema cuadrático, cuestión que se desarrolla en la sección (1.2).

Definición 1.1. Sea M una matriz $p \times p$, y sea \mathbf{q} un vector dado. El problema lineal complementario (PLC) consiste en determinar los vectores \mathbf{w} y \mathbf{z} tales que:

$$\mathbf{w} - M\mathbf{z} = \mathbf{q} \tag{1.1}$$

$$w_j \geq 0, z_j \geq 0, j = 1, \dots, p. \tag{1.2}$$

$$w_j z_j = 0, j = 1, \dots, p. \tag{1.3}$$

A (w_j, z_j) se les llama variables complementarias. Una solución (\mathbf{w}, \mathbf{z}) al sistema se llama solución factible básica complementaria (SFBC) si cumple (1.1) y (1.2) y si sólo una de las variables del par (w_j, z_j) es básica para $j = 1, \dots, p$.

¹Señalar que en la bibliografía aparecen cuatro referencias [1-4] que no aparecen citadas explícitamente a lo largo del trabajo. Estas son las que han constituido la base del trabajo, de [1] se han extraído la mayoría de los resultados estudiados y de [2], [3] y [4] se ha obtenido la información general sobre los problemas, su aplicabilidad, etc. Las citas se omiten deliberadamente para evitar continuas repeticiones de ellas, si bien se pone de manifiesto por medio de esta nota.

1.1.1. Resolución del problema complementario lineal.

Para la resolución de este problema, vamos a necesitar construir un problema auxiliar cuando el vector \mathbf{q} dado sea negativo. Para ello introducimos una variable artificial de la siguiente forma, construyendo así el problema auxiliar:

$$\mathbf{w} - M\mathbf{z} - \mathbf{1}z_0 = \mathbf{q} \quad (1.4)$$

$$w_j \geq 0, z_j \geq 0, j = 1, \dots, p. \quad (1.5)$$

$$w_j z_j = 0, j = 1, \dots, p. \quad (1.6)$$

Tomando $z_0 = \max\{-q_i : 1 < i \leq p\}$, $\mathbf{z} = \mathbf{0}$ y $\mathbf{w} = \mathbf{q} + \mathbf{1}z_0$, siendo $\mathbf{1}$ el vector de unos, tenemos una solución inicial para este nuevo sistema (PLC1). El objetivo es conseguir que la variable artificial z_0 sea nula a través del algoritmo y así tendremos una solución al PLC.

Definición 1.2. Consideramos el sistema (PLC1), una solución factible $(\mathbf{w}, \mathbf{q}, z_0)$ para el sistema se llama solución casi factible básica complementaria (SCFBC) si es solución factible básica para (1.4) y (1.5) en el sistema (PLC1), si además ni w_s , ni z_s son básicas para algún $s \in \{1, \dots, p\}$ y solamente una variable del par complementario (w_j, z_j) es básica para $j = 1, \dots, p$.

Definición 1.3. Dada una SCFBC $(\mathbf{w}, \mathbf{z}, z_0)$, donde w_s y z_s son ambas no básicas, una SCFBC adyacente $(\widehat{\mathbf{w}}, \widehat{\mathbf{z}}, \widehat{z}_0)$ se obtiene introduciendo w_s o z_s en la base y sacando cualquier otra variable que no sea z_0 de esta.

Es claro por la definición anterior que cada SCFBC tiene, a lo sumo, dos soluciones adyacentes. Y además, si lo que ocurre es que el algoritmo nos conduce a sacar a z_0 de la base, obtendremos una terminación de rayo².

Algoritmo del pivote complementario o método de Lemke.

Este algoritmo desarrollado por Lemke en 1968 para resolver el problema complementario lineal se basa en ir construyendo soluciones adyacentes hasta que se encuentre una solución factible básica complementaria o, por el contrario, una terminación de rayo.

Paso inicial:

Si $\mathbf{q} \geq 0$, $(\mathbf{w}, \mathbf{z}) = (\mathbf{q}, \mathbf{0})$ es una solución factible básica complementaria. STOP.

Si no, creamos una tabla con (PLC1). Sea $-q_s = \max\{-q_i : 1 \leq i \leq p\}$ y obtenemos otra tabla pivotando la fila s y la columna z_0 . Entonces, las variables básicas son ahora z_0 y w_j $j = 1, \dots, p$ y $j \neq s$ que son no negativas todas. Llamamos $y_s = z_s$ y continuamos con el paso principal.

Paso principal:

1. Sea \mathbf{d}_s la nueva columna de la actual tabla de la variable y_s .

Si $\mathbf{d}_s \leq 0$, continuamos en (4). De lo contrario, buscamos el índice r donde se alcanza el siguiente mínimo:

$$\frac{\tilde{q}_r}{d_{rs}} = \min\left\{\frac{\tilde{q}_i}{d_{is}} : d_{is} > 0\right\}$$

Si la variable básica de la fila r es z_0 , continuamos en (3). De lo contrario, seguimos en (2).

²Una terminación en rayo hace referencia a un problema en el que la región factible es no acotada y su solución es una semirrecta.

2. La variable básica de la fila r será ahora w_l o z_l para algún $l \neq s$. La variable y_s entra a la base pivotando la fila r y la columna asignada a y_s . Si la variable que deja la base es w_l , entonces $y_s = z_l$, o viceversa, en caso de que sea z_l la que salga de la base. Volvemos al paso (1).
3. En este punto, y_s ha entrado a la base y z_0 la ha dejado. Pivotamos la columna asignada a y_s con la fila asignada a z_0 obteniendo una solución factible básica complementaria. STOP.
4. Finalización en rayo, debemos encontrar la dirección que llamaremos R . Sabemos que $R = \{(\mathbf{w}, \mathbf{z}, z_0) + \lambda \mathbf{d} : \lambda \geq 0\}$ verifica (1.4), (1.5) y (1.6) donde $(\mathbf{w}, \mathbf{z}, z_0)$ es la SCFBC que hemos obtenido en la última tabla y \mathbf{d} es una dirección del conjunto y tiene un 1 en la fila correspondiente a y_s , $-d_s$ en las filas de las variables básicas y cero en el resto. STOP.

1.1.2. Convergencia del método de Lemke.

Lema 1.4. *Suponemos que cada solución casi factible básica complementaria del sistema (PLC1) es no degenerada, es decir, que cada variable básica sea positiva y distinta de cero. Entonces ninguno de los puntos generados por el algoritmo se repite y, además, este debe terminar en un número finito de pasos.*

Demostración. Sea $(\mathbf{w}, \mathbf{z}, z_0)$ una solución casi factible básica complementaria, donde w_s y z_s son ambas no básicas. Entonces, esta solución tiene, a lo sumo, dos soluciones adyacentes casi factibles básicas complementarias, una introduciendo w_s en la base y la otra haciendo lo mismo con z_s .

Por hipótesis, cada solución es no degenerada, por tanto cada una de estas nuevas soluciones adyacentes es distinta a la primera dada. Queremos probar que ninguna de las soluciones nuevas generadas por el algoritmo se repite. Sea $(\mathbf{w}, \mathbf{z}, z_0)_v$ el punto generado en la iteración v . Por contradicción, suponemos que $(\mathbf{w}, \mathbf{z}, z_0)_{\kappa+\alpha} = (\mathbf{w}, \mathbf{z}, z_0)_\kappa$ para algún κ , $\alpha \in \mathbb{Z}$, donde $\kappa + \alpha$ es el índice mas pequeño donde se observa la primera repetición. Por hipótesis de no degeneración $\alpha > 1$ y, además, por la propia construcción del algoritmo $\alpha > 2$. Pero si $(\mathbf{w}, \mathbf{z}, z_0)_{\kappa+\alpha-1}$ es adyacente a $(\mathbf{w}, \mathbf{z}, z_0)_{\kappa+\alpha}$, tendrá que ser también adyacente a $(\mathbf{w}, \mathbf{z}, z_0)_\kappa$. Si $\kappa = 1$, y como $(\mathbf{w}, \mathbf{z}, z_0)$ tiene exactamente una solución adyacente, por el paso inicial del algoritmo, $(\mathbf{w}, \mathbf{z}, z_0)_{\kappa+\alpha-1} = (\mathbf{w}, \mathbf{z}, z_0)_{\kappa+1}$ y, por lo tanto, aparece una repetición en la iteración $\kappa + \alpha - 1$, que contradice a la hipótesis de que la primera repetición aparece en $\kappa + \alpha$. Si $\kappa \geq 2$, entonces $(\mathbf{w}, \mathbf{z}, z_0)_{\kappa+\alpha-1}$ es adyacente a $(\mathbf{w}, \mathbf{z}, z_0)_\kappa$ y, por tanto, debe ser igual a $(\mathbf{w}, \mathbf{z}, z_0)_{\kappa+1}$ o $(\mathbf{w}, \mathbf{z}, z_0)_{\kappa-1}$. En ese caso, tenemos una repetición en la iteración $\kappa + \alpha - 1$, que contradice de nuevo la hipótesis. Por tanto, las soluciones generadas por el algoritmo son todas diferentes, esto implica que haya solo un número finito de soluciones casi factibles básicas complementarias y como ninguna está repetida, el algoritmo terminará en un número finito de pasos con una solución factible básica complementaria o con terminación en rayo. □

Con el resultado previo, hemos demostrado que el algoritmo funciona. La cuestión ahora será ver si el algoritmo es consistente y acaba en una SFBC o es inconsistente, que entonces acabará con una terminación en rayo.

Lema 1.5. *Suponemos que cada solución casi factible básica complementaria del sistema (PLC1) es no degenerada. Suponemos también que el método de Lemke se usa para resolver el sistema y que obtenemos una terminación en rayo. En particular, tenemos la solución casi factible básica complementaria $(\tilde{\mathbf{w}}, \tilde{\mathbf{z}}, \tilde{z}_0)$ y la dirección $(\hat{\mathbf{w}}, \hat{\mathbf{z}}, \hat{z}_0)$, obteniendo $R = \{(\tilde{\mathbf{w}}, \tilde{\mathbf{z}}, \tilde{z}_0) + \lambda(\hat{\mathbf{w}}, \hat{\mathbf{z}}, \hat{z}_0) : \lambda \geq 0\}$. Entonces,*

1. $(\hat{\mathbf{w}}, \hat{\mathbf{z}}, \hat{z}_0) \neq (\mathbf{0}, \mathbf{0}, 0)$, $\hat{\mathbf{w}}, \hat{\mathbf{z}} \geq \mathbf{0}$, $\hat{z}_0 \geq 0$
2. $\hat{\mathbf{w}} - M\hat{\mathbf{z}} - \mathbf{1}\hat{z}_0 = \mathbf{0}$
3. $\tilde{\mathbf{w}}^t \tilde{\mathbf{z}} = \tilde{\mathbf{w}}^t \hat{\mathbf{z}} = \hat{\mathbf{w}}^t \tilde{\mathbf{z}} = \hat{\mathbf{w}}^t \hat{\mathbf{z}} = 0$

$$4. \hat{\mathbf{z}} \neq \mathbf{0}$$

$$5. \hat{\mathbf{z}}^t M \hat{\mathbf{z}} = -\mathbf{1} \hat{\mathbf{z}} \hat{z}_0 \leq 0$$

Demostración. Como $(\hat{\mathbf{w}}, \hat{\mathbf{z}}, \hat{z}_0)$ es una dirección del conjunto definido por (PLC1), entonces (1) y (2) son inmediatas por el teorema de caracterización de dirección extrema. Notar que cada terminación en rayo cumple las condiciones del sistema, en concreto la (3) y, por tanto, $0 = (\tilde{\mathbf{w}} + \lambda \hat{\mathbf{w}})^t (\tilde{\mathbf{z}} + \lambda \hat{\mathbf{z}})$ para cada $\lambda \geq 0$. Esto junto a la no negatividad de $\tilde{\mathbf{w}}, \hat{\mathbf{w}}, \tilde{\mathbf{z}}$ y $\hat{\mathbf{z}}$, implica que $\tilde{\mathbf{w}}^t \tilde{\mathbf{z}} = \tilde{\mathbf{w}}^t \hat{\mathbf{z}} = \hat{\mathbf{w}}^t \tilde{\mathbf{z}} = \hat{\mathbf{w}}^t \hat{\mathbf{z}} = 0$. Luego, se cumple (3).

Veamos que $\hat{\mathbf{z}} \neq \mathbf{0}$. Por contradicción, suponemos que $\hat{\mathbf{z}} = \mathbf{0}$, entonces $\hat{z}_0 > 0$, ya que si fuera nulo $\hat{\mathbf{w}} = \mathbf{0}$ que contradice a (1). Entonces, $\hat{\mathbf{w}} = \mathbf{1} \hat{z}_0$. Sabemos por (3) que $\hat{\mathbf{w}}^t \hat{\mathbf{z}} = 0$. Entonces, $\mathbf{1}^t \tilde{\mathbf{z}} = 0$ y como $\tilde{\mathbf{z}} \geq 0$, entonces $\tilde{\mathbf{z}} = \mathbf{0}$. Por la hipótesis de no degeneración, cada componente de $\tilde{\mathbf{z}}$ es no básica. Además, $\tilde{z}_0 = 0$ es básica y $\tilde{\mathbf{w}}$ tiene que tener exactamente $p - 1$ componentes básicas. En particular, como $\tilde{\mathbf{w}} - M \tilde{\mathbf{z}} - \mathbf{1} \tilde{z}_0 = \mathbf{q}$ y $\tilde{\mathbf{z}} = \mathbf{0}$, entonces $\tilde{z}_0 = \max\{-q_i : 1 \leq i \leq p\}$. Entonces, la SCFCB $(\tilde{\mathbf{w}}, \tilde{\mathbf{z}}, \tilde{z}_0)$ es la solución inicial, lo que es imposible por el lema 2.4. Por tanto, $\hat{\mathbf{z}} \neq \mathbf{0}$ y se cumple (4). Multiplicando $\hat{\mathbf{w}} - M \hat{\mathbf{z}} - \mathbf{1} \hat{z}_0 = \mathbf{0}$ por $\hat{\mathbf{z}}^t$ y sabiendo que se cumple (3) obtenemos (5). \square

Definición 1.6. Sea M una matriz $p \times p$, se dice que M es copositiva³ si $\mathbf{z}^t M \mathbf{z} \geq 0$ para cada $\mathbf{z} \geq 0$. Además, M se dice copositiva-plus si es copositiva y si $\mathbf{z} \geq 0$ y $\mathbf{z}^t M \mathbf{z} = 0$ implica $(M + M^t) \mathbf{z} = \mathbf{0}$.

Teorema 1.7. Suponemos que cada solución casi factible básica complementaria del sistema (PLC1) es no degenerada, es decir, que cada variable básica sea positiva y suponemos que M es copositiva-plus. Entonces el algoritmo finaliza en un número finito de pasos. En particular, si el sistema (1.1) (1.2) es consistente, el algoritmo termina con una solución factible básica complementaria al sistema (PLC). Por otro lado, si el sistema es inconsistente, el algoritmo finalizará con una terminación en rayo.

Demostración. Por el lema 1.4, el algoritmo finaliza en un número finito de pasos. Ahora suponemos que el algoritmo acaba en una terminación en rayo. En particular, suponemos que $(\tilde{\mathbf{w}}, \tilde{\mathbf{z}}, \tilde{z}_0)$ es la solución casi factible básica complementaria y $(\hat{\mathbf{w}}, \hat{\mathbf{z}}, \hat{z}_0)$ es la dirección extrema asociada de la última tabla. Por el lema 1.5, $\hat{\mathbf{z}} \geq \mathbf{0}$, $\hat{\mathbf{z}} \neq \mathbf{0}$ y $\hat{\mathbf{z}}^t M \hat{\mathbf{z}} = -\mathbf{1} \hat{\mathbf{z}} \hat{z}_0 \leq 0$. Pero como M es copositiva plus, $\mathbf{z}^t M \mathbf{z} \geq 0$ y, de lo anterior se sigue, $\hat{\mathbf{z}}^t M \hat{\mathbf{z}} = -\mathbf{1} \hat{\mathbf{z}} \hat{z}_0 = 0$. Como $\hat{\mathbf{z}} \neq \mathbf{0}$, $\hat{z}_0 = 0$ y como $(\hat{\mathbf{w}}, \hat{\mathbf{z}}, \hat{z}_0)$ es una dirección del conjunto, $\hat{\mathbf{w}} - M \hat{\mathbf{z}} - \mathbf{1} \hat{z}_0 = \mathbf{0}$ y, por tanto, $\hat{\mathbf{w}} = M \hat{\mathbf{z}}$.

Ahora, demostremos que $\mathbf{q}^t \hat{\mathbf{z}} < 0$. Como $\hat{\mathbf{z}}^t M \hat{\mathbf{z}} = 0$ y M es copositiva plus, $(M + M^t) \mathbf{z} = \mathbf{0}$. Esto junto a la parte (3) del lema 1.5 y que $\tilde{\mathbf{w}} = \mathbf{q} + M \tilde{\mathbf{z}} + \mathbf{1} \tilde{z}_0 = \mathbf{0}$ implica $0 = \tilde{\mathbf{w}}^t \hat{\mathbf{z}} = (\mathbf{q} + M \tilde{\mathbf{z}} + \mathbf{1} \tilde{z}_0)^t \hat{\mathbf{z}} = \mathbf{q}^t \hat{\mathbf{z}} - \tilde{\mathbf{z}}^t M \hat{\mathbf{z}} + \tilde{z}_0 \mathbf{1}^t \hat{\mathbf{z}}$. Como $\hat{\mathbf{w}} = M \hat{\mathbf{z}}$ y, por la parte (3) del lema 1.5, se sigue que $\tilde{\mathbf{z}}^t M \hat{\mathbf{z}} = 0$. Además, $\tilde{z}_0 \geq 0$ y $\mathbf{1}^t \hat{\mathbf{z}} > 0$ sustituyendo en la cadena de igualdades anterior se sigue que $\mathbf{q}^t \hat{\mathbf{z}} < 0$. Entonces, el sistema $M^t \mathbf{y} \leq \mathbf{0}$, $-\mathbf{1} \mathbf{y} \leq \mathbf{0}$ y, entonces, $\mathbf{q}^t \mathbf{y} < 0$ tiene solución $\mathbf{y} = \hat{\mathbf{z}}$. Por el teorema de Farkas, enunciado en A.4, el sistema $\mathbf{w} - M \mathbf{z} = \mathbf{q}$, $\mathbf{w} \geq \mathbf{0}$, $\mathbf{z} \geq \mathbf{0}$ no tiene solución.

Entonces, si el sistema (1.1), (1.2) es consistente, el algoritmo nos aportará una solución factible básica complementaria. Si el sistema (1.1), (1.2) fuera inconsistente, el algoritmo terminaría igualmente en un número finito de pasos con una terminación en rayo, como ya se demostró en el lema 1.4. \square

Corolario 1.8. Sea M matriz $p \times p$ que no tiene elementos negativos y los elementos de la diagonal son distintos de cero. Entonces el método de Lemke termina en un número finito de pasos con una solución factible básica complementaria.

Demostración. Si vemos que el problema es factible, basta tomar \mathbf{z} suficientemente grande, esto implica $\mathbf{w} = M \mathbf{z} + \mathbf{q} \geq \mathbf{0}$ y, por tanto, el sistema es consistente. Por otro lado, M cumpliendo la hipótesis es copositiva plus, por tanto, se cumple el teorema anterior y el algoritmo acabará en un número finito de pasos y con la solución del sistema, que por ser consistente acaba en una SFBC. \square

³Notar que este término es una generalización de una forma bilineal semidefinida positiva a cualquier matriz, sin necesidad de que esta sea simétrica.

1.2. El problema de programación cuadrática.

En esta sección se desarrolla el algoritmo de resolución del PPC, presentado al comienzo del capítulo. Para ello escribiremos las KKT asociadas y veremos como reformularlas para el PPC. A partir de ahí veremos bajo que condiciones el algoritmo del PLC nos proporciona la solución para el PPC. Denotamos a los vectores de los multiplicadores de Lagrange de las restricciones $A\mathbf{x} \leq \mathbf{0}$ y $\mathbf{x} \geq \mathbf{0}$, \mathbf{u} y \mathbf{v} respectivamente y al vector de las variables de holgura como \mathbf{y} . Quedando el sistema expresado como:

$$\begin{aligned} A\mathbf{x} + \mathbf{y} &= \mathbf{b} \\ -H\mathbf{x} - A^t\mathbf{u} + \mathbf{v} &= \mathbf{c} \\ \mathbf{x}^t\mathbf{v} = 0, \mathbf{u}^t\mathbf{y} &= 0 \\ \mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} &\geq \mathbf{0} \end{aligned} \quad (1.7)$$

Para llevar dichas condiciones al PLC definimos M , \mathbf{q} , \mathbf{w} y \mathbf{z} de la siguiente forma:

$$M = \begin{pmatrix} 0 & -A \\ A^t & H \end{pmatrix}, \mathbf{q} = \begin{pmatrix} \mathbf{b} \\ \mathbf{c} \end{pmatrix}, \mathbf{w} = \begin{pmatrix} \mathbf{y} \\ \mathbf{v} \end{pmatrix}, \mathbf{z} = \begin{pmatrix} \mathbf{u} \\ \mathbf{x} \end{pmatrix},$$

podemos reescribir las condiciones de Kuhn-Tucker como el sistema de complementariedad lineal:

$$\begin{aligned} \mathbf{w} - M\mathbf{z} &= \mathbf{q} \\ \mathbf{w}^t\mathbf{z} &= 0 \\ \mathbf{w}, \mathbf{z} &\geq \mathbf{0} \end{aligned} \quad (1.8)$$

Ahora, podemos usar el método de Lemke descrito en la sección anterior. para resolver este problema. En la siguiente sección veremos bajo que condiciones aseguramos la convergencia del método de Lemke para el problema de programación cuadrática.

1.3. Análisis de la convergencia del método de Lemke para el problema de programación cuadrática.

Durante el comienzo de este capítulo, hemos visto que podíamos resolver el problema lineal complementario y que el método de Lemke finalizará en un número finito de pasos en una solución única factible básica o terminación en rayo. En particular, si la matriz es copositiva plus, el algoritmo produce una SFBC. Vamos a realizar el mismo análisis con el problema cuadrático.

Teorema 1.9. *Sea A una matriz $m \times n$, y sea H una matriz simétrica $n \times n$. Si $\mathbf{y}^t H \mathbf{y} \geq 0$ para cada $\mathbf{y} \geq \mathbf{0}$, entonces la matriz $M = \begin{pmatrix} 0 & -A \\ A^t & H \end{pmatrix}$ es copositiva. Además, si $\mathbf{y} \geq \mathbf{0}$ y $\mathbf{y}^t H \mathbf{y} = 0$ implica que $H\mathbf{y} = \mathbf{0}$, entonces M es copositiva plus.*

Demostración. Vamos a demostrar que M es copositiva. Sea $\mathbf{z}^t = (\mathbf{x}^t, \mathbf{y}^t) \geq \mathbf{0}$. Entonces

$$\mathbf{z}^t M \mathbf{z} = (\mathbf{x}^t, \mathbf{y}^t) \begin{pmatrix} 0 & -A \\ A^t & H \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{y}^t H \mathbf{y} \quad (1.9)$$

Por hipótesis, $\mathbf{y}^t H \mathbf{y} \geq 0$, entonces M es copositiva.

Para demostrar que es copositiva plus, supongamos que $\mathbf{z} \geq \mathbf{0}$ y que $\mathbf{z}^t M \mathbf{z} = 0$, como $M + M^t = \begin{pmatrix} 0 & 0 \\ 0 & 2H \end{pmatrix}$ y entonces $(M + M^t)\mathbf{z} = \begin{pmatrix} 0 \\ 2H\mathbf{y} \end{pmatrix}$. Como $\mathbf{z}^t M \mathbf{z} = 0$, entonces por (1.8) $\mathbf{y}^t H \mathbf{y} = 0$. Por hipótesis, como $\mathbf{y} \geq \mathbf{0}$ e $\mathbf{y}^t H \mathbf{y} = 0$, entonces $H\mathbf{y} = \mathbf{0}$ y, por tanto $(M + M^t)\mathbf{z} = \mathbf{0}$, es decir, M es copositiva plus. \square

Corolario 1.10. *Si H es semidefinida positiva, entonces $\mathbf{y}^t H \mathbf{y} = 0$ implica que $H\mathbf{y} = \mathbf{0}$ y, por tanto, M es copositiva plus.*

Demostración. Basta con demostrar que $\mathbf{y}'H\mathbf{y} = 0$ implica que $H\mathbf{y} = \mathbf{0}$. Sea $H\mathbf{y} = \mathbf{d}$, y notemos que H es semidefinida positiva, tenemos que $0 \leq (\mathbf{y}' - \lambda\mathbf{d}')H(\mathbf{y} - \lambda\mathbf{d}) = \mathbf{y}'H\mathbf{y} + \lambda^2\mathbf{d}'H\mathbf{d} - 2\lambda\|\mathbf{d}\|^2$. Como $\mathbf{y}'H\mathbf{y} = 0$, si dividimos la desigualdad anterior por λ , y hacemos tender λ a 0^+ , se tiene que $\mathbf{0} = \|\mathbf{d}\|$ y, por tanto, $\mathbf{0} = \mathbf{d} = H\mathbf{y}$. \square

Corolario 1.11. *Si H tiene todos sus elementos no negativos, entonces M es copositiva. Además, si H tiene elementos no negativos con los elementos de las diagonal distintos de cero, entonces M es copositiva plus.*

Demostración. Si $\mathbf{y} \geq \mathbf{0}$ y $\mathbf{y}'H\mathbf{y} = 0$, entonces $\mathbf{y} = \mathbf{0}$, y entonces $H\mathbf{y} = \mathbf{0}$. Por el teorema 1.8, M es copositiva plus. \square

Teorema 1.12. *Consideramos el problema:*

$$\begin{aligned} & \text{mín } \mathbf{c}'\mathbf{x} + \frac{1}{2}\mathbf{x}'H\mathbf{x} \\ & \text{suje } \text{to a } \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Suponemos que la región factible es no vacía y que hacemos uso del método de Lemke para dar solución al sistema de Kuhn-Tucker $\mathbf{w} - M\mathbf{z} = \mathbf{q}$, $\mathbf{w}, \mathbf{z} \geq \mathbf{0}$, $\mathbf{w}'\mathbf{z} = 0$, donde

$$M = \begin{pmatrix} 0 & -A \\ A' & H \end{pmatrix}, \mathbf{q} = \begin{pmatrix} \mathbf{b} \\ \mathbf{c} \end{pmatrix}, \mathbf{w} = \begin{pmatrix} \mathbf{y} \\ \mathbf{v} \end{pmatrix}, \mathbf{z} = \begin{pmatrix} \mathbf{u} \\ \mathbf{x} \end{pmatrix},$$

siendo \mathbf{y} el vector de las variables de holgura, \mathbf{u} y \mathbf{v} el vector formado por los multiplicadores de Lagrange asociados a las restricciones $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ y $\mathbf{x} \geq \mathbf{0}$.

Si el sistema es no degenerado, bajo alguna de las siguientes condiciones, el algoritmo acaba en un número finito de iteraciones con una solución de Kuhn-Tucker.

1. H es semidefinida positiva y $\mathbf{c} = \mathbf{0}$.
2. H es definida positiva.
3. H tiene todos sus elementos no negativos y, además, los elementos de la diagonal son distintos de cero.

Además, si H es semidefinida positiva, la terminación en rayo implica que la solución óptima no está acotada.

Demostración. Suponemos que $H = H'$ porque si no fuera así podríamos reemplazar H por $\frac{1}{2}(H + H')$. Por el lema 1.4, el algoritmo de Lemke acaba en un número finito de pasos en punto de Kuhn-Tucker o en una terminación de rayo. Si H es semidefinida positiva, definida positiva o sus elementos son todos no negativos con los elementos de la diagonal distintos de cero, por los corolarios anteriores, M es copositiva plus. Ahora supongamos el algoritmo finaliza en una terminación de rayo. Por el teorema 1.9, como M es copositiva plus, esto solo es posible si el sistema siguiente no tiene solución:

$$\begin{aligned} \mathbf{A}\mathbf{x} + \mathbf{y} &= \mathbf{b} \\ -H\mathbf{x} - A'\mathbf{u} + \mathbf{v} &= \mathbf{c} \\ \mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} &\geq \mathbf{0} \end{aligned}$$

Por el Teorema de Farkas, enunciado en A.4, el siguiente sistema debe tener una solución (\mathbf{d}, \mathbf{e}) .

$$\mathbf{A}\mathbf{d} \leq \mathbf{0} \tag{1.10}$$

$$A^t \mathbf{e} - H\mathbf{d} \geq \mathbf{0} \quad (1.11)$$

$$\mathbf{e} \geq \mathbf{0} \quad (1.12)$$

$$\mathbf{d} \geq \mathbf{0} \quad (1.13)$$

$$\mathbf{b}^t \mathbf{e} + \mathbf{c}^t \mathbf{d} < 0 \quad (1.14)$$

Multiplicando (1.11) por $\mathbf{d}^t \geq 0$, y sabiendo que se satisfacen (1.10) y (1.12), se tiene que:

$$0 \leq \mathbf{d}^t A^t \mathbf{e} - \mathbf{d}^t H\mathbf{d} \leq 0 - \mathbf{d}^t H\mathbf{d}. \quad (1.15)$$

Por hipótesis, existen $\hat{\mathbf{x}}$ e $\hat{\mathbf{y}}$ tal que $A\hat{\mathbf{x}} + \hat{\mathbf{y}} = \mathbf{b}$, $\hat{\mathbf{x}}, \hat{\mathbf{y}} \geq \mathbf{0}$. Sustituyendo esta expresión para \mathbf{b} en (1.14) y teniendo en cuenta (1.11) y que $\hat{\mathbf{e}}, \hat{\mathbf{x}}, \hat{\mathbf{y}} \geq \mathbf{0}$, obtenemos:

$$0 > \mathbf{b}^t \mathbf{e} + \mathbf{c}^t \mathbf{d} = (A\hat{\mathbf{x}} + \hat{\mathbf{y}})^t \mathbf{e} + \mathbf{c}^t \mathbf{d} \geq \hat{\mathbf{x}}^t A^t \mathbf{e} + \mathbf{c}^t \mathbf{d} \geq \hat{\mathbf{x}}^t H\mathbf{d} + \mathbf{c}^t \mathbf{d}. \quad (1.16)$$

Supongamos que H es definida positiva, (1.15) implica que $\mathbf{d}^t H\mathbf{d} = 0$, y por el corolario 1.10 tenemos que $H\mathbf{d} = \mathbf{0}$. Por (1.16), tenemos que $\mathbf{c}^t \mathbf{d} < 0$. Por (1.10) y (1.13), \mathbf{d} es un dirección de la región factible, entonces $\hat{\mathbf{x}} + \lambda \mathbf{d}$ es factible $\forall \lambda \geq 0$. Ahora consideramos $f(\hat{\mathbf{x}} + \lambda \mathbf{d})$, donde $f(\mathbf{x}) = \mathbf{c}^t \mathbf{x} + \frac{1}{2} \mathbf{x}^t H \mathbf{x}$. Como $H\mathbf{d} = \mathbf{0}$, obtenemos:

$$f(\hat{\mathbf{x}} + \lambda \mathbf{d}) = f(\hat{\mathbf{x}}) + \lambda(\mathbf{c}^t + \mathbf{x}^t H)\mathbf{d} + \frac{1}{2} \lambda^2 \mathbf{d}^t H\mathbf{d} = f(\hat{\mathbf{x}}) + \lambda \mathbf{c}^t \mathbf{d}.$$

Como $\mathbf{c}^t \mathbf{d} < 0$, $f(\hat{\mathbf{x}} + \lambda \mathbf{d})$ tiende a $-\infty$ escogiendo un λ lo suficientemente grande, entonces tenemos una solución no acotada.

Por último, demostremos que no es posible obtener una terminación en rayo bajo las condiciones 1, 2 y 3 del teorema. Por contradicción, suponemos que obtenemos una terminación en rayo bajo alguna de las tres condiciones. Por (1.15), $\mathbf{d}^t H\mathbf{d} \leq 0$, bajo las condiciones 2 o 3 tendríamos $\mathbf{d} = \mathbf{0}$, lo que es imposible por (1.16), entonces se cumple 1. Si se cumple la condición 1, $H\mathbf{d} = \mathbf{0}$. Esto, junto a la hipótesis de que $\mathbf{c} = \mathbf{0}$ contradice (1.16), por tanto es imposible obtener una terminación en rayo.

Por tanto, el algoritmo debe acabar en un punto de Kuhn-Tucker bajo las condiciones consideradas. \square

Resumiendo, a lo largo de este capítulo hemos visto que bajo ciertas hipótesis se garantiza que el método de Lemke finalice en un número finito de pasos para el problema de programación cuadrática. Estas hipótesis sobre la matriz H pueden ser, o bien que esta sea semidefinida positiva y $\mathbf{c} = \mathbf{0}$, o bien que H sea definida positiva, o bien que todos sus elementos sean no negativos y los elementos de la diagonal distintos de cero, obteniendo así una solución óptima o una terminación en rayo, es decir, una solución no acotada.

Capítulo 2

Programación separable.

En este capítulo abordaremos el uso del método simplex para obtener soluciones a la programación no lineal donde su función objetivo y las restricciones pueden ser expresada como suma de funciones que solo dependen cada una de ellas de una variable. Denotaremos a este problema como programación no lineal separable (PNLS) y se podrá expresar de la siguiente manera:

$$\begin{aligned} & \text{mín} \sum_{j=1}^n f_j(x_j) \\ & \text{suje}to \text{ a } \sum_{j=1}^n g_{ij}(x_j) \leq p_i, \quad i = 1, \dots, m. \\ & \quad \quad \quad x_j \geq 0 \text{ para } j = 1, \dots, n. \end{aligned} \tag{2.1}$$

El algoritmo que se presenta a continuación se basa en construir una aproximación lineal del problema en cuestión, mostrando como proceder para conseguir una aproximación de la solución que tenga una calidad adecuada, es decir, que se pueda considerar igual a la solución del problema original.

2.1. Aproximación del problema separable.

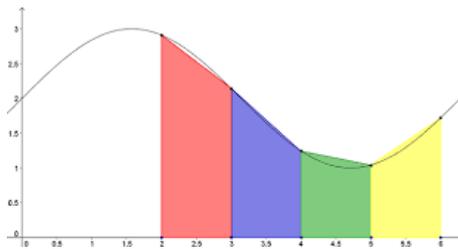
En este caso, como he comentado previamente, no se va a proporcionar la solución exacta al problema, sino que se propone un método para calcular una aproximación del óptimo. Vamos a definir un nuevo problema que aproxime al original PNLS. Este problema se obtiene cambiando cada función no lineal por una función a trozos lineal que la aproxime. Supongamos que queremos aproximar la función θ de variable μ de la que nos interesan sus valores sobre el intervalo $[a, b]$. Para construir su aproximación $\hat{\theta}$, partimos el intervalo en trozos más pequeños definidos por los puntos $a = \mu_1, \mu_2, \dots, \mu_k = b$. En cada subintervalo $[\mu_v, \mu_{v+1}]$ aproximamos θ como:

$$\hat{\theta} = \lambda \theta(\mu_v) + (1 - \lambda) \theta(\mu_{v+1})$$

De forma general, podemos aproximar la función θ sobre el intervalo $[a, b]$ por medio de los puntos $a = \mu_1, \mu_2, \dots, \mu_k = b$ por medio de la función a trozos $\hat{\theta}$ definida por

$$\hat{\theta}(\mu) = \sum_{v=1}^k \lambda_v \theta(\mu_v), \quad \sum_{v=1}^k \lambda_v = 1, \quad \lambda_v \geq 0, \quad v = 1, \dots, k, \tag{2.2}$$

donde como máximo dos λ_v son distintos de cero y, además, estos λ_v deben ser adyacentes.



Notar que los puntos pueden ser o no equidistantes y que mejorará la exactitud de la aproximación cuanto más puntos haya, pero también aumentará la dificultad del problema. Ahora definimos,

$$L = \{j : f_j \text{ y } g_{ij} \text{ lineales, } i = 1, \dots, m.\}$$

Entonces, para cada $j \notin L$, tomamos el intervalo $[a_j, b_j]$ con $a_j, b_j \geq 0$ y definimos los puntos x_{vj} para cada $v = 1, \dots, k_j$ donde $x_{1j} = a_j$ y $x_{k_j j} = b_j$. Usando esta partición, realizando la aproximación dada en (2.2), sustituimos f_j y g_{ij} por sus aproximaciones y nos queda el sistema:

$$\begin{aligned} \widehat{f}_j(x_j) &= \sum_{v=1}^{k_j} \lambda_{vj} f_j(x_{vj}), \quad j \notin L \\ \widehat{g}_{ij}(x_j) &= \sum_{v=1}^{k_j} \lambda_{vj} g_{ij}(x_{vj}), \quad i = 1, \dots, m \text{ y } j \notin L \\ \sum_{v=1}^{k_j} \lambda_{vj} &= 1, \quad j \notin L \\ \lambda_{vj} &\geq 0, \quad v = 1, \dots, k_j \text{ y } j \notin L. \end{aligned} \tag{2.3}$$

Haremos esto con las funciones dadas que no sean lineales, pero para $i = 1, \dots, m$, f_j y g_{ij} son lineales si $j \in L$. Por tanto, estas funciones no necesitan aproximación. Definiremos por tanto la aproximación al PNLs que denotaremos problema aproximado (PA) de la siguiente manera:

$$\begin{aligned} \text{mín} & \sum_{j \in L} f_j(x_j) + \sum_{j \notin L} \widehat{f}_j(x_j) \\ \text{sujeto a} & \sum_{j \in L} g_{ij}(x_j) + \sum_{j \notin L} \widehat{g}_{ij}(x_j) \leq p_i, \quad i = 1, \dots, m. \\ & x_j \geq 0, \quad j = 1, \dots, n, \end{aligned} \tag{2.4}$$

donde la función objetivo y las restricciones son funciones a trozos lineales aproximadas de la forma indicada. Si escribimos el problema con las expresiones explícitas de las funciones aproximadas obtenemos el problema lineal aproximado (PLA):

$$\begin{aligned} \text{mín} & \sum_{j \in L} f_j(x_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \lambda_{vj} f_j(x_{vj}) \\ \text{sujeto a} & \sum_{j \in L} g_{ij}(x_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \lambda_{vj} g_{ij}(x_{vj}) \leq p_i, \quad i = 1, \dots, m. \\ & \sum_{v=1}^{k_j} \lambda_{vj} = 1, \quad j \notin L \\ & \lambda_{vj} \geq 0, \quad v = 1, \dots, k_j \text{ y } j \notin L \\ & x_j \geq 0, \quad j \in L \\ & \text{Como máximo dos } \lambda_{vj} \text{ adyacentes son positivos para } j \notin L. \end{aligned} \tag{2.5}$$

2.2. Resolución del problema lineal aproximado.

Notar que hemos especificado antes que como máximo dos λ_{vj} adyacentes podían ser positivas. Para resolver el PLA podemos usar el método simplex con la siguiente regla de entrada:

Una variable no básica λ_{vj} entra en la base sólo si mejora la función objetivo y si la nueva base no tiene más de dos λ_{vj} adyacentes positivas para cada $j \notin L$.

El siguiente teorema nos dará las condiciones para poder descartar esta regla de entrada y tomar el método simplex tal y como lo usamos para el problema de programación lineal.

Teorema 2.1. Consideramos el PNL. Sea L tal y como lo hemos definido anteriormente y suponemos que para cada $j \notin L$, f_j es estrictamente convexa y que g_{ij} es convexa.¹ Suponemos que para cada $j \notin L$, sustituimos f_j y g_{ij} por sus aproximaciones lineales a trozos por medio de los puntos x_{vj} para $v = 1, \dots, k_j$: \hat{f}_j y \hat{g}_{ij} respectivamente, obteniendo el problema²:

$$\begin{aligned} & \min \sum_{j \in L} f_j(x_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \lambda_{vj} f_j(x_{vj}) \\ & \text{sujeta a } \sum_{j \in L} g_{ij}(x_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \lambda_{vj} g_{ij}(x_{vj}) \leq p_i, \quad i = 1, \dots, m. \\ & \sum_{v=1}^{k_j} \lambda_{vj} = 1, \quad j \notin L \\ & \lambda_{vj} \geq 0, \quad v = 1, \dots, k_j \text{ y } j \notin L \\ & x_j \geq 0, \quad j \in L. \end{aligned} \tag{2.6}$$

Sea \hat{x}_j para $j \in L$ y $\hat{\lambda}_{vj}$ para $v = 1, \dots, k_j$ y $j \notin L$ la solución al problema anterior. Entonces:

1. Para cada $j \notin L$, como máximo dos λ_{vj} son positivas y estas deben ser adyacentes.
2. Sea $\hat{x}_j = \sum_{v=1}^{k_j} \hat{\lambda}_{vj} x_{vj}$ para $j \notin L$, entonces el vector $\hat{\mathbf{x}}$ cuya componente j -ésima es \hat{x}_j para $j = 1, \dots, n$ es factible para el PNL.

Demostración. (1) Basta probar que para cada $j \notin L$, si $\hat{\lambda}_{lj}$ y $\hat{\lambda}_{pj}$ son positivas, entonces los puntos x_{lj} y x_{pj} deben ser adyacentes. Por contradicción, suponemos que existen $\hat{\lambda}_{lj}$ y $\hat{\lambda}_{pj} > 0$, donde los puntos x_{lj} y x_{pj} no son adyacentes. Entonces, existe un punto $x_{\gamma j} \geq 0 \in (x_{lj}, x_{pj})$ que puede ser expresado como $x_{\gamma j} = \alpha_1 x_{lj} + \alpha_2 x_{pj}$, donde $\alpha_1, \alpha_2 > 0$ y $\alpha_1 + \alpha_2 = 1$. Ahora, consideramos la solución óptima al problema definido en (3.6). Sean $u_i \geq 0$ para $i = 1, \dots, m$ los multiplicadores de Lagrange óptimos asociados a las primeras m restricciones y para cada $j \in L$, sea v_j el multiplicador de Lagrange óptimo asociado a la restricción $\sum_{v=1}^{k_j} \lambda_{vj} = 1$.

Entonces el siguiente subconjunto de condiciones necesarias de Kuhn-Tucker se satisfacen:

$$f_j(x_{\gamma j}) + \sum_{i=1}^m u_i g_{ij}(x_{\gamma j}) + v_j = 0 \tag{2.7}$$

$$f_j(x_{pj}) + \sum_{i=1}^m u_i g_{ij}(x_{pj}) + v_j = 0 \tag{2.8}$$

¹Una función f se dice convexa sobre S si $f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2)$. Se dirá que f es estrictamente convexa, si la desigualdad es estricta.

²(2.6) es igual que (2.1) excepto por el hecho de no obligar a que haya como máximo dos $\lambda_{vj} \neq 0$ y que estos sean adyacentes.

$$f_j(x_{vj}) + \sum_{i=1}^m u_i g_{ij}(x_{vj}) + v_j \geq 0 \text{ para } v = 1, \dots, k_j. \quad (2.9)$$

Por la hipótesis de convexidad estricta para f_j y convexidad para g_{ij} y por (2.7) y (2.8) tenemos:

$$f_j(x_{\gamma j}) + \sum_{i=1}^m u_i g_{ij}(x_{ij}) + v_j < \alpha_1 f_j(x_{ij}) + \alpha_2 f_j(x_{pj}) + \sum_{i=1}^m u_i [\alpha_1 g_{ij}(x_{ij}) + \alpha_2 g_{ij}(x_{pj})] + v_j = 0$$

Esto contradice a (2.9) para $v = \gamma$ y, por tanto, x_{ij} y x_{pj} deben ser adyacentes.

(2) Por la convexidad de g_{ij} para cada $j \notin L$ y para cada $i = 1, \dots, m$, y sabiendo que \hat{x}_j para $j \in L$ y $\hat{\lambda}_{vj}$ para $v = 1, \dots, k_j$ y $j \notin L$ satisfacen las restricciones de (2.6), tenemos:

$$g_i(\hat{\mathbf{x}}) = \sum_{j \in L} g_{ij}(\hat{x}_j) + \sum_{j \notin L} g_{ij}(\hat{x}_j) = \sum_{j \in L} g_{ij}(\hat{x}_j) + \sum_{j \notin L} g_{ij} \left(\sum_{v=1}^{k_j} \hat{\lambda}_{vj} x_{vj} \right) \leq \sum_{j \in L} g_{ij}(\hat{x}_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \hat{\lambda}_{vj} g_{ij}(x_{vj}) \leq p_i,$$

$i = 1, \dots, m$. Además, $\hat{x}_j \geq 0$ para $j \in L$, y $\hat{x}_j = \sum_{v=1}^{k_j} \hat{\lambda}_{vj} x_{vj} \geq 0$ para $j \notin L$, como $\hat{\lambda}_{vj}, x_{vj} \geq 0$ para $v = 1, \dots, k_j$ y $j \notin L$, se tiene $\hat{\mathbf{x}}$ es factible para el PNLS. \square

En resumen, hemos desarrollado el PLA y hemos demostrado que, sin necesidad de ninguna regla de entrada diferente a la conocida para el algoritmo Simplex, podemos encontrar la solución a este problema. Además, hemos visto que esta solución es válida para el PNLS.

2.3. Selección o generación de los puntos de la partición.

Como ya se ha dicho anteriormente, todo el procedimiento mostrado anteriormente depende del número de puntos que se tomen en la partición y lo cerca que estén los unos de los otros. Sin embargo, si tomamos más puntos, el número de variables en el PLA también crece. En este apartado se va a discutir cuando es necesario aumentar estos puntos, estableciendo un procedimiento o algoritmo para seleccionar esta partición.

Consideramos el PLA definido en (2.5). Sean x_{vj} para $v = 1, \dots, k_j$ y $j \notin L$ los puntos de la partición seleccionados. Sean \hat{x}_j para $j \in L$ y $\hat{\lambda}_{vj}$ para $v = 1, \dots, k_j$ y $j \notin L$ las soluciones al PLA. Además, sean $\hat{u}_i \geq 0$ para $i = 1, \dots, m$ los multiplicadores de Lagrange óptimos asociados a las primeras m restricciones del problema y sean \hat{v}_j para $j \notin L$ los multiplicadores de Lagrange asociado a la restricción $\sum_{v=1}^{k_j} \lambda_{vj} = 1$. Notar que las soluciones óptimas a todas estas variables cumplen las condiciones de Kuhn-Tucker para el PLA. Se quiere saber si se necesita considerar un punto nuevo en la partición para alguna de las variables x_j para $j \notin L$ para conseguir una aproximación mejor para el problema original, es decir, si tomáramos este nuevo punto en nuestro PLA, el valor de la función objetivo disminuirá. Para $j \notin L$, se considera un punto $x_{\gamma j}$, entonces se verifica:

$$f_j(x_{\gamma j}) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_{\gamma j}) + \hat{v}_j \geq 0, \quad (2.10)$$

si tomamos $\hat{\lambda}_{\gamma j} = 0$ se cumplirán las condiciones de Kuhn-Tucker para el nuevo PLA. Sin embargo, no sabemos dónde debe estar localizado este nuevo punto, para saberlo, si todos los x_j tal que $a_j \leq x_j \leq b_j$ para $j \notin L$ cumplen (2.10) resolvemos el siguiente problema de selección (PS) para cada $j \notin L$:

$$\begin{aligned} & \text{mín } f_j(x_j) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_j) + \hat{v}_j \\ & \text{sujeto a } a_j \leq x_j \leq b_j. \end{aligned}$$

El teorema siguiente concluirá en que si el valor mínimo de la función objetivo es no negativo $\forall j \notin L$, entonces la solución actual es óptima para el PNLS original; si no fuera así, es decir, si el valor mínimo de la función objetivo es negativo para algún $j \notin L$ se puede conseguir una aproximación mejor al PNLS. Notar que, gracias a los teoremas A.11 y A.12 incluidos en el Anexo, podemos calcular la desviación máxima de la solución aproximada respecto a la del PNLS, que es $c(1 + \sum_{i=1}^m \hat{u}_i)$, de forma que si disminuimos la distancia entre los puntos de la partición, c será mas pequeña y, por tanto, la aproximación será mejor.

Teorema 2.2. Consideramos el PNLS definido en (2.1). Sea $L = \{j : f_j \text{ y } g_{ij} \text{ son lineales para } i = 1, \dots, m.\}$. Suponemos, sin pérdida de generalidad, que $f_j(x_j)$ es de la forma $c_j x_j$ y $g_{ij}(x_j)$ es $a_{ij} x_j$ para $i = 1, \dots, m$ y para $j \in L$. Usando la partición x_{v_j} , $v = 1, \dots, k_j$ para $j \notin L$, sea el PLA definido en (2.5). Para $j \notin L$, suponemos que f_j y g_{ij} son convexas para $i=1, \dots, m$. Sea \hat{x}_j para $j \in L$ y $\hat{\lambda}_{v_j}$ para $v = 1, \dots, k_j$ y $j \notin L$ óptimos para el PLA con su correspondiente valor en la función objetivo \hat{z} . Sean $\hat{u}_i \geq 0$ para $i = 1, \dots, m$, los multiplicadores de Lagrange asociados a las primeras m restricciones, y sea \hat{v}_j para $j \notin L$ los multiplicadores de Lagrange asociados a la restricción $\sum_{v=1}^{k_j} \lambda_{v_j} = 1$ en el PLA. Para cada $j \notin L$, consideramos el siguiente problema:

$$\begin{aligned} \text{mín } f_j(x_j) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_j) \\ \text{suje to a } a_j \leq x_j \leq b_j \end{aligned} \quad (2.11)$$

donde $[a_j, b_j]$ con $a_j, b_j \geq 0$ es el intervalo de interés para x_j . Sea \tilde{z}_j el valor óptimo para la función objetivo sobre el problema (2.11). Entonces, se tiene:

1. $\sum_{j \notin L} \tilde{z}_j - \sum_{i=1}^m \hat{u}_i p_i \leq \sum_{j=1}^n f_j(\tilde{x}_j) \leq \sum_{j=1}^n f_j(\hat{x}_j) \leq \hat{z}$, donde $\hat{x}_j = \sum_{v=1}^{k_j} \hat{\lambda}_{v_j} x_{v_j}$ para $j \notin L$, entonces $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n)^t$ es una solución óptima al PNLS.
2. Si $\tilde{z}_j + \hat{v}_j \geq 0$ para $j \notin L$, entonces $\hat{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_n)^t$ es una solución óptima al PNLS. Además, $\sum_{j=1}^n f_j(\hat{x}_j) = \hat{z}$
3. Si $\tilde{z}_j + \hat{v}_j < 0$ para algún $j \notin L$, sea x_{γ_j} la solución óptima que hacía que $\tilde{z}_j < -\hat{v}_j$. Entonces, añadiendo el punto x_{γ_j} a la partición en el PLA, tendremos un nuevo problema de aproximación con un valor para la función objetivo menor o igual que \hat{z} .

Demostración. Como \hat{u}_i y \hat{v}_j son los multiplicadores de Lagrange óptimos asociados al PLA, se tiene:

$$c_j + \sum_{i=1}^m \hat{u}_i a_{ij} \geq 0 \text{ para } j \in L$$

Multiplicando por $x_j \geq 0$, y usando que $f_j(x_j) = c_j x_j$ y $g_{ij}(x_j) = a_{ij} x_j$, se tiene:

$$f_j(x_j) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_j) \geq 0 \text{ para } j \in L \text{ y } \forall x_j \geq 0. \quad (2.12)$$

Además, por definición de \tilde{z}_j , se tiene:

$$f_j(x_j) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_j) \geq \tilde{z}_j \text{ para } j \notin L \text{ y } \forall a_j \leq x_j \leq b_j. \quad (2.13)$$

Sumando (2.12) con $j \in L$ y (2.13) con $j \notin L$, y restando al resultado $\sum_{i=1}^m \hat{u}_i p_i$, se tiene:

$$\sum_{i=1}^n f_j(x_j) + \sum_{i=1}^m \hat{u}_i [g_{ij}(x_j) - p_i] \geq \sum_{j \notin L} \tilde{z}_j - \sum_{i=1}^m \hat{u}_i p_i \quad \forall a_j \leq x_j \leq b_j. \quad (2.14)$$

Como $a_j \leq x_j \leq b_j$, $\sum_{i=1}^n g_{ij}(\tilde{x}_j) \leq p_i$, y $\hat{u}_i \geq 0$, (2.14) implica que $\sum_{i=1}^n f_j(\tilde{x}_j) \geq \sum_{j \notin L} \tilde{z}_j - \sum_{i=1}^m \hat{u}_i p_i$, que es la primera desigualdad de la parte 1 del teorema. Por el teorema A.12, $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n)^t$ es factible

para el PNLS, entonces $\sum_{i=1}^n f_j(\tilde{x}_j) \leq \sum_{i=1}^n f_j(\hat{x}_j)$. Por último, por la convexidad de f_j para $j \notin L$, se tiene:

$$\sum_{i=1}^n f_j(\hat{x}_j) = \sum_{j \in L} f_j(\hat{x}_j) + \sum_{j \notin L} f_j(\hat{x}_j) = \sum_{j \in L} f_j(\hat{x}_j) + \sum_{j \notin L} f_j\left(\sum_{v=1}^{k_j} \hat{\lambda}_{vj} x_{vj}\right) \leq \sum_{j \in L} f_j(\hat{x}_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \hat{\lambda}_{vj} f_j(x_{vj}) = \hat{z},$$

y la parte 1 queda demostrada. Para probar el apartado 2, consideramos el PLA definido en (2.5). Gracias a las condiciones óptimas de holgura complementarias de Kuhn-Tucker se tiene:

$$f_j(\hat{x}_j) + \sum_{i=1}^m \hat{u}_i g_{ij}(\hat{x}_j) = 0 \text{ para } j \in L. \quad (2.15)$$

$$\hat{\lambda}_{vj} [f_j(x_{vj}) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_{vj}) + \hat{v}_j] = 0 \text{ para } j \notin L \text{ y } v = 1, \dots, k_j. \quad (2.16)$$

$$\hat{u}_i \left[\sum_{j \in L} g_{ij}(\hat{x}_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \hat{\lambda}_{vj} g_{ij}(x_{vj}) - p_i \right] = 0 \text{ para } i = 1, \dots, m. \quad (2.17)$$

Sumando (2.15) en $j \in L$, y (2.16) con $v = 1, \dots, k_j$ y $j \notin L$, tenemos:

$$\left[\sum_{j \in L} f_j(\hat{x}_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \hat{\lambda}_{vj} f_j(x_{vj}) \right] + \sum_{i=1}^m \hat{u}_i \left[\sum_{j \in L} g_{ij}(\hat{x}_j) + \sum_{j \notin L} \sum_{v=1}^{k_j} \hat{\lambda}_{vj} g_{ij}(x_{vj}) \right] + \sum_{j \notin L} \sum_{v=1}^{k_j} \hat{\lambda}_{vj} \hat{v}_j = 0. \quad (2.18)$$

El primer término de (2.18) es \hat{z} por definición y el segundo término es igual a $\sum_{i=1}^m \hat{u}_i p_i$ por (2.17). Además, $\sum_{v=1}^{k_j} \hat{\lambda}_{vj} = 1$ para $j \notin L$ y como λ_{vj} es factible para el PLA definido en (2.5), se tiene:

$$\hat{z} + \sum_{i=1}^m \hat{u}_i p_i + \sum_{j \notin L} \hat{v}_j = 0. \quad (2.19)$$

Y por la parte 1 del teorema:

$$\sum_{j \notin L} \tilde{z}_j - \sum_{i=1}^m \hat{u}_i p_i \leq \sum_{j=1}^n f_j(\tilde{x}_j). \quad (2.20)$$

Sumando (2.20) con (2.19), obtenemos $\sum_{j \notin L} (\tilde{z}_j + \hat{v}_j) + \hat{z} \leq \sum_{j=1}^n f_j(\tilde{x}_j)$. Pero por hipótesis de la parte 2, $\tilde{z}_j + \hat{v}_j \geq 0$ para $j \notin L$. Por tanto, $\hat{z} \leq \sum_{j=1}^n f_j(\tilde{x}_j)$ y usando la parte 1 del teorema, obtenemos, $\hat{z} \leq \sum_{j=1}^n f_j(\tilde{x}_j) \leq \sum_{j=1}^n f_j(\hat{x}_j) \leq \hat{z}$. Esto implica que $\sum_{j=1}^n f_j(\tilde{x}_j) = \sum_{j=1}^n f_j(\hat{x}_j)$. Como $\hat{\mathbf{X}} = (\hat{x}_1, \dots, \hat{x}_n)^t$ es factible para el PNLS, queda demostrada la parte 2.

Para probar la última parte del teorema, suponemos que x_{vj} es la solución óptima dada por $\tilde{z}_j < -v_j$. Tenemos que $f_j(x_{vj}) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_{vj}) + \hat{v}_j < 0$. Pero si el punto x_{vj} se incluye en la partición del PLA, entonces una de las condiciones de Kuhn-Tucker no se cumplirá: $f_j(x_{vj}) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_{vj}) + \hat{v}_j \geq 0$, si introducimos x_{vj} en las bases la función objetivo del PLA no es mayor a \hat{z} y queda demostrada la parte 3. \square

2.3.1. Algoritmo de selección de puntos para la partición.

Gracias al teorema 2.2 se puede desarrollar un algoritmo que nos permite decidir sobre el número de puntos que debemos tomar en la partición. A continuación, se va a describir este algoritmo que se construye apoyándose en los resultados que hemos visto en la sección anterior.

Paso inicial:

Se definen $a_j, b_j \geq 0$ tal que todas los puntos factibles cumplen $x_j \in [a_j, b_j]$ para $j \notin L$. Para cada $j \notin L$, elegimos un conjunto de puntos que nos den una partición de $[a_j, b_j]$ formado por k_j elementos

para $j \notin L$. Vamos al paso principal.

Paso principal:

1. Resolver el PLA definido en (2.5). Sea \hat{x}_j para $j \in L$ y $\hat{\lambda}_{vj}$ para $v = 1, \dots, k_j$ y $j \notin L$. Sean $\hat{u}_i \geq 0$ para $i = 1, \dots, m$, los multiplicadores de Lagrange asociados a las primeras m restricciones, y sean \hat{v}_j para $j \notin L$ los multiplicadores de Lagrange asociados a la restricción $\sum_{v=1}^{k_j} \lambda_{vj} = 1$. Ir a paso 2.
2. Para cada $j \notin L$, resolver el problema de $\min f_j(x_j) + \sum_{i=1}^m \hat{u}_i g_{ij}(x_j)$ sujeto a $a_j \leq x_j \leq b_j$. Sea \tilde{z}_j el valor óptimo de la función objetivo para $j \notin L$. Si $\tilde{z}_j + \hat{v}_j \geq 0 \forall j \notin L$, STOP; La solución óptima al problema es $\tilde{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_n)^t$ donde $\hat{x}_j = \sum_{v=1}^{k_j} \hat{\lambda}_{vj} x_{vj}$ para $j \in L$. De lo contrario, ir a paso 3.
3. Sea $\tilde{z}_p + \hat{v}_p = \min_{j \notin L} (\tilde{z}_j + \hat{v}_j) < 0$. Sea x_{vp} la solución óptima dada por $\tilde{z}_p < -\hat{v}_p$. Sea $v = k_p + 1$, sustituir k_p por $k_p + 1$ y volver al paso 1.

Capítulo 3

Programación lineal fraccional.

A lo largo de este capítulo vamos a considerar el problema en el que la función objetivo es la proporción entre dos funciones afines y las restricciones son lineales. Estos problemas se llaman problemas de programación lineal fraccional (PLF) y se expresan de la siguiente forma:

$$\begin{aligned} & \text{mín } \frac{\mathbf{p}^t \mathbf{x} + \alpha}{\mathbf{q}^t \mathbf{x} + \beta} \\ & \text{suje } a \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \tag{3.1}$$

donde \mathbf{p} , \mathbf{q} son vectores n -dimensionales, \mathbf{b} es un vector m -dimensional, A es una matriz $m \times n$, y α , β son escalares.

El capítulo se centra en construir dos alternativas para resolver este problema. La primera es el algoritmo Simplex para problemas convexos y la segunda el método de Charnes y Cooper, ambas basadas en el método Simplex. De hecho, la última de ellas consiste en reformular el PLF para poder aplicar el algoritmo Simplex en su formato original.

3.1. Algoritmo simplex para problemas convexos.

En esta sección, se van a exponer los resultados necesarios para justificar el uso del algoritmo Simplex para problemas convexos para resolver el PLF.

Lema 3.1. Sea $f(\mathbf{x}) = (\mathbf{p}^t \mathbf{x} + \alpha) / (\mathbf{q}^t \mathbf{x} + \beta)$, y sea S un conjunto convexo tal que $\mathbf{q}^t \mathbf{x} + \beta \neq 0$ en S . Entonces f es pseudoconvexa y pseudocóncava sobre S ¹.

Demostración. Primero, notemos que o $\mathbf{q}^t \mathbf{x} + \beta > 0 \forall \mathbf{x} \in S$ o $\mathbf{q}^t \mathbf{x} + \beta < 0 \forall \mathbf{x} \in S$. De otra forma, existiría un \mathbf{x}_1 y un \mathbf{x}_2 en S tal que $\mathbf{q}^t \mathbf{x}_1 + \beta > 0$ y $\mathbf{q}^t \mathbf{x}_2 + \beta < 0$, y existiría alguna combinación convexa entre \mathbf{x}_1 y \mathbf{x}_2 que sería igual a cero, lo que contradice nuestra hipótesis.

Primero, demostremos que f es pseudoconvexa. Suponemos que $\mathbf{x}_1, \mathbf{x}_2 \in S$ con $(\mathbf{x}_2 - \mathbf{x}_1)^t \nabla f(\mathbf{x}_1) \geq 0$. Tenemos que ver que $f(\mathbf{x}_2) \geq f(\mathbf{x}_1)$. Notar que

$$\nabla f(\mathbf{x}_1) = \frac{(\mathbf{q}^t \mathbf{x}_1 + \beta) \mathbf{p} - (\mathbf{p}^t \mathbf{x}_1 + \alpha) \mathbf{q}}{(\mathbf{q}^t \mathbf{x}_1 + \beta)^2}$$

Como $(\mathbf{x}_2 - \mathbf{x}_1)^t \nabla f(\mathbf{x}_1) \geq 0$ y $(\mathbf{q}^t \mathbf{x}_1 + \beta)^2 > 0$, tenemos que:

¹ f se dice pseudoconvexa si $\forall x_1, x_2 \in S$ con $(\mathbf{x}_2 - \mathbf{x}_1)^t \nabla f(\mathbf{x}_1) \geq 0$, implica $f(\mathbf{x}_2) \geq f(\mathbf{x}_1)$; f se dice pseudocóncava si $-f$ es pseudoconvexa.

$$0 \leq (\mathbf{x}_2 - \mathbf{x}_1)^t [(\mathbf{q}'\mathbf{x}_1 + \beta)\mathbf{p} - (\mathbf{p}'\mathbf{x}_1 + \alpha)\mathbf{q}] = (\mathbf{p}'\mathbf{x}_2 + \alpha)(\mathbf{q}'\mathbf{x}_1 + \beta) - (\mathbf{q}'\mathbf{x}_2 + \beta)(\mathbf{p}'\mathbf{x}_1 + \alpha).$$

Entonces, $(\mathbf{p}'\mathbf{x}_2 + \alpha)(\mathbf{q}'\mathbf{x}_1 + \beta) \geq (\mathbf{q}'\mathbf{x}_2 + \beta)(\mathbf{p}'\mathbf{x}_1 + \alpha)$. Pero, $\mathbf{q}'\mathbf{x}_1 + \beta$ y $\mathbf{q}'\mathbf{x}_2 + \beta$ son ambos positivos o negativos, dividiendo la expresión anterior por $(\mathbf{q}'\mathbf{x}_1 + \beta)(\mathbf{q}'\mathbf{x}_2 + \beta) > 0$, obtenemos:

$$\frac{\mathbf{p}'\mathbf{x}_2 + \alpha}{\mathbf{q}'\mathbf{x}_2 + \beta} \geq \frac{\mathbf{p}'\mathbf{x}_1 + \alpha}{\mathbf{q}'\mathbf{x}_1 + \beta},$$

que implica $f(\mathbf{x}_2) \geq f(\mathbf{x}_1)$. Entonces, f es pseudoconvexa.

De manera análoga podemos demostrar que $(\mathbf{x}_2 - \mathbf{x}_1)^t \nabla f(\mathbf{x}_1) \leq 0$ implica que $f(\mathbf{x}_2) \leq f(\mathbf{x}_1)$, y por tanto, f es pseudocóncava. \square

Del lema anterior se desprenden las siguientes consideraciones:

- Como la función objetivo es pseudoconvexa y pseudocóncava sobre S , por el teorema A.8 es también cuasiconvexa, cuasicóncava, estrictamente cuasiconvexa y estrictamente cuasicóncava.²
- Como la función objetivo es pseudoconvexa y pseudocóncava sobre S , por el teorema A.9, un punto que satisfaga las condiciones de Kuhn-Tucker para el problema de mínimo es también un mínimo global sobre la región factible. De manera análoga, un punto que satisfaga las condiciones de Kuhn-Tucker para el problema de máximo es también un máximo global sobre la región factible.
- Como la función objetivo es estrictamente pseudoconvexa y pseudocóncava sobre S , por el teorema A.7, un mínimo local es también global sobre la región factible. De manera análoga, un máximo local es también global sobre la región factible.
- Como la función objetivo es pseudoconvexa y pseudocóncava sobre S , si la región factible está acotada, por el teorema A.6, la función objetivo alcanza su máximo y su mínimo en algún punto extremo de la región factible.

Todas estas observaciones sobre la función objetivo f nos permiten desarrollar procedimientos para resolver el problema PLF. En particular, en las siguientes secciones vamos a describir métodos que parten del método simplex para el problema de programación lineal.

Dada la estructura especial de la función objetivo f para el PLF, el algoritmo Simplex para problemas convexos, que es una modificación del método simplex para programación lineal, nos permitirá encontrar la solución óptima.

Notar que se trata de un algoritmo Simplex que en lugar de trabajar con los costes marginales, va a trabajar con las proyecciones de vector gradiente en el punto extremo actual en las diferentes aristas que parten del punto.

Lema 3.2. Sea $f(\mathbf{x}) = \frac{\mathbf{p}'\mathbf{x} + \alpha}{\mathbf{q}'\mathbf{x} + \beta}$, y sea S un conjunto convexo. Además, suponemos que $\mathbf{q}'\mathbf{x} + \beta \neq 0$ sobre S . Dado $\mathbf{x} \in S$, sea \mathbf{d} tal que $\nabla f(\mathbf{x})^t \mathbf{d} < 0$. Entonces, $f(\mathbf{x} + \lambda \mathbf{d})$ es una función decreciente de λ .

Demostración. Notar que

$$\nabla f(\mathbf{y}) = \frac{(\mathbf{q}'\mathbf{y} + \beta)\mathbf{p} - (\mathbf{p}'\mathbf{y} + \alpha)\mathbf{q}}{(\mathbf{q}'\mathbf{y} + \beta)^2}$$

² f se dice cuasiconvexa si para cada $\mathbf{x}_1, \mathbf{x}_2 \in S$, se tiene que $f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$ para cada $\lambda \in (0, 1)$. f es cuasicóncava si $-f$ es cuasiconvexa. f se dice estrictamente cuasiconvexa si la desigualdad es estricta.

Tomando $\mathbf{y} = \mathbf{x} + \lambda \mathbf{d}$, $s = [\mathbf{q}'(\mathbf{x} + \lambda \mathbf{d}) + \beta]^2 > 0$, y $s' = (\mathbf{q}'\mathbf{x} + \beta)^2 > 0$, obtenemos:

$$\nabla f(\mathbf{x} + \lambda \mathbf{d}) = \frac{[\mathbf{q}'(\mathbf{x} + \lambda \mathbf{d}) + \beta]\mathbf{p} - [\mathbf{p}'(\mathbf{x} + \lambda \mathbf{d}) + \alpha]\mathbf{q}}{s} = \frac{s'}{s}\nabla f(\mathbf{x}) + \frac{\lambda}{s}[(\mathbf{q}'\mathbf{d})\mathbf{p} - (\mathbf{p}'\mathbf{d})\mathbf{q}].$$

Por tanto,

$$\nabla f(\mathbf{x} + \lambda \mathbf{d})' \mathbf{d} = \frac{s'}{s}\nabla f(\mathbf{x})' \mathbf{d} + \frac{\lambda}{s}[(\mathbf{q}'\mathbf{d})(\mathbf{p}'\mathbf{d}) - (\mathbf{p}'\mathbf{d})(\mathbf{q}'\mathbf{d})] = \frac{s'}{s}\nabla f(\mathbf{x})' \mathbf{d}. \quad (3.2)$$

Tomando $\theta(\lambda) = f(\mathbf{x} + \lambda \mathbf{d})$. Entonces, por (3.2), $\theta'(\lambda) = \nabla f(\mathbf{x} + \lambda \mathbf{d})' \mathbf{d} < 0 \forall \lambda$ y se tiene el resultado. \square

Para resumir, dado un punto extremo \mathbf{x} y la dirección \mathbf{d} con $\nabla f(\mathbf{x})' \mathbf{d} < 0$, no es necesario minimizar a lo largo de la dirección \mathbf{d} , ya que $f(\mathbf{x} + \lambda \mathbf{d})$ es una función decreciente de λ basta con movernos a lo largo de \mathbf{d} hasta alcanzar otro punto extremo y repetir el proceso.

3.2. Resumen del algoritmo de programación fraccional de Gilmore y Gomory.

En esta sección vamos a presentar el algoritmo de Gilmore y Gomory [1963] para resolver el PLF expresado como en (3.1). Suponemos que el conjunto S está acotado y que $\mathbf{q}'\mathbf{x} + \beta \neq 0 \forall \mathbf{x} \in S$.

Paso inicial:

Se busca una solución factible básica \mathbf{x}_1 para el sistema $A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$. Se escribe la tabla correspondiente representada por $\mathbf{x}_B + B^{-1}N\mathbf{x}_N = B^{-1}\mathbf{b}$. Tomamos $k=1$ y vamos al paso principal.

Paso principal:

1. Calcular el vector $\mathbf{r}_N^t = \nabla_N f(\mathbf{x}_k)^t - \nabla_B f(\mathbf{x}_k)^t B^{-1}N$. Si $\mathbf{r}_N \geq 0$ STOP; el punto \mathbf{x}_k es la solución óptima. En otro caso, ir al paso 2.
2. Sea $-r_j = \max\{-r_i : r_i \leq 0\}$, donde r_i es la i -ésima componente del vector \mathbf{r}_N . Seleccionamos la variable \mathbf{x}_B , que abandona la base según el siguiente criterio:

$$\frac{\tilde{b}_r}{y_{rj}} = \min_{1 \leq i \leq m} \left\{ \frac{\tilde{b}_i}{y_{ij}} : y_{ij} > 0 \right\}$$

donde $\tilde{\mathbf{b}} = B^{-1}\mathbf{b}, y_j = B^{-1}a_j$, y a_j es la j -ésima columna de A . Ir al paso 3.

3. Se sustituye la variable x_{B_r} , por la variable x_j . Se escribe la tabla correspondiente pivotando con y_{rj} . Tomamos la solución \mathbf{x}_{k+1} . Sustituimos k por $k+1$, vamos al paso 1.

Ahora, asumimos que $\mathbf{x}_B > \mathbf{0}$ para cada punto extremo. El algoritmo se mueve de un punto a otro. Por el lema 3.2 y la hipótesis de no degeneración, la función objetivo decrece estrictamente en cada iteración y, por tanto, los puntos obtenidos mediante el algoritmo son todos diferentes. Hay un número finito de puntos extremos, lo que implica que el algoritmo termina en un número finito de pasos. Además, al finalizar, el gradiente es no negativo se genera un punto de Kuhn Tucker por el teorema A.11 y, por el lema 3.1 este punto es óptimo.

3.3. Método de Charnes y Cooper.

En esta sección se va a describir otro procedimiento haciendo uso del método simplex para resolver el problema de programación lineal fraccional. A diferencia del método anterior, este método no se basa en desarrollar un nuevo algoritmo, sino en reformular el problema para que tome el formato de PL y poder aplicar el algoritmo Simplex. Consideramos el siguiente problema:

$$\begin{aligned} & \text{mín } \frac{\mathbf{p}'\mathbf{x} + \alpha}{\mathbf{q}'\mathbf{x} + \beta} \\ & \text{suje } a \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (3.3)$$

Supongamos que el conjunto $S = \{\mathbf{x} : \mathbf{Ax} \leq \mathbf{b} \text{ y } \mathbf{x} \geq 0\}$ es compacto y que $\mathbf{q}'\mathbf{x} + \beta > 0 \forall \mathbf{x} \in S$. Definimos $z = 1/(\mathbf{q}'\mathbf{x} + \beta)$ e $\mathbf{y} = z\mathbf{x}$, con estas nuevas dos variables el problema (3.2) se transforma en:

$$\begin{aligned} & \text{mín } \mathbf{p}'\mathbf{y} + \alpha z \\ & \text{suje } a \mathbf{Ay} - \mathbf{bz} \leq \mathbf{0} \\ & \mathbf{q}'\mathbf{y} + \beta z = 1 \\ & \mathbf{y} \geq \mathbf{0} \\ & z \geq 0. \end{aligned} \quad (3.4)$$

Lema 3.3. Sea $S = \{\mathbf{x} : \mathbf{Ax} \leq \mathbf{b} \text{ y } \mathbf{x} \geq 0\}$ compacto, supongamos que $\mathbf{q}'\mathbf{x} + \beta > 0 \forall \mathbf{x} \in S$. Definimos $z = 1/(\mathbf{q}'\mathbf{x} + \beta)$ e $\mathbf{y} = z\mathbf{x}$ y consideramos el problema definido en (3.4).

1. Si (\mathbf{y}, z) es una solución factible para el problema (3.4), entonces $z > 0$.
2. Si $(\tilde{\mathbf{y}}, \tilde{z})$ es una solución óptima al problema lineal (3.4), entonces $\tilde{\mathbf{x}} = \frac{\tilde{\mathbf{y}}}{\tilde{z}}$ es una solución óptima al problema lineal fraccional.

Demostración. (1) Por contradicción, sea $z = 0$, entonces $\mathbf{y} \neq \mathbf{0}$ tal que $\mathbf{Ay} \leq \mathbf{0}$ e $\mathbf{y} \geq 0$, lo que significa que \mathbf{y} es una dirección de S , lo que contradice a la hipótesis de que S está acotado.

(2) Notar que $\mathbf{A}\tilde{\mathbf{x}} \leq \mathbf{0}$ y $\tilde{\mathbf{x}} \geq 0$, por tanto $\tilde{\mathbf{x}}$ es solución factible para el PLF. Ahora, demostremos que $\tilde{\mathbf{x}}$ es óptima: sea \mathbf{x} tal que $\mathbf{Ax} \leq \mathbf{b}$ y $\mathbf{x} \geq 0$. Por hipótesis $\mathbf{q}'\mathbf{x} + \beta > 0 \forall \mathbf{x} \in S$, y el vector (\mathbf{y}, z) es una solución factible al problema lineal (3.4), donde $\mathbf{y} = \frac{\mathbf{x}}{(\mathbf{q}'\mathbf{x} + \beta)}$ y $z = \frac{1}{(\mathbf{q}'\mathbf{x} + \beta)}$. Como $(\tilde{\mathbf{y}}, \tilde{z})$ es una solución óptima al problema lineal (3.4), $\mathbf{p}'\tilde{\mathbf{y}} + \alpha\tilde{z} \leq \mathbf{p}'\mathbf{y} + \alpha z$. Sustituyendo $\tilde{\mathbf{y}}, \mathbf{y}$ y z , la desigualdad nos queda $\tilde{z}(\mathbf{p}'\tilde{\mathbf{x}} + \alpha) \leq \frac{(\mathbf{p}'\mathbf{x} + \alpha)}{(\mathbf{q}'\mathbf{x} + \beta)}$ y obtenemos el resultado dividiendo el lado izquierdo de la desigualdad por $1 = \mathbf{q}'\tilde{\mathbf{y}} + \beta\tilde{z}$.

De manera análoga, si $\mathbf{q}'\mathbf{x} + \beta < 0 \forall \mathbf{x} \in S$ podemos hacer el cambio $-z = \frac{1}{(\mathbf{q}'\mathbf{x} + \beta)}$, e $\mathbf{y} = z\mathbf{x}$, obteniendo el siguiente problema lineal:

$$\begin{aligned} & \text{mín } -\mathbf{p}'\mathbf{y} - \alpha z \\ & \text{suje } a \mathbf{Ay} - \mathbf{bz} \leq \mathbf{0} \\ & -\mathbf{q}'\mathbf{y} - \beta z = 1 \\ & \mathbf{y} \geq \mathbf{0} \\ & z \geq 0. \end{aligned} \quad (3.5)$$

Análogamente, si $(\tilde{\mathbf{y}}, \tilde{z})$ es solución para el PPL (3.5), entonces $\tilde{\mathbf{x}} = \frac{\tilde{\mathbf{y}}}{\tilde{z}}$ es solución para el PLF. \square

Resumiendo, hemos probado que el PLF puede ser resuelto como un PPL con una variable y una restricción adicionales. La forma del problema lineal que usemos para la resolución dependerá de si $\mathbf{q}'\mathbf{x} + \beta > 0 \forall \mathbf{x} \in S$ o $\mathbf{q}'\mathbf{x} + \beta < 0 \forall \mathbf{x} \in S$.

En otro caso, si existiera \mathbf{x}_1 y un \mathbf{x}_2 en S tal que $\mathbf{q}'\mathbf{x}_1 + \beta > 0$ y $\mathbf{q}'\mathbf{x}_2 + \beta < 0$, entonces el problema tendría, obviamente, solución no acotada.

Capítulo 4

Ejemplos y aplicaciones.

A lo largo de este capítulo mostraremos ejemplos que permitan entender mas fácilmente todos los algoritmos desarrollados a lo largo de los capítulos anteriores para resolver los problemas expuestos. Trataremos también de mostrar diferentes aplicaciones y entornos donde podemos encontrar este tipo de problemas.

4.1. Programación cuadrática.

El desarrollo de la programación cuadrática comenzó en los años cincuenta, fecha desde la cuál mantiene su importancia. En los últimos años, los avances realizados en este campo son sobretodo el desarrollo de algoritmos eficientes que resuelven el problema de programación cuadrática como por ejemplo, los métodos del punto interior o el método de Newton. De todos estos algoritmos podemos encontrar una descripción completa en [10] y [6]. También es clásico el algoritmo de Wolfe para programación cuadrática que es una versión modificada del algoritmo de Gilmore y Gomory en el que la variable que entra a la base es la de mayor coste marginal siempre que no se junte en la base con su variable complementaria. En cualquier caso, este algoritmo requiere que la hessiana de la función objetivo sea definida positiva, lo cuál es mas restrictivo que las hipótesis que plantea el método de Lemke que ha sido presentado en el trabajo. Se puede encontrar más información de este algoritmo en [9].

Respecto a las aplicaciones de este problema, son variadas y pertenecen a muchos campos. Quizá, las más conocidas o utilizadas en el presente es la economía y empresa, dado que estos problemas aparecen con frecuencia en la selección de inversiones, análisis de políticas de impuestos y también en otros ámbitos más generales como es el control óptimo, técnica aplicada en sistemas que evolucionan en el tiempo y podrían ser influenciados o modificados por fuerzas externas. Todo esto lo podemos encontrar desarrollado de forma mas extensa en [7].

Ejemplo 1. *A continuación se muestra la resolución paso a paso de un problema de programación cuadrática.*

Consideramos el siguiente PPC:

$$\begin{aligned} \text{mín} \quad & -2x_1 - 6x_2 + x_1^2 - 2x_1x_2 + 2x_2^2 \\ \text{suje}to \quad & a \quad x_1 + x_2 \leq 2 \\ & -x_1 + 2x_2 \leq 0 \\ & x_1, x_2 \geq 0. \end{aligned}$$

Primero, comprobemos que el método de Lemke finalizará en un número finito de pasos, para ello escribimos:

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}, H = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \mathbf{c} = \begin{pmatrix} 6 \\ -2 \end{pmatrix}.$$

H es definida positiva, entonces M es copositiva plus por el corolario 1.10. Y por el teorema 1.7, el sistema el algoritmo terminará en un número finito de pasos. Vamos a usar el método de Lemke para resolver el problema:

Denotamos al vector de las variables de holgura como \mathbf{y} y el vector de los multiplicadores de Lagrange de las restricciones $A\mathbf{x} \leq \mathbf{b}$ y $\mathbf{x} \geq 0$ como \mathbf{u} y \mathbf{v} . Y sean:

$$M = \begin{pmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & -2 \\ 1 & -1 & 2 & -2 \\ 1 & 2 & -2 & 4 \end{pmatrix} \text{ y } \mathbf{q} = \begin{pmatrix} 2 \\ 2 \\ -2 \\ -6 \end{pmatrix}.$$

Introducimos la variable artificial z_0 y formamos la siguiente tabla:

	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	-
w_1	1	0	0	0	0	0	1	1	-1	2
w_2	0	1	0	0	0	0	-1	2	-1	2
w_3	0	0	1	0	-1	1	-2	2	-1	-2
w_4	0	0	0	1	-1	-2	2	-4	-1	-6

Calculamos $\min\{q_i : 1 \leq i \leq 4\} = q_4$, entonces pivotamos la fila 4 y la columna de z_0 .

	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	-
w_1	1	0	0	-1	1	2	-1	5	0	8
w_2	0	1	0	-1	1	2	-3	6	0	8
w_3	0	0	1	-1	0	3	-4	6	0	4
z_0	0	0	0	-1	1	2	-2	4	1	6

Por el test de cocientes, w_3 deja la base y, por tanto, z_3 entra. Pivotamos la fila de w_3 con la columna z_3 .

	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	-
w_1	1	0	$\frac{-5}{6}$	$\frac{-1}{6}$	1	$\frac{-1}{2}$	$\frac{7}{3}$	0	0	$\frac{14}{3}$
w_2	0	1	-1	0	1	-1	1	0	0	4
z_4	0	0	$\frac{1}{6}$	$\frac{-1}{6}$	0	$\frac{1}{2}$	$\frac{-2}{3}$	1	0	$\frac{2}{3}$
z_0	0	0	$\frac{-2}{3}$	$\frac{-1}{3}$	1	0	$\frac{2}{3}$	0	1	$\frac{10}{3}$

Por el test de cocientes, w_1 deja la base y, por tanto z_1 entra en ella. Pivotamos, entonces, la fila de w_1 con la columna z_3 .

	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	-
z_3	$\frac{3}{7}$	0	$\frac{-5}{14}$	$\frac{-1}{14}$	$\frac{3}{7}$	$\frac{-3}{14}$	1	0	0	2
w_2	$\frac{-3}{7}$	1	$\frac{-9}{14}$	$\frac{1}{14}$	$\frac{4}{7}$	$\frac{-11}{14}$	0	0	0	2
z_4	$\frac{2}{7}$	0	$\frac{-1}{14}$	$\frac{-3}{14}$	$\frac{2}{7}$	$\frac{5}{14}$	0	1	0	2
z_0	$\frac{-2}{7}$	0	$\frac{-3}{7}$	$\frac{-2}{7}$	$\frac{5}{7}$	$\frac{1}{7}$	0	0	1	2

Por el test de cocientes, z_0 deja las bases y entra z_1 . Pivotando la fila de z_0 con la columna de z_1 nos queda la última tabla:

	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	-
z_3	$\frac{3}{5}$	0	$\frac{-1}{10}$	$\frac{1}{10}$	0	$\frac{-3}{10}$	1	0	$\frac{-3}{5}$	$\frac{4}{5}$
w_2	$\frac{-1}{5}$	1	$\frac{-3}{10}$	$\frac{3}{10}$	0	$\frac{-9}{10}$	0	0	$\frac{-4}{5}$	$\frac{2}{5}$
z_4	$\frac{2}{5}$	0	$\frac{1}{10}$	$\frac{-1}{10}$	0	$\frac{3}{10}$	0	1	$\frac{-2}{5}$	$\frac{6}{5}$
z_1	$\frac{-2}{5}$	0	$\frac{-3}{5}$	$\frac{-2}{5}$	1	$\frac{1}{5}$	0	0	$\frac{7}{5}$	$\frac{14}{5}$

Entonces el algoritmo produce el punto $(w_1, w_2, w_3, w_4, z_1, z_2, z_3, z_4) = (0, \frac{2}{5}, 0, 0, \frac{14}{5}, 0, \frac{4}{5}, \frac{6}{5})$. *STOP*.
 1 Para obtener la solución a nuestro problema original basta con tomar la solución asociada las variables originales, que en nuestro caso son z_3 y z_4 . Por tanto, la solución a nuestro problema cuadrático es $(x_1, x_2) = (\frac{4}{5}, \frac{6}{5})$.

Calculamos también el valor de la función objetivo en este punto, $f(\frac{4}{5}, \frac{6}{5}) = -\frac{186}{25}$.

Ejemplo 2. Vamos a presentar otro problema cuadrático, en este caso vamos a obtener una terminación en rayo. Como ya hemos ilustrado el algoritmo en el ejemplo anterior, mostraremos tan solo la última tabla.

$$\begin{aligned} \text{mín} \quad & -2x_1 - 4x_2 + x_1^2 - 2x_1x_2 + x_2^2 \\ \text{su jeto a} \quad & -x_1 + x_2 \leq 1 \\ & x_1 - 2x_2 \leq 4 \\ & x_1, x_2 \geq 0. \end{aligned}$$

	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	-
w_1	1	0	$-\frac{3}{4}$	$-\frac{1}{4}$	$-\frac{1}{2}$	$\frac{1}{4}$	0	0	0	$\frac{7}{2}$
w_2	0	1	0	-1	1	-2	-1	0	0	8
z_4	0	0	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{2}$	$-\frac{3}{4}$	-1	1	0	$\frac{1}{2}$
z_0	0	0	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{1}{2}$	0	0	1	3

En este punto, z_3 debería entrar en las bases, pero todos los elementos asociados a su columna son no positivos. *STOP*. Hemos encontrado una terminación en rayo,

$$R = \{(\frac{7}{2}, 8, 0, 0, 0, 0, \frac{1}{2}, 3) + \lambda(0, 1, 0, 0, 0, 0, 1, 1, 0) : \lambda \geq 0\}.$$

Se ha obtenido tomando la solución asociada a la última tabla y como dirección del vector formado por un uno en la variable que debería haber entrado en la base, en este caso z_3 , el vector columna de esta misma variable cambiado de signo y ceros en el resto de las posiciones. Para obtener la solución a nuestro problema original basta con tomar la solución asociada las variables originales, que en nuestro caso son z_3 y z_4 . Entonces, la solución para nuestro problema original es:

$$R = \{(0, \frac{1}{2}) + \lambda(1, 1) : \lambda \geq 0\}.$$

Calculamos también el valor para la función objetivo que es $f(\lambda, \frac{1}{2} + \lambda) = -6\lambda - \frac{7}{4}$, que en efecto, tiende a $-\infty$ cuando $\lambda \rightarrow \infty$.

4.2. Programación separable.

Un problema de programación separable, como ya se ha expuesto en el capítulo 3, es aquel en el que su función objetivo y restricciones pueden ser expresadas como suma de funciones dependientes de una de las variables solamente cada una.

Este tipo de programación aparece principalmente en el campo de la ingeniería, por ejemplo en el planteamiento de sistemas de recursos hídricos; también encontraremos algún problema de programación separable aplicado al ámbito de la economía y las finanzas, donde es bastante frecuente querer encontrar máximos y mínimos en funciones, que en ocasiones, dependiendo del problema, podrán ser separables. De todas estas aplicaciones económicas podemos encontrar más información en [7].

Ejemplo 3. A continuación se muestra el algoritmo presentado en el segundo capítulo para aproximar la solución del siguiente problema separable:

$$\text{mín} \quad x_1^2 - 6x_1 + x_2^2 - 8x_2 - \frac{1}{2}x_3$$

$$\begin{aligned} \text{suje}to \ a \ x_1 + x_2 + x_3 &\leq 5 \\ x_1^2 - x_2 &\leq 3 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

En este ejemplo la variable x_3 es lineal en la función objetivo y las restricciones, luego $L = \{3\}$. Comenzamos el método de generación de los puntos de partición con la partición inicial $x_{11} = x_{12} = 0$. Definimos las variables de holgura x_4 y x_5 y creamos la primera tabla, en esta iteración x_3 entra a las bases y x_4 la deja:

Iteración 1.

	z	λ_{11}	λ_{12}	x_3	x_4	x_5	-
z	1	0	0	0.5	0	0	0
x_4	0	0	0	1	1	0	5
x_5	0	0	0	0	0	1	3
λ_{11}	0	1	0	0	0	0	1
λ_{12}	0	0	1	0	0	0	1

	z	λ_{11}	λ_{12}	x_3	x_4	x_5	-
z	1	0	0	0	-0.5	0	-2.5
x_3	0	0	0	1	1	0	5
x_5	0	0	0	0	0	1	3
λ_{11}	0	1	0	0	0	0	1
λ_{12}	0	0	1	0	0	0	1

Decidamos si un nuevo punto se incluye en la partición, resolvemos los siguientes problemas:

$$\text{mín } x_1^2 - 5.5x_1 \text{ sujeto a } 0 \leq x_1 \leq 5.$$

$$\text{mín } x_2^2 - 7.5x_2 \text{ sujeto a } 0 \leq x_2 \leq 5.$$

Las soluciones son $\tilde{x}_1 = 2.75$ y $\tilde{x}_2 = 3.75$ y sus valores para la función objetivo $\tilde{z}_1 = -7.56$ y $\tilde{z}_2 = -14.06$, respectivamente. Vamos a introducir un nuevo punto en la partición, x_{22} , ya que $\text{mín}\{\tilde{z}_1 + \hat{v}_1, \tilde{z}_2 + \hat{v}_2\} = \tilde{z}_2 + \hat{v}_2 = -14.06$.

Iteración 2.

Observamos que la variable λ_{22} acaba de ser incluida en la tabla con los valores $(3.75, -3.75, 0, 1)^t$ como columna debido a las funciones g_{i2} y λ 's, esta columna se actualiza multiplicando por $B^{-1} = I$. Para obtener la segunda tabla, pivotamos la columna de la nueva variable con λ_{12} .

	z	λ_{11}	λ_{12}	λ_{22}	x_3	x_4	x_5	-
z	1	0	0	14.06	0	-0.5	0	-2.5
x_3	0	0	0	3.75	1	1	0	5
x_5	0	0	0	-3.75	0	0	1	3
λ_{11}	0	1	0	0	0	0	0	1
λ_{12}	0	0	1	1	0	0	0	1

	z	λ_{11}	λ_{12}	λ_{22}	x_3	x_4	x_5	-
z	1	0	-14.06	0	0	-0.5	0	-16.56
x_3	0	0	-3.75	0	1	1	0	1.25
x_5	0	0	3.75	0	0	0	1	6.75
λ_{11}	0	1	0	0	0	0	0	1
λ_{22}	0	0	1	1	0	0	0	1

Procedemos de manera análoga y vamos a introducir un nuevo punto en la partición, x_{21} , ya que $\min\{\tilde{z}_1 + \hat{v}_1, \tilde{z}_2 + \hat{v}_2\} = \tilde{z}_1 + \hat{v}_1 = -7.56$.

Iteración 3.

Procedemos análogamente a la iteración 2, ya que en este caso se introduce la columna asociada a λ_{21} , actualizamos la columna multiplicando por B^{-1} . Para la siguiente tabla, λ_{21} pivota con la fila de x_3 .

	z	λ_{11}	λ_{21}	λ_{12}	λ_{22}	x_3	x_4	x_5	-
z	1	0	7.56	-14.06	0	0	-0.5	0	-16.56
x_3	0	0	2.75	-3.75	0	1	1	0	1.25
x_5	0	0	7.56	3.75	0	0	0	1	6.75
λ_{11}	0	1	1	0	0	0	0	0	1
λ_{22}	0	0	0	1	1	0	0	0	1

	z	λ_{11}	λ_{21}	λ_{12}	λ_{22}	x_3	x_4	x_5	-
z	1	0	0	-3.78	0	-2.72	-3.22	0	-19.96
λ_{21}	0	0	1	-1.36	0	0.36	0.36	0	0.45
x_5	0	0	0	14.03	0	-2.72	-2.72	1	3.35
λ_{11}	0	1	0	1.36	0	-0.36	-0.36	0	0.55
λ_{22}	0	0	0	1	1	0	0	0	1

Por tanto, deducimos de las expresiones de las funciones aproximadas y de esta última tabla que $\hat{\mathbf{x}} = (1.25, 3.75, 0)^t$. Podríamos seguir añadiendo puntos en la partición y mejorando la aproximación, pero en nuestro caso el valor de la función objetivo original está entre -23.74 y -21.88; y para nuestro problema aproximado hemos encontrado un valor para la función objetivo de -21.88 y lo consideramos suficiente. STOP.

4.3. Programación lineal fraccional.

En 1937, Von Neumann comenzó el estudio de los problemas de programación fraccional, aunque en ese momento no se denominara como en la actualidad, por su aparición en una aplicación económica. El interés de estos problemas es obvio, ya que tanto en economía como en ingeniería o logística aparecen constantemente funciones cociente de magnitudes que nos gustaría optimizar como coste/tiempo o coste/unidad entre otras. Este problema también aparecerá en campos de la informática o de la matemática como el análisis numérico o la programación estocástica, una recopilación de estas aplicaciones aparecen en [5] y en [8].

Ejemplo 4. A continuación planteamos el siguiente problema de programación fraccional:

$$\begin{aligned} \min \quad & \frac{-2x_1 + x_2 + 2}{x_1 + 3x_2 + 4} \\ \text{sujo a} \quad & -x_1 + x_2 \leq 4 \\ & x_2 \leq 6 \\ & 2x_1 + x_2 \leq 14 \\ & x_1, x_2 \geq 0, \end{aligned}$$

- Resolvamos el ejemplo por el método de Gilmore y Gomory.

Introduciendo las variables de holgura x_3 , x_4 y x_5 obtenemos el punto extremo inicial $\mathbf{x}_1 = (0, 0, 4, 6, 14)^t$. El gradiente de la función objetivo es $\nabla f(\mathbf{x}) = \left(\frac{-7x_2 - 10}{(x_1 + 3x_2 + 4)^2}, \frac{7x_1 - 2}{(x_1 + 3x_2 + 4)^2}, 0, 0, 0 \right)$, en esta fórmula sustituiremos el punto extremo de cada iteración para calcular la primera fila.

Creamos la primera tabla,
Iteración 1

	x_1	x_2	x_3	x_4	x_5	-
$\nabla f(\mathbf{x}_1)$	$-\frac{10}{16}$	$-\frac{2}{16}$	0	0	0	-
x_3	-1	1	1	0	0	4
x_4	0	1	0	1	0	6
x_5	2	1	0	0	1	14
\mathbf{r}	$-\frac{10}{16}$	$-\frac{2}{16}$	0	0	0	-

Calculamos $\max\{-r_i : i = 1, \dots, 5\} = \frac{10}{16}$. Por tanto, x_1 entra a la base y por el test del cociente x_5 abandona las bases.

Iteración 2

	x_1	x_2	x_3	x_4	x_5	-
$\nabla f(\mathbf{x}_2)$	$-\frac{10}{121}$	$\frac{47}{121}$	0	0	0	-
x_3	0	$\frac{3}{2}$	1	0	$\frac{1}{2}$	11
x_4	0	1	0	1	0	6
x_1	1	$\frac{1}{2}$	0	0	$\frac{1}{2}$	7
\mathbf{r}	0	$\frac{52}{121}$	0	0	$\frac{5}{121}$	-

Obtenemos $\mathbf{x}_2 = (7, 0, 11, 6, 0)$ y como todos los valores de la fila \mathbf{r} son positivos, *STOP*.

Y $\mathbf{x}_2 = (7, 0, 11, 6, 0)$ es la solución óptima. Por tanto, la solución de nuestro problema original es $\mathbf{x}_2 = (7, 0)$ y calculamos también el valor de la función objetivo que es $f(7, 0) = -2$.

- Ahora, realicemos el cambio necesario para reformular este problema de programación fraccional como un PPL.

Definimos $z = \frac{1}{x_1 + 3x_2 + 4}$ e $y = z\mathbf{x}$ y obtenemos:

$$\begin{aligned} & \text{mín } -2y_1 + y_2 + 2z \\ & \text{suje } a \quad -y_1 + y_2 - 4z \leq 0 \\ & \quad \quad \quad y_2 - 6z \leq 0 \\ & \quad \quad \quad 2y_1 + y_2 - 14z \leq 0 \\ & \quad \quad \quad y_1 + 3y_2 + 4z = 1 \\ & \quad \quad \quad y_1, y_2, z \geq 0. \end{aligned}$$

Ahora, este problema puede ser resuelto mediante el algoritmo Simplex. La solución que se obtiene al realizar el método es $(\frac{7}{11}, 0, \frac{1}{11})$ y, por tanto, deshaciendo el cambio la solución al problema original es $(7, 0)$. Calculamos también el valor de la función objetivo que es $f(7, 0) = -2$.

Bibliografía

- [1] MAKHTAR, S. BAZARAA, M.S., *Nonlinear Programming. Theory and Algorithms*, C.M. Shetty, 1979.
- [2] WAYNE L. WINSTON, *Investigación de operaciones. Aplicaciones y algoritmos*, Grupo Editorial Iberoamérica.
- [3] HILLIER Y LIEBERMAN, *Introducción a la investigación de operaciones*, McGraw-Hill
- [4] MERINO MAESTRE, MARÍA, *Técnicas clásicas de optimización. Parte I: Programación lineal y no lineal.*, Apuntes del máster en modelización e investigación matemática.
- [5] RUÍZ MEZA, J.L., VERGARA RODRÍGUEZ, C.J. Y OTROS, *Enfoques, Teorías y Perspectivas de la Ingeniería Industrial y sus Programas académicos.*, ed. digital, Cekar, colección investigación.
- [6] MORALES ALMORA, J. L., *Programación cuadrática: Teoría, algoritmos y aplicaciones.*, 1990.
- [7] MARTÍNEZ ANGELES, L. A. , *Modelo de Programación Cuadrática y Ratios Financieros para minimizar el riesgo de las inversiones en la Bolsa de Valores de Lima.*, 2013.
- [8] ZIANI, A., *Optimización de funciones fraccionarias.*, Trabajo fin de grado, Universidad de Murcia, facultad de matemáticas, 2017.
- [9] DECASTRO ICOGO, E., *On Wolfe's algorithm for Quadratic Programming.*, Institute of Mathematical Sciences and Physics, University of the Philippines at Los Banos.
- [10] HUNGERLÄNDER, P., *Algorithms for Convex Quadratic Programming.*, Institute of Mathematical Sciences and Physics, University of the Philippines at Los Banos.

Anexos

Anexo A

Anexo

A.1. El problema de programación lineal.

El algoritmo Simplex desarrollado por George Dantzing es el procedimiento básico para resolver el problema de programación lineal, pero además es el punto de partida para la creación de muchos otros algoritmos. A lo largo de este capítulo definiremos este problema y expondremos el método Simplex, algoritmo que será la base para la construcción de la mayoría de los métodos que definiremos durante todo el documento.

Un problema de programación lineal (PPL) es un problema de optimización para el que cumple:

1. La función objetivo que hay que optimizar, maximizar o minimizar, es una función lineal de las variables de decisión.
2. Cada restricción puede escribirse como una ecuación o inecuación lineal.

Matemáticamente, el PPL puede formularse como se muestra a continuación.

$$\begin{aligned} \text{máx(mín)} \quad z &= c_1x_1 + c_2x_2 + \dots + c_nx_n \\ \text{sujeto a} \quad a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \\ x_i &\geq 0 \quad \forall i = 1, \dots, n. \end{aligned} \tag{A.1}$$

Equivalentemente, presentamos el PPL en su forma matricial:

$$\begin{aligned} \text{mín} \quad \mathbf{c}'\mathbf{x} \\ \text{sujeto a} \quad \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}, \end{aligned}$$

donde \mathbf{c} es un vector n dimensional, A es una matriz $m \times n$ de rango m y \mathbf{b} es un vector m dimensional. En realidad, un PPL admite muchos formatos diferentes, el presentado en (A.1) se denomina PPL en forma estándar en el que todas las restricciones son de igualdad y las variables de decisión son no negativas.

Los siguientes tres teoremas son necesarios para el desarrollo del algoritmo Simplex:

Teorema A.1. Caracterización de punto extremo. Minkowski. Sea $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, donde A es una matriz $m \times n$ de rango m , y \mathbf{b} es un vector de dimensión m . Un punto $\tilde{\mathbf{x}}$ es un punto extremo de S si y sólo si A puede descomponerse como $[\mathbf{B}, \mathbf{N}]$, donde B es una matriz $m \times m$ tal que $\mathbf{B}^{-1}\mathbf{b} \geq \mathbf{0}$ y $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_B, \tilde{\mathbf{x}}_N] = [\mathbf{B}^{-1}\mathbf{b}, \mathbf{0}]$

Teorema A.2. Caracterización de dirección extrema. Sea $S = \{\mathbf{x} : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, donde \mathbf{A} es una matriz $m \times n$ de rango m , y \mathbf{b} es un vector de dimensión m . Un vector no nulo $\tilde{\mathbf{d}}$ es una dirección extrema de S si y solo si \mathbf{A} se puede descomponer en $[\mathbf{B}, \mathbf{N}]$ de tal forma que $\mathbf{B}^{-1}\mathbf{a}_j \leq \mathbf{0}$ para alguna columna \mathbf{a}_j de \mathbf{N} , y $\tilde{\mathbf{d}}$ es un múltiplo positivo de $\mathbf{d} = \begin{pmatrix} \mathbf{B}^{-1}\mathbf{a}_j \\ \mathbf{e}_j \end{pmatrix}$, donde \mathbf{e}_j es un vector $n - m$ -dimensional cuyos elementos son todo ceros, excepto un 1 en la posición j .

Teorema A.3. Caracterización de solución óptima. Consideramos el PPL en su forma matricial. Suponemos que la región factible es no vacía y sean $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ puntos extremos y $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_l$ direcciones extremas de la región factible. Una condición necesaria y suficiente para obtener una solución finita es que $\mathbf{c}^t \mathbf{d}_j \geq 0$ para $j = 1, \dots, l$. En este caso, existe un punto extremo \mathbf{x}_i que es la solución óptima al problema.

A.2. El algoritmo Simplex.

El algoritmo Simplex es un procedimiento para resolver el PPL. Se basa en moverse de un punto extremo a otro del problema mejorando cada vez la función objetivo. El proceso acaba cuando se alcanza el punto extremo óptimo o se encuentra una dirección extrema de no acotación.

Paso inicial: Considerar el PPL expresado en su forma matricial. Encontrar un punto extremo inicial \mathbf{x} con su base B .

Paso principal:

1. Sea \mathbf{x} un punto extremo con su base B . Calcular $\mathbf{c}_B^t B^{-1} \mathbf{a}_j - \mathbf{c}_j \forall j$. Si existe j_0 tal que $\mathbf{y}_{j_0} = B^{-1} \mathbf{a}_{j_0} \leq \mathbf{0}$, STOP. La función objetivo es no acotada a lo largo de la semirrecta

$$\{\mathbf{x} + \lambda \begin{pmatrix} -\mathbf{y}_{j_0} \\ \mathbf{e}_{j_0} \end{pmatrix} : \lambda \geq 0\},$$

donde \mathbf{e}_{j_0} es el vector de ceros excepto en la componente j_0 . Si \mathbf{y}_{j_0} tiene alguna componente positiva, ir al paso 2.

2. Calcular r de la siguiente manera

$$\lambda = \min_{1 \leq i \leq m} \left\{ \frac{\tilde{b}_i}{y_{ij_0}} \right\} = \frac{\tilde{b}_r}{y_{rj_0}},$$

donde y_{ij_0} es la i -ésima componente de \mathbf{y}_{j_0} . En este caso, las componentes de $\mathbf{x} = \tilde{\mathbf{x}} + \lambda \mathbf{d}_{j_0}$ vienen dadas por

$$x_{B_i} = \tilde{b}_i - \frac{\tilde{b}_r}{y_{rj_0}} \text{ para } i = 1, \dots, m,$$

$$x_{j_0} = \frac{\tilde{b}_r}{y_{rj_0}},$$

y el resto de valores de x_j son cero. Con esto, formamos un nuevo punto extremo \mathbf{x} . Formar las nuevas bases eliminando la columna \mathbf{a}_B de B e introduciendo \mathbf{a}_{j_0} en su lugar. Volver al paso 1.

A.3. Resultados útiles para la programación cuadrática y fraccional.

A continuación, presentamos los teoremas necesarios para la justificación de los resultados que presentamos a lo largo del documento relacionados con los capítulos primero y cuarto.

Teorema A.4. Teorema de Farkas. Sea A un matriz $m \times n$ y sea \mathbf{c} un vector n dimensional. Entonces solamente uno de los siguientes dos sistemas tiene solución:

$$\text{Sistema 1 : } \mathbf{Ax} \leq 0 \text{ y } \mathbf{c}'\mathbf{x} > 0 \text{ para algún } \mathbf{x} \in E_n$$

$$\text{Sistema 2 : } \mathbf{A}'\mathbf{y} = \mathbf{c} \text{ e } \mathbf{y} \geq 0 \text{ para algún } \mathbf{y} \in E_m.$$

Teorema A.5. Sea S un conjunto abierto y convexo en E_n , y sea $f : S \rightarrow E_1$ diferenciable sobre S . Entonces, f es convexa si y solo si para algún $\tilde{\mathbf{x}} \in S$, se tiene

$$f(\mathbf{x}) \geq f(\tilde{\mathbf{x}}) + \nabla f(\tilde{\mathbf{x}})'(\mathbf{x} - \tilde{\mathbf{x}}) \text{ para cada } \tilde{\mathbf{x}} \in S$$

Análogamente, f es estrictamente convexa si y solo si para cada $\tilde{\mathbf{x}} \in S$, se tiene

$$f(\mathbf{x}) > f(\tilde{\mathbf{x}}) + \nabla f(\tilde{\mathbf{x}})'(\mathbf{x} - \tilde{\mathbf{x}}) \text{ para cada } \mathbf{x} \neq \tilde{\mathbf{x}} \in S.$$

Teorema A.6. Sea S un conjunto no vacío compacto poliédrico en E_n , y sea $f : E_n \rightarrow E_1$ cuasiconvexa y continua sobre S . Consideramos el problema de

$$\begin{aligned} & \text{máx } f(\mathbf{x}) \\ & \text{sujeta a } \mathbf{x} \in S. \end{aligned}$$

Entonces existe $\tilde{\mathbf{x}}$ solución óptima del problema, donde $\tilde{\mathbf{x}}$ es un punto extremo de S .

Teorema A.7. Sea S un conjunto no vacío convexo en E_n , y sea $f : E_n \rightarrow E_1$ estrictamente cuasiconvexa. Consideramos el problema de

$$\begin{aligned} & \text{mín } f(\mathbf{x}) \\ & \text{sujeta a } \mathbf{x} \in S. \end{aligned}$$

Si $\tilde{\mathbf{x}}$ solución local óptima del problema, entonces también es una solución global.

Teorema A.8. Sea S un conjunto no vacío convexo en E_n , y sea $f : E_n \rightarrow E_1$ diferenciable y pseudoconvexa sobre S . Entonces f es estrictamente cuasiconvexa y cuasicóncava.

Teorema A.9. Condiciones de Karush-Kuhn-Tucker. Sea X un subconjunto no vacío de \mathbb{R}^n , y sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ para $i = 1, \dots, m$, y $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ para $i = 1, \dots, l$. Consideramos el siguiente problema:

$$\begin{aligned} & \text{mín } f(\mathbf{x}) \\ & \text{suje to a } g_i(\mathbf{x}) \leq 0 \text{ para } i = 1, \dots, m \\ & \quad h_i(\mathbf{x}) = 0 \text{ para } i = 1, \dots, l \\ & \quad \mathbf{x} \in X \end{aligned}$$

Sea $\tilde{\mathbf{x}}$ una solución factible, y sea $I = \{i : g_i(\tilde{\mathbf{x}}) = 0\}$. Suponemos que f y g_i para $i \in I$ son diferenciables en $\tilde{\mathbf{x}}$, y que cada g_i para $i \notin I$ es continua en $\tilde{\mathbf{x}}$ y que cada h_i para $i = 1, \dots, l$ es continua y diferenciable en $\tilde{\mathbf{x}}$. Además, suponemos que $\nabla g_i(\tilde{\mathbf{x}})$ para $i \in I$ y $\nabla h_i(\tilde{\mathbf{x}})$ para $i = 1, \dots, l$ son linealmente independientes. Si $\tilde{\mathbf{x}}$ es la solución local al problema, entonces existen unos únicos escalares u_i para cada $i \in I$ y v_i para $i = 1, \dots, l$ tal que:

$$\nabla f(\tilde{\mathbf{x}}) + \sum_{i \in I} u_i g_i(\tilde{\mathbf{x}}) + \sum_{i=1}^l v_i \nabla h_i(\tilde{\mathbf{x}}) = \mathbf{0}, \quad u_i \geq 0 \quad \forall i \in I$$

Además, si cada g_i para $i \notin I$ es también diferenciable en $\tilde{\mathbf{x}}$, las condiciones de Kuhn-Tucker (KKT) se pueden escribir como:

$$\begin{aligned} \nabla f(\tilde{\mathbf{x}}) + \sum_{i=1}^m u_i g_i(\tilde{\mathbf{x}}) + \sum_{i=1}^l v_i \nabla h_i(\tilde{\mathbf{x}}) &= \mathbf{0} \\ u_i g_i(\tilde{\mathbf{x}}) &= 0 \quad \text{para } i = 1, \dots, m \\ u_i &\geq 0 \quad \text{para } i = 1, \dots, m. \end{aligned}$$

Teorema A.10. Sea X un conjunto no vacío en E_n , y sea $f : E_n \rightarrow E_1$, $g : E_n \rightarrow E_m$ y $h : E_n \rightarrow E_l$. Supongamos que existe $\tilde{\mathbf{x}} \in X$ y $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ con $\tilde{\mathbf{u}} \geq \mathbf{0}$, tal que

$$\phi(\tilde{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leq \phi(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}}) \leq \phi(\mathbf{x}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}}), \quad (\text{A.2})$$

para todo $\mathbf{x} \in X$ y todo (\mathbf{u}, \mathbf{v}) con $\mathbf{u} \geq \mathbf{0}$, donde $\phi(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \mathbf{u}^t g(\mathbf{x}) + \mathbf{v}^t h(\mathbf{x})$. Entonces, $\tilde{\mathbf{x}}$ y $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ son la solución al problema primal y al dual respectivamente. Recíprocamente, suponemos que X , f y g son convexas y que h es una función lineal afín. Además, suponemos que $\mathbf{0} \in \text{inth}(X)$ y que existe $\hat{\mathbf{x}} \in X$ con $g(\hat{\mathbf{x}}) < \mathbf{0}$ y $h(\hat{\mathbf{x}}) = \mathbf{0}$. Si $\tilde{\mathbf{x}}$ es una solución óptima para el problema primal, entonces existe $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ con $\tilde{\mathbf{u}} \geq \mathbf{0}$ tal que se cumple (A.2).

A.4. Resultados útiles para la programación no lineal separable.

Como hemos visto en la sección 2.2, bajo hipótesis de convexidad sobre las funciones, podemos encontrar soluciones factibles para el problema PLA. En esta sección, vamos a ver enunciar resultados que nos permitan saber si los puntos de la partición elegida están seleccionados lo suficientemente cerca entre ellos, entonces las soluciones a ambos problemas, el original y el aproximado, también lo están, es decir, las soluciones a ambos problemas pueden ser consideradas las mismas, o si, por el contrario, debemos añadir algún punto mas a la partición para mejorar la calidad de la solución. Estos teoremas son la base para la construcción del algoritmo presentado al final del capítulo 2.

Teorema A.11. Consideramos los PNLs y PA definidos en (2.1) y (2.4) respectivamente. Para $j \notin L$ suponemos que f_j y g_{ij} para $i = 1, \dots, m$ son convexas y además, sean \hat{f}_j y \hat{g}_{ij} sus aproximaciones lineales a trozos sobre el intervalo $[a_j, b_j]$. Para $j \notin L$ y para $i = 1, \dots, m$, sea c_{ij} tal que se satisface que $|g'_{ij}(x_j)| \leq c_{ij}$ para $x_j \in [a_j, b_j]$. Además, para $j \notin L$, sea c_j tal que $|f'_j(x_j)| \leq c_j$ para $x_j \in [a_j, b_j]$. Para $j \notin L$, sea δ_j la longitud máxima usada para la variable x_j . Entonces,

$$\begin{aligned} \hat{f}(\mathbf{x}) &\geq f(\mathbf{x}) \geq \hat{f}(\mathbf{x}) - c \\ \hat{g}_i(\mathbf{x}) &\geq g_i(\mathbf{x}) \geq \hat{g}_i(\mathbf{x}) - c \quad \forall i = 1, \dots, m, \end{aligned}$$

donde $c = \max_{0 \leq i \leq m} \{\tilde{c}_i\}$ y

$$\tilde{c}_0 = \sum_{j \notin L} 2c_j \delta_j \quad \text{y} \quad \tilde{c}_i = \sum_{j \notin L} 2c_{ij} \delta_j \quad \text{para } i = 1, \dots, m.$$

Teorema A.12. Consideramos el PNLs definido en (2.1). Sea $L = \{j : f_j \text{ y } g_{ij} \text{ lineales } i = 1, \dots, m.\}$. Para $j \notin L$, sean \hat{f}_j y \hat{g}_{ij} funciones a trozos que aproximan a f_j y g_{ij} respectivamente para $i=1, \dots, m$. Consideramos también los PA y PLA definidos en (2.4) y (2.5), problemas equivalentes al PNLs. Para $j \notin L$ suponemos que f_j y g_{ij} para $i = 1, \dots, m$ son convexas. Sea $\tilde{\mathbf{x}}$ una solución óptima al PNLs. Sea \hat{x}_j para $j \in L$ y $\hat{\lambda}_{vj}$ para $v = 1, \dots, k_j$ y $j \notin L$ una solución óptima al PLA tal que el vector $\hat{\mathbf{x}}$, cuyas componentes son \hat{x}_j para $j \in L$ y $\hat{x}_j = \sum_{v=1}^{k_j} \hat{\lambda}_{vj} x_{vj}$ para $j \notin L$, es una solución óptima para el problema definido en (3.4). Sea \hat{u}_i sean los multiplicadores de Lagrange óptimos asociados a las restricciones $\hat{g}_i(\mathbf{x}) \leq 0$ para $i=1, \dots, m$. Entonces:

1. $\hat{\mathbf{x}}$ es una solución factible al PNLs.
2. $0 \leq f(\hat{\mathbf{x}}) - f(\tilde{\mathbf{x}}) \leq c(1 + \sum_{i=1}^m \hat{u}_i)$, donde c se define como en el teorema 2.2.

