UNIVERSIDAD DE MURCIA

TRABAJO FIN DE GRADO
FACULTAD DE MATEMÁTICAS

Optimización de funciones fraccionarias

Autor:
Ziani ADIL
adil.ziani@um.es

supervisor:
Dr. Pelegrín Pelegrín BLAS
pelegrin@um.es

06/02/2017

Declaración de Originalidad

Ziani Adil, autor del TFG, «Optimización de funciones fraccionarias», bajo la tutela del Dr. Pelegrín Pelegrín Blas, declara que el trabajo que presenta es original, en el sentido de que ha puesto el mayor empeño en citar debidamente todas las fuentes utilizadas.

En Murcia, a 15 de Septiembre de 2017.

Firma:

Nota: En la Secretaría de la Facultad de Matemáticas se ha presentado una copia firmada de esta declaración.

UNIVERSIDAD DE MURCIA

Resumen

Optimización de funciones fraccionarias

por Ziani Adil

Desde el desarrollo de la programación lineal por Leonid Kantorovich en 1939 y el desarrollo del método Símplex por George Dantzig en 1947, esta herramienta de las matemáticas encontró diversas aplicaciones prácticas en la actividad del hombre y desde entonces diversas investigaciones se fueron desarrollando dentro de la Optimización.
La modelización de problemas donde se busca una solución que optimice un objetivo cumpliendo unas restricciones fue necesitando de expresiones más complejas que dejaron de ser lineales, una de ellas es la optimización fraccionaria, se trata de optimizar una función objetivo no lineal como cociente de dos funciones, mas en concreto el modelo general del problema y el que trataremos a lo largo del documento sería:

$$\text{(PF)} \qquad \text{mín} \left\{ \frac{f(x)}{h(x)} : g(x) \leq 0, x \in X \right\} \tag{0.1}$$

donde $X \subset \mathbb{R}^n$ un conjunto no vacio, $f, h : X \to \mathbb{R}$ funciones reales de variable real, $g : X \to \mathbb{R}^m$ función vectorial de variable real.
Cuando las funciones f, h, g son lineales, se habla de Optimización Fraccionaria Lineal.
En la práctica aparecen diversas situaciones donde la función objetivo es dada como cociente de dos funciones, lo que hizo necesario un estudio del problema transformándose en un área de investigación propia desde que Von Neumann, en 1937, tratase el problema fraccional (aunque entonces no fue considerado como tal) en un modelo económico donde se maximiza una tasa de crecimiento dada como cociente de dos funciones, dicho trabajo titulado *über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes* puede encontrarse en el libro Menger (1998, pág 453-463).

El interés de este tema reside en el hecho de que muchos problemas de Optimización en la ingeniería y economía principalmente consideran la optimización de una relación entre funciones, por ejemplo: coste/tiempo, coste/unidad, coste/beneficio u otras cantidades que miden la eficiencia de un sistema.
También tiene un interés fuera de estos dos campos por su frecuente aparición en la teoría de la información, el análisis numérico, la programación estocástica, etc.

El objetivo de este trabajo es doble, por una parte mostrar los principales contenidos teóricos de la Optimización Fraccionaria junto con métodos de resolución y ejemplos ilustrativos, y por otra parte, usando las herramientas informáticas, dar una posible implementación de algoritmos más generales posible para la Optimización Fraccionaria.

En el capítulo 1, en una primera parte, se dan resultados generales de la Optimización no Lineal que giran en torno a la convexidad y diferenciabilidad, y se dan los resultados que se alcanzan disponiendo de éstas propiedades.
En una segunda parte, se adaptan los resultados anteriores al caso fraccional, con el objetivo de identificar con mayor facilidad si la función objetivo f/h cumple alguna propiedad de convexidad. En la ultima sección se da la transformación de Manas-Schaible, con un especial interés ya que transforma un problema de Optimización Fraccionaria en otro equivalente no fraccional, además, si $f, -h, g$ son convexas, el problema es convexo.
En el capítulo 2, se dan condiciones de optimalidad. También se muestra la relación con la Optimización Paramétrica que se explota posteriormente en el siguiente capítulo. Se estudia también el caso en el que la región factible es un poliedro y finalmente se introduce el concepto de Dualidad en la Optimización Fraccionaria.
En el capítulo 3, usando la teoría de los capítulos previos, se dan métodos de resolución: métodos paramétricos como el algoritmo de Dinkilbach, métodos simpleciales, métodos duales y se termina el capítulo dedicando una sección a la Optimización Fraccionaria Lineal.
El capítulo 4, se dedica a la implementación del algoritmo de Dinkilbach y el método de la Secante para resolver el problema paramétrico, usando el lenguaje Python, y se realizan pruebas computacionales sobre un conjunto de problemas seleccionados.

UNIVERSIDAD DE MURCIA

Résumé

Optimisation de fonctions fractionnaires

par Ziani ADIL

Depuis le développement de la programmation linéaire par Leonid Kantorovich en 1939 et le développement de la méthode Símplex par George Dantzig en 1947, cet outil des mathématiques a trouvé des diverses applications pratiques dans l'activité de l'homme et depuis lors de diverses recherches se sont développées à l'intérieur de l'Optimisation.
La modélisation de problèmes où est cherchée une solution qui optimise un objectif en accomplissant quelques restrictions avait besoin d'expressions plus complexes qu'elles ont cessé d'être linéaires, l'une d'elles est l'optimisation fractionnaire. Il s'agit d'optimiser une fonction objectif non linéaire comme quotient de deux fonctions, plus en somme le modèle général du problème et celui que nous traiterons dans l'ensemble du document serait:

$$\text{(PF)} \qquad \text{mín} \left\{ \frac{f(x)}{h(x)} : g(x) \leq 0, x \in X \right\} \tag{0.2}$$

lorsque $X \subset \mathbb{R}^n$ un ensemble ne pas vide, $f, h : X \to \mathbb{R}$ des fonctions réelles de variable réel, $g : X \to \mathbb{R}^m$ fonction vectorielle de variable réel.
Quand les fonctions f, h, g sont linéaires, on parle d'une Optimisation Fractionnaire Linéaire.
Dans la pratique apparaissent des diverses situations où la fonction objective est donné comme quotient de deux fonctions rendant nécessaire une étude du problème en se transformant dans un propre domaine de recherche depuis que, en 1937, Von Neumann traitait le problème fractionnel (même si à ce moment-là n'était pas considéré comme tel) dans un modèle économique où un taux de croissance est maximisé donné comme quotient de deux fonctions, dudit travail intitulé über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes, on peut le trouver dans le livre Menger (1998, pág 453-463).

L'intérêt de ce sujet réside dans le fait de que des nombreux problèmes d'optimisation dans l'ingénierie et l'économie considèrent principalement l'optimisation d'une relation entre des fonctions, par exemple : coût / temps, coût / unité, coût / bénéfice ou d'autres quantités qui mesurent l'efficacité d'un système.
Il a aussi un intérêt outre ces deux champs par sa fréquente apparition dans la théorie de l'information, de l'analyse numérique, la programmation stochastique, etc.
L'objectif de ce travail est double, d'un côté, montrer les contenus principaux théoriques de l'Optimisation Fractionnaire avec des méthodes de résolution et d'exemples illustrants, et par ailleurs, en utilisant les outils informatiques, donner une possible implémentation d'algorithmes les plus généraux possibles pour l'Optimisation Fractionnaire.

Dans le chapitre 1, dans une première partie on donne les résultats généraux de l'Optimisation non linéaire qui tournent autour de la convexité et la différentiabilité, et on donne les résultats qu'on atteint en disposant de ces propriétés.
Dans la deuxième partie, les résultats antérieurs s'adaptent au cas fractionnel, avec l'objectif d'identifier avec plus de facilité si la fonction objective f/h satisfait une certaine propriété de convexité.
Dans la dernière section on donne la transformation de Manas-Schaible, avec un intérêt spécial puisqu'elle transforme un problème d'Optimisation Fractionnaire dans un autre équivalent non fractionnel, en plus, si f,-h, g sont convexes, le problème est convexe.
Dans le chapitre 2, on donne des conditions d'optimalité. On montre aussi la relation avec l'Optimisation Paramétrique et qui est exploité par la suite dans le chapitre suivant. On étudie aussi le cas dans lequel la région faisable est un polyèdre et finalement le concept de Dualité s'introduit à l'Optimisation Fractionnaire.
Dans le chapitre 3, en utilisant la théorie des chapitres préalables, on donne des méthodes de résolution : des méthodes paramétrique comme l'algorithme de Dinkilbach, des méthodes simples, des méthodes duales et le chapitre se termine en dédiant une section à l'Optimisation Fractionnaire Linéaire.
Le chapitre 4 s'occupe de l'implémentation de l'algorithme de Dinkilbach et la méthode de la Sécante pour résoudre le problème paramétrique, en utilisant le langage Python, et des preuves d'informatique sont réalisées sur l'ensemble de problèmes sélectionnées.

Índice general

Capítulo 1

Preliminares

Dentro de la Optimización, la Optimización Convexa es la que dispone de resultados teóricos eficientes. Desde la segunda mitad del siglo XX, el tema de generalizaciones de convexidad y sus aplicaciones a la Optimización fue atrayendo cada vez más investigadores, naciendo así el Grupo de Trabajo Internacional de Convexidad (WGGC),y es por eso que se muestra aquí conceptos de convexidad para la Optimización no lineal y que se particularizan al caso fraccionario, así como conceptos de diferenciabilidad. En este capítulo también se mostrarán ejemplos y aplicaciones de la Optimización Fraccionaria.

Estructura del capítulo
Sección 1.1. Se muestran algunos ejemplos y aplicaciones de la Optimización Fraccionaria.
Sección 1.2. Se dan las definiciones de conjunto convexo, función convexa y sus generalizaciones, junto con propiedades en la Optimización.
Sección 1.3. Se dan las definiciones de función diferenciable y dos veces diferenciable junto con caracterizaciones. También se dan las generalizaciones de diferenciabilidad junto con propiedades en la Optimización.
Sección 1.4. Se particularizan las dos secciones anteriores para el caso f/h, analizando las condiciones sobre f y h para obtener tipos de convexidad y aplicar resultados, en esta línea el resultado clave es el Teorema 1.1, que permite estudiar la convexidad y sus generalizaciones de la función f/h en función del tipo de convexidad de f y h por separado.
Sección 1.5. Se muestra la transformación de Manas-Schaible, que bajo ciertas condiciones de convexidad transforma el problema 0.1 en otro equivalente convexo. Dicha transformación tiene una útil aplicación teórica y en el caso lineal, nos da un método de resolución estudiado en Sección 3.4.

1.1. Aplicaciones de la Optimización Fraccionaria

En esta sección veremos algunos problemas clásicos de Optimización fraccionaria para que sirvan de motivación. Pueden hallarse más en Minasian (1997) y Frenk y Schaible (2004).

1.1.1. Teoría de la información

Meister y Oettli (1967) y Aggarwal y Sharma (1970) aplicaron la optimización fraccionaria para calcular la velocidad de transmisión máxima de un canal de comunicación, es decir, lo que se llama capacidad del canal en teoría de la información. Consideramos un canal discreto caracterizado por un alfabeto de entrada, de m símbolos $A = \{a_1, a_2, \cdots, a_m\}$ y un alfabeto de salida, de n símbolos $B = \{b_1, b_2, \cdots, b_n\}$. Sea $P = (p_{ij})$ la matriz del canal, esto es:

$$p_{ij} = \text{Pr(recibir } b_i \text{ habiendo enviada } a_j) \quad i = 1 \cdots n, \quad j = 1 \cdots m.$$
$$p_{ij} \geq 0 \qquad \sum_j p_{ij} = 1 \quad \forall i$$

Supongamos que la matriz P no tiene filas nulas, esto es, que no hay símbolos del alfabeto de salida con probabilidad nula de recibirse en la transmisión. Sea $t_j > 0$ el coste (tiempo, energía,...) asociado a la transmisión del símbolo a_j. Denotamos por $x = (x_j)$ la distribución de probabilidad de los símbolos de entrada, se tiene $x_j \geq 0 \quad \sum_j x_j = 1$.

La definición de velocidad de transmisión de un canal, según Reza (1961), es:

$$T(x) = \frac{\sum_i \sum_j x_j p_{ij} \log(p_{ij} / \sum_k x_k p_{ik})}{\sum_j t_j x_j}$$

La capacidad del canal,C, es definida como el máximo de velocidad de transmisión, es decir, máximo de $T(x)$. Entonces encontramos el siguiente problema de Optimización fraccionaria:

$$\text{máx}\{T(x) : x_j \geq 0, \sum_j x_j = 1\} \tag{1.1}$$

1.1.2. Camino más corto

En teoría de grafos, el problema de buscar el camino más corto que une dos nodos es un problema clásico, para los cuales existen algoritmos eficientes de complejidad algorítmica como el algoritmo de Dijkstra. Dicho problema también, en ocasiones, se puede plantear como problema de optimización fraccionaria: Consideramos un grafo $G = (E, V)$ con $E = \{x_1, x_2, \cdots, x_n\}$ conjunto de n nodos y $V = \{e_1, e_2, \cdots, e_m\}$ conjunto de m aristas. Supongamos que del nodo x_1 se quiere mandar una determinada mercancía hacia en nodo x_n, transportar dicha mercancía de un nodo x_i hasta x_j tiene un coste de c_{ij} y se requiere t_{ij} unidades de tiempo. El objetivo es minimizar es coste total del trayecto por unidad de tiempo, con lo cual, el problema se puede plantear como:
Consideramos x_{ij} variables binarias que toman el valor 1 si en el camino se considera la arista (x_i, x_j)

$$\begin{aligned}
&\text{mín} \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{m} t_{ij} x_{ij}} \\
&\text{s.a} \\
&\qquad \sum_j x_{1j} = 1, \sum_i x_{in} = 1 \\
&\qquad \sum_j x_{ij} = \sum_j x_{ji} \quad \forall i
\end{aligned} \tag{1.2}$$

Para evitar ciclos se podría usar la técnica de etiquetar nodos.
La solución a dicho problema ofrece la ruta con el menor coste/tiempo.

1.1.3. Transporte con incertidumbre

Supongamos que se tiene n puntos de repartición de cierto producto y m destinos. Enviar una cantidad x_{ij} del punto i hacia el destino j supone una ganancia de b_{ij} y un gasto de c_{ij}. En la realidad, no se suele conocer a ciencia exacta la demanda del mercado en el destino m, y entonces se estima mediante probabilidades(lo que en la literatura puede encontrarse bajo el nombre de optimización fraccionaria estocástica). El objetivo del problema es transportar al destino m cierta cantidad de producto con el objetivo de maximizar el beneficio por unidad de producto. Con lo cual el problema se plantea como sigue:

$$\begin{aligned}
&\text{máx } \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(b_{ij}-c_{ij})x_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m}x_{ij}} \\
&\text{s.a} \\
&\qquad \sum_{j} x_{ij} \leq a_i \quad \forall i \\
&\qquad P_r\left(\sum_{i} x_{ij} \geq r_j\right) \geq 1 - l_j \ \forall j \\
&\qquad x_{ij} \geq 0 \quad \forall i, j
\end{aligned} \tag{1.3}$$

donde $b_{ij} \sim N(\mu_{b,ij}, \sigma^2_{b,ij})$ y $c_{ij} \sim N(\mu_{c,ij}, \sigma^2_{c,ij})$. Las cantidades a_i son las cantidades disponibles en punto de repartición i y los parámetros r_j y l_j son cantidades dadas obtenidas por un estudio estadístico del mercado.
Un ejemplo resuelto puede hallarse en Charles (2007).

1.1.4. Análisis descriminante.

Teniendo grupos clasificados según unas medidas(variables aleatorias), se pretende clasificar otros objetos según sus puntuaciones-medidas en dichas variables. La estrategia para afrontar estos problemas de clasificación es buscar funciones discriminadoras que clasifiquen con cierta probabilidad de acierto. Para nuestros intereses, supongamos que tenemos dos grupos o poblaciones, normales y con la misma matriz de covarianzas V definida positiva. Identificamos a los grupos como $X = (x_1, x_2, \cdots, x_k)'$ y $Y = (y_1, y_2, \cdots, y_k)'$ y con vectores media $\mu_X = E(X), \ \ \mu_Y = E(Y)$.
Sea un nuevo objeto Z con medidas $Z = (z_1, z_2, \cdots, z_k)'$ que queremos clasificar o bien en grupo X o en grupo Y. La idea de R. A. Fisher Sc.D. (1936) fue definir alguna función lineal $L(Z) = a'Z$ para clasificar bajo normalidad, de modo que se precisa buscar el vector a que disperse al máximo la proyección de las medias $a'\mu_X$ y $a'\mu_Y$ y disminuir la varianza común. Esta idea se plasma resolviendo el problema de optimización

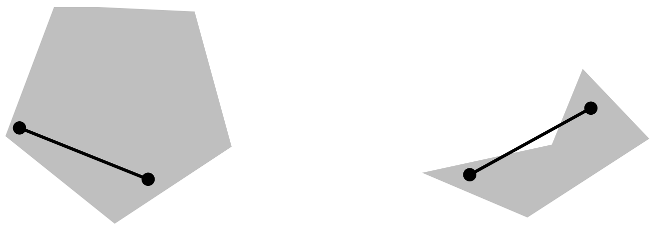
$$\text{máx}_a \left\{ \frac{(a'\mu_X - a'\mu_Y)^2}{a'Va} : a \in \mathbb{R}^n \right\} \tag{1.4}$$

Con solución general $a = \lambda V^{-1}(\mu_X - \mu_Y), \lambda \neq 0$.

1.2. Funciones convexas y sus generalizaciones

El objetivo de esta sección es mostrar algunos resultados de convexidad que serán útiles para el caso de programación fraccionaría. Para posteriormente ver como podemos relacionar dichos resultados al caso fraccionario y poder aplicar las propiedades conocidas en la optimización no lineal general al caso fraccionario. Como referencias bibliográficas podemos citar a Bazaraa, Sherali y Shetty (2006), Cambini y Martein (2009), Boyd y Vandenberghe (2004) y Minasian (1997).

Definición 1.1 (Conjunto convexo). *Dado* $A \subset \mathbb{R}^n$ *se dice convexo si* $\forall x_1, x_2 \in A$, *el segmento* $\lambda x_1 + (1-\lambda)x_2 \in A \;\; \forall \lambda \in [0,1]$



(a) Conjunto convexo (b) Conjunto no convexo

Definición 1.2 (Función convexa). *Sea* $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A *convexo, diremos que* f *es convexa en* A *si* $\forall x_1, x_2 \in A$ *y* $x_\lambda := \lambda x_1 + (1-\lambda)x_2$ *con* $\lambda \in [0,1]$ *se tiene que* $f(x_\lambda) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$

Definición 1.3 (Función estrictamente convexa). *Sea* $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A *convexo, diremos que* f *es estrictamente convexa en* A *si* $\forall x_1, x_2 \in A, x_1 \neq x_2$ *y* $x_\lambda := \lambda x_1 + (1-\lambda)x_2$ *con* $\lambda \in (0,1)$ *se tiene que* $f(x_\lambda) < \lambda f(x_1) + (1-\lambda) f(x_2)$

Definición 1.4 (mínimo local y global). *Dado un problema de optimización* mín$\{f(x) : x \in X\}$.
Diremos que $\overline{x}$ *es un mínimo local si* $\overline{x} \in X$ *y existe un entorno* $\beta_\epsilon(\overline{x})$ *tales que* $f(\overline{x}) \leq f(x) \; \forall x \in X \cap \beta_\epsilon(\overline{x})$.
Diremos que $\overline{x}$ *es un mínimo global si* $\overline{x} \in X$ *y* $f(\overline{x}) \leq f(x) \; \forall x \in X$.

Veamos a continuación algunas propiedades cuya demostración puede encontrarse en cualquier libro sobre optimización no lineal como es Bazaraa, Sherali y Shetty (2006).

Proposición 1.1. *Si* f *es convexa en* X, *entonces el conjunto de minimizadores es convexo.*

Proposición 1.2. *Si f es convexa en X, entonces todo mínimo local es mínimo global.*

Proposición 1.3. *Si f es estrictamente convexa en X, entonces sólo puede tener un minimizador.*

Proposición 1.4. *Si f es no constante y cóncava en X, $\overline{x}$ mínimo global, entonces $\overline{x} \in \partial X$*

Definición 1.5 (Función cuasiconvexa). *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$ decimos que es cuasiconvexa en X si $f(\lambda x_1 + (1-\lambda)x_2) \leq$ máx$\{f(x_1), f(x_2)\}$ $\forall \lambda \in [0,1], \forall x_1, x_2 \in X$*

Definición 1.6 (Función estrictamente cuasiconvexa). *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$ decimos que es estrictamente cuasiconvexa en X si $f(\lambda x_1 + (1-\lambda)x_2) <$ máx$\{f(x_1), f(x_2)\}$ $\forall \lambda \in (0,1), \forall x_1, x_2 \in X$ con $f(x_1) \neq f(x_2)$*

Definición 1.7 (Función explícitamente cuasiconvexa). *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$ decimos que es explícitamente cuasiconvexa si es cuasiconvexa y estrictamente cuasiconvexa en X*

Definición 1.8 (Función fuertemente cuasiconvexa). *f es fuertemente cuasiconvexa en X si $f(\lambda x_1 + (1-\lambda)x_2) <$ máx$\{f(x_1), f(x_2)\}$ $\forall \lambda \in (0,1), \forall x_1, x_2 \in X$*

Proposición 1.5. *Si f es cuasiconvexa en X, entonces, el conjunto de minimizadores es convexo.*

Proposición 1.6. *Si f es estrictamente cuasiconvexa en X, entonces, todo mínimo local es global.*

Proposición 1.7. *Si f es explícitamente cuasiconvexa en X, entonces, el conjunto de minimizadores es convexo y todo mínimo local es global.*

Proposición 1.8. *Si f es fuertemente cuasiconvexa en X, entonces, el mínimo global si se alcanza es único.*

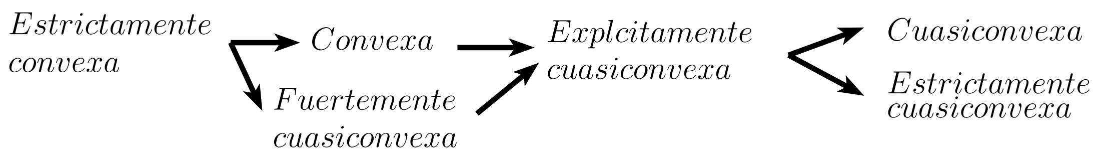


FIGURA 1.1: Relación entre las generalizaciones de funciones convexas.

Definición 1.9 (función cóncava). *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A convexo, diremos que f es cóncava en A si $\forall x_1, x_2 \in A$ y $x_\lambda := \lambda x_1 + (1-\lambda)x_2$ con $\lambda \in [0,1]$ se tiene que $f(x_\lambda) \geq \lambda f(x_1) + (1-\lambda)f(x_2)$*

Proposición 1.9. *f convexa en $X \Leftrightarrow -f$ cóncava en X*

Dicho resultado permite adaptar los resultados que se verán para convexidad al caso de funciones cóncavas.

1.3. Funciones convexas diferenciables y sus generalizaciones

Al igual que se hizo en la sección anterior, vemos algunas propiedades de las funciones convexas diferenciables que adaptaremos en la sección posterior al caso f/h.

Definición 1.10 (Función diferenciable). *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$, X abierto, diremos que es diferenciable en $\overline{x} \in X$ si existe un vector $\nabla f(\overline{x})$ tal que:*

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})(x-\overline{x}) + \alpha(\overline{x}, x-\overline{x})\|x-\overline{x}\| \text{ con } \alpha(x, x-\overline{x}) \underset{x\to\overline{x}}{\to} 0$$

Definición 1.11 (Función dos veces difrenciable). *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$, X abierto, diremos que es dos veces diferenciable en $\overline{x} \in X$ si existe un vector $\nabla f(\overline{x})$ y una matriz $H(\overline{x})$ tales que:*

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})(x-\overline{x}) + \tfrac{1}{2}(x-\overline{x})' H(\overline{x})(x-\overline{x}) + \alpha(\overline{x}, x-\overline{x})\|x-\overline{x}\|$$

con $\alpha(x, x-\overline{x}) \underset{x\to\overline{x}}{\to} 0$

Proposición 1.10. *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$, X abierto y convexo, f dos veces diferenciable en X. Entonces: f convexa en $X \Leftrightarrow H(x)$ es semidefinida positiva $\forall x \in X$*

Proposición 1.11. *Si $H(x)$ es definida positiva en X, entonces, f es estrictamente convexa.*

Proposición 1.12. *Si f es convexa y diferenciable en X, $\overline{x} \in X$ es mínimo global de f en $X \Leftrightarrow \nabla f(\overline{x})(x-\overline{x}) \geq 0 \ \forall x \in X$*

Proposición 1.13. *Si f es convexa y diferenciable en X, $\overline{x} \in X$ es mínimo global de f en X abierto $\Leftrightarrow \nabla f(\overline{x}) = 0$*

Las demostraciones pueden hallarse en cualquier libro de análisis real de varias variables, por ejemplo Vera (2008).

Definición 1.12 (Función pseudoconvexa). *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$, X abierto y convexo, f convexa y diferenciable en X, diremos que es pseudoconvexa en X si $\forall x, \overline{x} \in X$ tales que si $\nabla f(\overline{x})(x-\overline{x}) \geq 0$, entonces, $f(x) \geq f(\overline{x})$*

Definición 1.13 (Función estrictamente pseudoconvexa). *Dada $f : X \subset \mathbb{R}^n \to \mathbb{R}$, X abierto y convexo, f convexa y diferenciable en X, diremos que es estrictamente pseudoconvexa en X si $\forall x \neq \overline{x} \in X$ tales que si $\nabla f(\overline{x})(x-\overline{x}) \geq 0$, entonces, $f(x) > f(\overline{x})$*

Proposición 1.14. *Si f pseudoconvexa en X abierto y convexo y $\nabla f(\overline{x}) = 0$, entonces, $\overline{x}$ es mínimo global.*

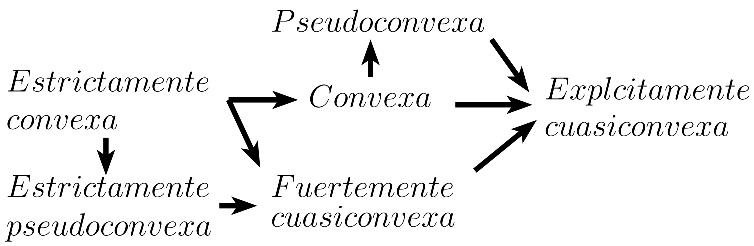


FIGURA 1.2: Relación entre las generalizaciones de funciones convexas diferenciables.

1.4. Convexidad, diferenciabilidad y sus generalizaciones para el caso f/h

Se dan resultados que nos permiten conocer cuando f/h es algún tipo de generalización de convexidad y diferenciabilidad vistos en las secciones 1.2 y 1.3 en función del tipo de convexidad de f y h por separado. Los resultados aquí mostrados pueden hallarse en Minasian (1997, capítulo 2), así como en el artículo Mangasarian (1970) re-editado en el libro Ziemba y Vickson (1975, pág 33-41)

Nuestro objetivo es conocer la convexidad de una función de la forma f/h en función del tipo de convexidad de f y h por separado, pero en general no existe ninguna regla. Por ejemplo: considérese la función constante 1 y la función convexa en todo $\mathbb{R}$ $e^{x^2} > 0$, el cociente,e^{-x^2} es una función que ni es convexa ni es cóncava en todo $\mathbb{R}$. Por tanto descartamos la posibilidad de que convexa/convexa o cóncava/convexa nos diera una función convexa. El siguiente ejemplo muestra que tampoco es cierto en general que convexa/cóncava o cóncava/cóncava nos diese una función convexa: tómese $f(x) = x$ definida para $x > 0$ y $h(x) = \sqrt{x}$ definida para $x > 0$, el cociente es $\sqrt{x}$ función no convexa.

Proposición 1.15. *Si f es convexa y no negativa, h convexa y positiva, y se cumple $(f(x_2)-f(x_1))(h(x_1) - h(x_2)) \leq 0$ $\forall x_1, x_2 \in X$. Entonces $f \cdot h$ es convexa.*

Demostración.
Sean $x_1, x_2 \in X$, sea $\lambda \in [0, 1]$, entonces: $f \cdot h$ es convexa en X
$:\Leftrightarrow f(\lambda x_1 + (1-\lambda)x_2)h(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1)h(x_1) + (1-\lambda)f(x_1)h(x_1)$.
Usando que f y h son convexas y no negativas se tiene que:

$$f(\lambda x_1 + (1-\lambda)x_2)h(\lambda x_1 + (1-\lambda)x_2) \leq (\lambda f(x_1) + (1-\lambda)f(x_2))(\lambda h(x_1) + (1-\lambda)h(x_2)) =$$
$$\lambda f(x_1)h(x_1) + (1-\lambda)f(x_2)h(x_2) + \lambda(1-\lambda)((f(x_2) - f(x_1))(h(x_1) - h(x_2))) \leq$$
$$\lambda f(x_1)h(x_1) + (1-\lambda)f(x_2)h(x_2) \text{ como se quería probar.}$$

$\square$

Notese que en caso de f, h de una variable real, una condición suficiente para que se satisfaga la condición de $(f(x_2) - f(x_1))(h(x_1) - h(x_2)) \leq 0\ \forall x_1, x_2 \in X$ es que tengan la misma monotonía.

Proposición 1.16. *Si h es cóncava y positiva, entonces $1/h$ es convexa.*

Demostración.
$1/h$ es convexa $:\Leftrightarrow \quad \forall x_1, x_2 \in X, \lambda \in [0,1]$ se cumple $\frac{1}{h(\lambda x_1+(1-\lambda)x_2)} \leq \lambda \frac{1}{h(x_1)} + (1-\lambda)\frac{1}{h(x_2)}$.
Al ser cóncava tenemos $h(\lambda x_1 + (1-\lambda)x_2) \geq \lambda h(x_1) + (1-\lambda)h(x_2) \quad \Leftrightarrow \quad \frac{1}{h(\lambda x_1+(1-\lambda)x_2)} \leq \frac{1}{\lambda h(x_1)+(1-\lambda)h(x_2)}$
Veremos que $\frac{1}{\lambda h(x_1)+(1-\lambda)h(x_2)} \leq \lambda \frac{1}{h(x_1)} + (1-\lambda)\frac{1}{h(x_2)}$ lo que completaría la prueba.
Al ser h positiva:

$$\frac{1}{\lambda h(x_1)+(1-\lambda)h(x_2)} \leq \lambda \frac{1}{h(x_1)} + (1-\lambda)\frac{1}{h(x_2)} \Leftrightarrow$$
$$h(x_1)h(x_2) \leq (\lambda h(x_1) + (1-\lambda)h(x_2))(\lambda h(x_2) + (1-\lambda)h(x_1)) \Leftrightarrow$$

Haciendo el producto y agrupando términos

$$0 \leq \lambda(1-\lambda)(h(x_1) - h(x_2))^2$$

Lo cual es cierto y por tanto $\frac{1}{h(\lambda x_1+(1-\lambda)x_2)} \leq \lambda \frac{1}{h(x_1)} + (1-\lambda)\frac{1}{h(x_2)}$ □

Corolario 1.1. *Sea $X \subset \mathbb{R}^n$ convexo, $f, h : X \to \mathbb{R}$ funciones reales de variable real que satisfacen:*

1. *f convexa y no negativa en X*
2. *h cóncava y positiva en X*
3. *$(f(x_2) - f(x_1))(h(x_1) - h(x_2)) \leq 0\ \forall x_1, x_2 \in X$*

Entonces, f/h es convexa en X.

Demostración.
Es una generalización del Ejercicio 3.32 y Ejemplo 3.13 de Boyd y Vandenberghe (2004)
Basta usar las proposiciones 1.15 y 1.16 para tener que $f\frac{1}{h}$ es convexa. □

El resultado anterior es bastante exigente por la condición tres y el ejemplo mostrado en el párrafo anterior a la Proposición 1.15 muestra que es imprescindible. Entonces, por lo general no se obtiene convexidad. Veremos bajo qué condiciones se obtiene algún tipo de generalización más débil como pseudoconvexidad o cuasiconvexidad.

Proposición 1.17. *Sea X abierto, $f, h : X \to \mathbb{R}$ diferenciables, $h(x) \neq 0$ en X, entonces, f/h es diferenciable en X y además:*

$$\nabla \frac{f(x)}{h(x)} = \frac{h(x)\nabla f(x) - f(x)\nabla h(x)}{h(x)^2}$$

Demostración.
Basta ver que $\lim\limits_{x\to\overline{x}} \frac{\frac{f(x)}{h(x)}-\frac{f(\overline{x})}{h(\overline{x})}-\frac{h(x)\nabla f(x)-f(x)\nabla h(x)}{h(x)^2}(x-\overline{x})}{\|x-\overline{x}\|} = 0$
Tras realizar operaciones algebraicas obtenemos que:

$\lim\limits_{x\to\overline{x}} \frac{\frac{f(x)}{h(x)}-\frac{f(\overline{x})}{h(\overline{x})}-\frac{h(x)\nabla f(x)-f(x)\nabla h(x)}{h(x)^2}(x-\overline{x})}{\|x-\overline{x}\|} =$
$\lim\limits_{x\to\overline{x}} \frac{1}{h(x)}\frac{f(x)-f(\overline{x})-\nabla f(\overline{x})(x-\overline{x})}{\|x-\overline{x}\|} - \lim\limits_{x\to\overline{x}} \frac{f(x)}{h(x)^2}\frac{h(x)-h(\overline{x})-\nabla h(\overline{x})(x-\overline{x})}{\|x-\overline{x}\|} + \lim\limits_{x\to\overline{x}} \frac{f(x)}{h(x)}(1-\frac{h(\overline{x})}{h(x)})$, usando que f y h son diferenciables, obtenemos que los límites anteriores son todos nulos y por ende $\lim\limits_{x\to\overline{x}} \frac{\frac{f(x)}{h(x)}-\frac{f(\overline{x})}{h(\overline{x})}-\frac{h(x)\nabla f(x)-f(x)\nabla h(x)}{h(x)^2}(x-\overline{x})}{\|x-\overline{x}\|} = 0$ como se quería probar. □

Para dar el resultado principal de esta sección, necesitamos dar ciertas definiciones y notación. Sea $A \subset \mathbb{R}^m \times \mathbb{R}^k$, consideramos $x \in A$ y consideramos su descomposición en $x^1 = (x_1, x_2, \cdots, x_m) \in \mathbb{R}^m$ y $y^1 = (y_1, y_2, \cdots, y_k) \in \mathbb{R}^k$, entonces $x = (x^1, y^1)$
Tomando $a \in A, b \in A,\ a \le b \Leftrightarrow a_i \le b_i\ \forall i \in \{1, 2, \cdots, m, m+1, \cdots, m+k\}$. Lo mismo se considera en cada componente si nos restringimos a $\mathbb{R}^m$ o $\mathbb{R}^k$. Por ejemplo, $x^1 \le x^2 \Leftrightarrow x_i^1 \le x_i^2\ \forall i \in \{1, 2, \cdots, m\}$.

Definición 1.14. *Sea $\phi : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}$, se dice:*

- *creciente-decreciente si $\{x^1 \le x^2, y^2 \le y^1\} \Rightarrow \phi(x^1, y^1) \le \phi(x^2, y^2)$*
- *creciente-____ si $x^1 \ge x^2 \Rightarrow \phi(x^1, y) \ge \phi(x^2, y)$ para cualquier $y \in \mathbb{R}^k$.*
- *decreciente-____ si $x^2 \ge x^1 \Rightarrow \phi(x^1, y) \ge \phi(x^2, y)$ para cualquier $y \in \mathbb{R}^k$.*

Teorema 1.1 (Teoremas de composición). *Sea $X \in \mathbb{R}^n$ convexo, sean $f : X \to \mathbb{R}^m$, $g : X \to \mathbb{R}^k$, $\phi : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}$, y definimos:*

$$\theta : X \to \mathbb{R} \text{ como } \theta(x) = \phi(f(x), g(x))$$

Entonces:
Supongamos que se da alguna condición de cuadro 1.1, entonces:

1. *Si ϕ es estrictamente convexa en $\mathbb{R}^m \times \mathbb{R}^k$, entonces, θ es estrictamente convexa en X.*
2. *Si X es abierto, f y g diferenciables en X, ϕ estrictamente pseudoconvexa en $\mathbb{R}^m \times \mathbb{R}^k$, entonces, θ es estrictamente pseudoconvexa en X.*
3. *Si ϕ es explícitamente cuasiconvexa en $\mathbb{R}^m \times \mathbb{R}^k$, entonces, θ es explícitamente cuasiconvexa en X.*

	f (en X)	g (en X)	ϕ (en $\mathbb{R}^m \times \mathbb{R}^k$)
i)	convexa	cóncava	creciente-decreciente
ii)	lineal	lineal	_____ - _____
iii)	convexa	lineal	creciente-_____
iv)	cóncava	lineal	decreciente-____

Cuadro 1.1

Demostración.
1.i) Tenemos que f es convexa en X, g cóncava en X y ϕ creciente en la primera coordenada y decreciente en la segunda coordenada. Supongamos que ϕ es estrictamente convexa en $\mathbb{R}^m \times \mathbb{R}^k$ y queremos ver que θ es estrictamente convexa en X.
Tomamos $x_1, x_2 \in X, \lambda \in (0,1), x_1 \neq x_2, (f(x_1), g(x_1)) \neq (f(x_2), g(x_2))$
Consideramos $x_\lambda = \lambda x_1 + (1-\lambda)x_2$. Por las hipótesis consideradas en i), tenemos que:
$\theta(x_\lambda) = \phi(f(x_\lambda), g(x_\lambda)) \leq \phi(\lambda f(x_1)+(1-\lambda)f(x_2), \lambda g(x_1)+(1-\lambda)g(x_2)) = \phi(\lambda(f(x_1), g(x_1))+ (1-\lambda)(f(x_2), g(x_2))) \underset{\phi\text{Est.convexa}}{<} \lambda\phi(f(x_1), g(x_1)) + (1-\lambda)\phi(f(x_2), g(x_2)) = \lambda\theta(x_1) + (1-\lambda)\theta(x_2)$
Para 1.ii), 1.iii) y 1.iv) las demostraciones son similares a la anterior.

2.i) Consideramos las condiciones dadas en i) y ϕ estrictamente pseudoconvexa. Veamos que θ es estrictamente pseudoconvexa.
Sean $x_1, x_2 \in X$ y $\nabla\theta(x_2)'(x_1 - x_2) \geq 0$
Utilizando la regla de la cadena tenemos:

$$\nabla\theta(x_2)'(x_1 - x_2) = \nabla\phi(f(x_2), g(x_2))'(x_1 - x_2) =$$
$$\nabla_1\phi(f(x_2), g(x_2))'\nabla f(x_2)(x_1 - x_2) + \nabla_2\phi(f(x_2), g(x_2))'\nabla g(x_2)(x_1 - x_2)$$

Entonces:
$0 \leq \nabla_1\phi(f(x_2), g(x_2))'\nabla f(x_2)(x_1 - x_2) + \nabla_2\phi(f(x_2), g(x_2))'\nabla g(x_2)(x_1 - x_2)$
Usando las hipótesis de i):

- Al ser f convexa, podemos concluir que $\nabla f(x_2)(x_1 - x_2) \leq f(x_1) - f(x_2)$
  y al ser ϕ creciente en primera componente, tenemos que $\nabla_1\phi(f(x_2), g(x_2))'\nabla f(x_2)(x_1 - x_2) \leq \nabla_1\phi(f(x_2), g(x_2))'(f(x_1) - f(x_2))$
- Al ser g cóncava y ϕ decreciente en la segunda componente, tenemos que:
  $\nabla_1\phi(f(x_2), g(x_2))'\nabla g(x_2)(x_1 - x_2) \leq \nabla_1\phi(f(x_2), g(x_2))'(g(x_1) - g(x_2))$

Por tanto:
$0 \leq \nabla_1\phi(f(x_2), g(x_2))'\nabla f(x_2)(x_1 - x_2) + \nabla_2\phi(f(x_2), g(x_2))'\nabla g(x_2)(x_1 - x_2) \leq$
$\nabla_1\phi(f(x_2), g(x_2))'(f(x_1) - f(x_2)) + \nabla_1\phi(f(x_2), g(x_2))'(g(x_1) - g(x_2)) =$
$\nabla\phi(f(x_2), g(x_2))'((f(x_1), g(x_1)) - (f(x_2), g(x_2)))$, usando ahora que ϕ es estrictamente pseudoconvexa, se tiene que:

$$\phi(f(x_1), g(x_1)) > \phi(f(x_2), g(x_2))$$
$$\theta(x_1) > \theta(x_2) \text{ como se quería probar.}$$

Para 2.ii), 2.iii) y 2.iv) las demostraciones son similares a la anterior.

Veamos ahora 3.i). Supongamos que se dan las condiciones de i) y que ϕ es explícitamente cuasiconvexa en $\mathbb{R}^m \times \mathbb{R}^k$, queremos ver que entonces θ es explícitamente cuasiconvexa en X.
Sean $x_1, x_2 \in X, x_1 \neq x_2$, sea $\lambda \in (0,1)$, consideramos $x_\lambda = \lambda x_1 + (1-\lambda)x_2$
Supongamos que $\theta(x_1) < \theta(x_2)$ ¿$\theta(x_\lambda) < \theta(x_2)$?
$\phi(f(x_1), g(x_1)) < \phi(f(x_2), g(x_2))$, al ser ϕ explícitamente cuasiconvexa se tiene que:

$$\phi(\lambda(f(x_1), g(x_1)) + (1-\lambda)(f(x_2), g(x_2))) < \phi(f(x_2), g(x_2))$$
$$\phi(\lambda f(x_1) + (1-\lambda)f(x_2), \lambda g(x_1) + (1-\lambda)g(x_2)) < \phi(f(x_2), g(x_2))$$

Usando ahora las condiciones dadas en i), f convexa y ϕ creciente en primera componente, y por otra parte g cóncava y ϕ decreciente en segunda componente, se tiene:

$$\phi(f(x_\lambda), g(x_\lambda)) \leq \phi(\lambda f(x_1) + (1-\lambda)f(x_2), \lambda g(x_1) + (1-\lambda)g(x_2)) < \phi(f(x_2), g(x_2))$$

Es decir,

$$\theta(x_\lambda) < \theta(x_2)$$

Las pruebas para 3.ii), 3.iii) y 3.iv) son similares. □

Corolario 1.2. *Sea $X \in \mathbb{R}^n$ convexo, sean $f : X \to \mathbb{R}^m$,
$g : X \to \mathbb{R}^k$, $\phi : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}$, y definimos:*

$$\theta : X \to \mathbb{R} \text{ como } \theta(x) = \phi(f(x), g(x))$$

Supongamos que se da alguna condición de cuadro 1.2, entonces:

1. *Si ϕ es estrictamente cóncava en $\mathbb{R}^m \times \mathbb{R}^k$, entonces, θ es estrictamente cóncava en X.*

2. *Si X es abierto, f y g diferenciables en X, ϕ estrictamente pseudocóncava en $\mathbb{R}^m \times \mathbb{R}^k$, entonces, θ es estrictamente pseudocóncava en X.*

3. *Si ϕ es explícitamente cuasicóncava en $\mathbb{R}^m \times \mathbb{R}^k$, entonces, θ es explícitamente cuasicóncava en X.*

	f *(en X)*	g *(en X)*	ϕ *(en $\mathbb{R}^m \times \mathbb{R}^k$)*
i)	*cóncava*	*convexa*	*creciente-decreciente*
ii)	*lineal*	*lineal*	____ - ____
iii)	*cóncava*	*lineal*	*creciente-____*
iv)	*convexa*	*lineal*	*decreciente-____*

CUADRO 1.2

Teorema 1.2. *Sea $X \in \mathbb{R}^n$ convexo, $f : X \to \mathbb{R},\ h : X \to \mathbb{R}$ con $h(x) \neq 0\ \forall x \in X$*
Entonces:
a) Si f y h cumplen alguna condición de cuadro 1.3, entonces f/h es explícitamente cuasiconvexa en X.

	1	2	3	4	5	6
f	convexa ≥ 0	cóncava ≤ 0	convexa ≤ 0	cóncava ≥ 0	convexa	cóncava
h	cóncava > 0	convexa < 0	convexa > 0	cóncava < 0	lineal > 0	lineal < 0

Cuadro 1.3

b) Si f y h cumplen alguna condición de cuadro 1.4, entonces f/h es explícitamente cuasicóncava en X.

	1	2	3	4	5	6
f	cóncava ≥ 0	convexa ≤ 0	cóncava ≤ 0	convexa ≥ 0	cóncava	convexa
h	convexa > 0	cóncava < 0	cóncava > 0	convexa < 0	lineal > 0	lineal < 0

Cuadro 1.4

c) Si f y h cumplen alguna condición de cuadro 1.5, entonces f/h es estrictamente pseudoconvexa en X.

	1	2	3
f	convexa ≥ 0	cóncava ≤ 0	convexa ≤ 0
h	estr.cóncava > 0	estr.convexa < 0	estr.convexa > 0
	4	5	6
f	cóncava ≥ 0	estr.convexa	estr.cóncava
h	estr.cóncava < 0	lineal > 0	lineal < 0

Cuadro 1.5

d) Si f y h cumplen alguna condición de cuadro 1.6, entonces f/h es estrictamente pseudocóncava en X.

	1	2	3
f	cóncava ≥ 0	convexa ≤ 0	cóncava ≤ 0
h	estr.cóncava > 0	estr.convexa < 0	estr.convexa > 0
	4	5	6
f	convexa ≥ 0	estr.cóncava	estr.convexa
h	estr.cóncava < 0	lineal > 0	lineal < 0

Cuadro 1.6

Demostración.
Sea $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ definida como $\phi(x, y) = \frac{x}{y}$

Sea $\theta(x) = \phi(f(x), h(x)) = \frac{f(x)}{h(x)}$
Como en el problema que nos ocupa (PF) $h > 0$, consideramos como dominio de ϕ a $D = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y > 0\}$ que es un conjunto convexo.
Entonces la demostración sigue de Teorema 1.1 y Corolario 1.2, observando que para los casos a)y b) ϕ es pseudoconvexa (y por tanto explícitamente cuasiconvexa 1.2), en los casos c) y d) ϕ es estrictamente pseudoconvexa.

□

Observación 1.1. En los apartados c) y d), si no consideramos la condición estrictamente convexa ó cóncava, obtenemos que f/h es pseudoconvexa ó pseudocóncava sin estrictamente.

1.5. Transformación Manas-Schaible

Mostraremos la transformación de Manas-Schaible, usada con frecuencia posteriormente, que bajo ciertas condiciones transforma un problema de programación fraccionaria (PF) en un problema convexo no lineal y no fraccionario que llamaremos (PMS). Los contenidos aquí mostrados corresponden a los artículos Schaible (1976b) y C.Singh (1981).

Consideramos el problema:

$$\text{(PF)} \qquad \text{mín}\left\{\frac{f(x)}{h(x)} : g(x) \leq 0, x \in X\right\}$$

donde consideramos que $h(x) > 0 \ \ \forall x \in X \subset \mathbb{R}^n$.
Definimos:

$$z_0 = 1/h(x), \qquad z = xz_0 \quad \text{con } x \in X$$

$$W = \{(z_0, z) \in \mathbb{R} \times \mathbb{R}^n : z_0 > 0, z/z_0 \in X\}$$

Entonces, el problema (PF) es equivalente a este otro:

$$\text{(PMS)} \qquad \text{mín}\{z_0 f(z/z_0) : g(z/z_0) \leq 0, z_0 h(z/z_0) = 1, (z_0, z) \in W\} \tag{1.5}$$

Proposición 1.18. *Si X es convexo,f, g convexas en X, h lineal, entonces, el problema (PMS) es un problema convexo.*

Demostración.
Veamos primero que siendo X convexo, W es igualmente convexo.
Sean $(z_0^1, z^1), (z_0^2, z^2) \in W$, sea $\lambda \in [0, 1]$. Sea $(z_0^\lambda, z^\lambda) = \lambda(z_0^1, z^1) + (1 - \lambda)(z_0^2, z^2)$
$z_0^\lambda = \lambda z_0^1 + (1 - \lambda)z_0^2 > 0$
$\frac{z^\lambda}{z_0^\lambda} = \frac{\lambda z^1 + (1-\lambda)z^2}{\lambda z_0^1 + (1-\lambda)z_0^2} = \lambda \frac{z^1}{\lambda z_0^1 + (1-\lambda)z_0^2} + (1 - \lambda)\frac{z^2}{\lambda z_0^1 + (1-\lambda)z_0^2} = \lambda \frac{z^1}{\lambda z_0^1 + (1-\lambda)z_0^2}\frac{z_0^1}{z_0^1} + (1 - \lambda)\frac{z^2}{\lambda z_0^1 + (1-\lambda)z_0^2}\frac{z_0^2}{z_0^2}$
$= \lambda \frac{z_0^1}{\lambda z_0^1 + (1-\lambda)z_0^2}\frac{z^1}{z_0^1} + (1 - \lambda)\frac{z_0^2}{\lambda z_0^1 + (1-\lambda)z_0^2}\frac{z^2}{z_0^2} \in X$ al ser X convexo.
Veamos ahora que $z_0 f(z/z_0)$ es convexa.

Sean $(z_0^1, z^1), (z_0^2, z^2) \in W$, sea $\lambda \in [0,1]$.
$(\lambda z_0^1+(1-\lambda)z_0^2) f\left(\frac{\lambda z^1+(1-\lambda)z^2}{\lambda z_0^1+(1-\lambda)z_0^2}\right) = (\lambda z_0^1+(1-\lambda)z_0^2) f\left(\lambda \frac{z_0^1}{\lambda z_0^1+(1-\lambda)z_0^2}\frac{z^1}{z_0^1} + (1-\lambda)\frac{z_0^2}{\lambda z_0^1+(1-\lambda)z_0^2}\frac{z^2}{z_0^2}\right)$
$\underset{\text{f convexa}}{\leq} (\lambda z_0^1 + (1-\lambda)z_0^2)\left(\frac{\lambda z_0^1}{\lambda z_0^1+(1-\lambda)z_0^2} f(z^1/z_0^1) + \frac{(1-\lambda)z_0^2}{\lambda z_0^1+(1-\lambda)z_0^2} f(z^2/z_0^2)\right) = \lambda z_0^1 f(z^1/z_0^1) +$
$(1-\lambda)z_0^2 f(z^2/z_0^2)$
Considerando además que h es lineal, g convexa, tenemos finalmente que (PMS) es un problema convexo. $\square$

Proposición 1.19. *Si X es convexo,$f, -h, g$ convexas en X,ínf$\{f(x) : x \in X\} < 0$, entonces, el problema (PMS) es un problema convexo.*

Demostración.
En caso de que ínf$\{f(x) : x \in X\} < 0$, el problema (PMS) es equivalente a:

$$\text{(PMS)} \qquad \text{mín}\{z_0 f(z/z_0) : g(z/z_0) \leq 0, z_0 h(z/z_0) \leq 1, (z_0, z) \in W\}$$

Pues en tal caso, la restricción $z_0 h(z/z_0) \geq 1$ es siempre válida en los puntos donde $f(x) \leq 0$ ya que equivale a $\frac{f(x)}{h(z/z_0)} \geq \frac{f(x)}{h(x)}$ $\forall x \in X$ con $f(x) \leq 0$. Siendo $h > 0$ en todo X, tenemos que mín$\left\{\frac{f(x)}{h(x)} : g(x) \leq 0, x \in X\right\} < 0$ y se alcanza para $f(x) < 0$, con lo cual podemos considerar sólo la restricción $z_0 h(z/z_0) \leq 1$.
Con las hipótesis consideradas, el segundo problema (PMS) es convexo, con la misma demostración que la Proposición 1.18. $\square$

Capítulo 2

Condiciones de optimalidad en Optimización Franccionaria

Sea $X \subset \mathbb{R}^n$, sean $f, h : X \to \mathbb{R}$, funciones reales de variable real. Sea $g : X \to \mathbb{R}^m$ una función vectorial de variable real. Supongamos que $h(x) > 0 \ \forall x \in X$ y consideramos el problema:

$$\text{(PF)} \qquad \text{mín} \left\{ \frac{f(x)}{h(x)} : g(x) \leq 0, x \in X \right\} \tag{2.1}$$

Sea $\mathcal{S} = \{x \in X : g(x) \leq 0\}$ el conjunto factible. Nos preguntamos qué condiciones han de cumplir $f, h, \mathcal{S}$ para que $\overline{x} \in \mathcal{S}$ sea mínimo global del problema (PF).

Se establecen condiciones de optimalidad para el problema (PF), primero sin considerar diferenciabilidad y posteriormente considerando diferenciabilidad. Las condiciones que se dan son las de Kuhn-Tucker y Fritz-John para un problema de optimización no lineal especificadas para el caso fraccional.

El contenido aquí desarrollado es una adaptación del artículo "Optimality Conditions in Fractional Programming"de C.Singh, véase C.Singh (1981). En Minasian (1997) también pueden hallarse demostrada la condición necesaria y suficiente para el caso con diferenciablidad en teoremas 3.3.1 y 4.1.1. Demostraciones del caso particular de $h = 1$ puede hallarse en Mangasarian (1969) y Bazaraa, Sherali y Shetty (2006, Capítulos 5,7; Capítulos 4,6).

Cualesquiera que sean f, h y g podemos establecer condiciones de optimalidad mediante el uso de las funciones lagrangianas y los problemas de punto de silla asociadas a ellas. Para la condición suficiente sin diferenciabilidad no será necesaria ninguna condición de convexidad, en cambio, para la condición necesaria, la convexidad junto con alguna condición de regularidad serán imprescindibles.
En el caso de f, h, g diferenciables, para la condición necesaria no es necesario tener convexidad de f, h, g, en cambio, para la condición suficiente sí es necesario que $f, -h$ sean convexas y las restricciones g_i cuasiconvexas al menos.

Estructura del capítulo

Sección 2.1. Se dan condiciones de optimalidad de Fritz-John y Kuhn-Tucker para el caso sin diferenciabilidad adaptados al caso de Optimización Fraccionaria.
Sección 2.2. Se dan condiciones de optimalidad de Fritz-John y Kuhn-Tucker para el caso con diferenciabilidad adaptados al caso de Optimización Fraccionaria.
Sección 2.3. Se muestra la relación entre la Optimización Fraccionaria y Paramétrica. Relación que permite dar métodos de resolución como es el algoritmo de Dinkilbach basado en el resultado central Teorema 2.7.
Sección 2.4. Sección dedicada al caso donde la región factible es un poliedro. Los teoremas 2.11 y 2.12 muestran los casos en los que el óptimo del problema se encuentra en un vértice del poliedro.
Sección 2.5. Se dan las definiciones y teoremas de dualidad, tanto para el caso sin diferenciabilidad como el caso con diferenciabilidad.

2.1. Condiciones generales de optimalidad

Consideremos las siguientes funciones:

- Función lagrangiana Kuhn-Tucker (F.L.K.T):
  Para $x \in X, u \in \mathbb{R}^m, u \geq 0$, se define $\mathcal{K}(x,u) = \frac{f(x)}{h(x)} + ug(x)$.
- Función lagrangiana Fritz-John (F.L.F.J):
  Para $x \in X, r_0 \in \mathbb{R}, r \in \mathbb{R}^m, (r_0, r) \geq 0,\ (r_0, r) \neq 0$, se define $\mathcal{F}(x, r_0, r) = r_0 \frac{f(x)}{h(x)} + rg(x)$.

Diremos que $(\overline{x}, \overline{u})$,con $\overline{x} \in X$ y $\overline{u} \geq 0$, es un punto de silla de Kuhn-Tucker (P.S.K.T) si cumple:

$$\mathcal{K}(\overline{x}, u) \leq \mathcal{K}(\overline{x}, \overline{u}) \leq \mathcal{K}(x, \overline{u}), \qquad \forall x \in X, \forall u \geq 0$$

Diremos que $(\overline{x}, \overline{r}_0, \overline{r})$ con $\overline{x} \in X, (\overline{r}_0, \overline{r}) \geq 0, (\overline{r}_0, \overline{r}) \neq 0$, es un punto de silla de Fritz-John (P.S.F.J) si cumple:

$$\mathcal{F}(\overline{x}, \overline{r}_0, r) \leq \mathcal{F}(\overline{x}, \overline{r}_0, \overline{r}) \leq \mathcal{F}(x, \overline{r}_0, \overline{r}), \qquad \forall x \in X, \forall r \geq 0$$

Proposición 2.1. $\exists (\overline{x}, \overline{u})$ *P.S.K.T* $\Leftrightarrow$ $\exists (\overline{x}, \overline{r}_0, \overline{r})$ *P.S.F.J con* $\overline{r}_0 > 0$

Demostración.
$\Rightarrow]$ Si $(\overline{x}, \overline{u})$ es P.S.K.T $\Rightarrow (\overline{x}, 1, \overline{u})$ es P.S.F.J
$\Leftarrow]$ Si $\Rightarrow (\overline{x}, \overline{r}_0, \overline{r})$ es P.S.F.J con $\overline{r}_0 > 0 \Rightarrow (\overline{x}, \frac{\overline{r}}{\overline{r}_0})$ es P.S.K.T □

Condición suficiente

Teorema 2.1. *Si $(\overline{x}, \overline{u})$ es un P.S.K.T, entonces, $\overline{x}$ es solución de (PF) y $\overline{u}g(\overline{x}) = 0$*

Demostración.
Sea $(\overline{x}, \overline{u})$ P.S.K.T, entonces

$$\frac{f(\overline{x})}{h(\overline{x})} + ug(\overline{x}) \underset{(1)}{\leq} \frac{f(\overline{x})}{h(\overline{x})} + \overline{u}g(\overline{x}) \underset{(2)}{\leq} \frac{f(x)}{h(x)} + \overline{u}g(x), \quad \forall x \in X, \forall u \geq 0$$

Para el caso $u = 0$ tenemos de la desigualdad (1) que $\overline{u}g(\overline{x}) \geq 0$, y como u puede tomar valores tan grandes como se quiera manteniendo cierta la desigualdad (1) se tiene que verificar que $g(\overline{x}) \leq 0$. Esto nos garantiza que $\overline{x}$ es un punto factible y que $\overline{u}g(\overline{x}) \leq 0$, como $\overline{u}g(\overline{x}) \geq 0$ resulta que $\overline{u}g(\overline{x}) = 0$.
Usando lo anterior, la desigualdad (2) queda así: $\frac{f(\overline{x})}{h(\overline{x})} \leq \frac{f(x)}{h(x)} + \overline{u}g(x) \quad \forall x \in X$. En particular, si $x \in \mathcal{S}$ entonces

$$\overline{u}g(x) \leq 0 \text{ y por tanto: } \frac{f(x)}{h(x)} \geq \frac{f(x)}{h(x)} + \overline{u}g(x) \geq \frac{f(\overline{x})}{h(\overline{x})} \text{ por lo que } \overline{x} \text{ es solución de (PF).}$$

□

El recíproco no es en general cierto, veamos un contraejemplo:
Sean $\frac{f(x)}{h(x)} = \frac{-x_1+x_2}{x_1+1}$ definida en $X = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0\}$ y $g(x) = x_1^2 + x_2$
Con lo cual el conjunto factible se reduce al punto $\mathcal{S} = \{(0,0)\}$, que es la solución a (PF) y además $g(\overline{x}) = 0$. Pero no existe P.S.K.T.
Supongamos que sí existe P.S.K.T $(\overline{x}, \overline{u})$, entonces, por el teorema $\overline{x}$ es mínimo global y por tanto $\overline{x} = (0,0)$. Y las inecuaciones del problema de punto de silla

$$\frac{f(\overline{x})}{h(\overline{x})} + ug(\overline{x}) \underset{(1)}{\leq} \frac{f(\overline{x})}{h(\overline{x})} + \overline{u}g(\overline{x}) \underset{(2)}{\leq} \frac{f(x)}{h(x)} + \overline{u}g(x), \quad \forall x \in X, \forall u \geq 0$$

Quedan en

$$0 \leq \frac{-x_1+x_2}{x_1+1} + \overline{u}(x_1^2 + x_2) \quad \forall (x_1, x_2) \geq 0, \overline{u} \geq 0 \text{ fijo.}$$

Fijando $x_2 = 0$ resulta que

$$\left\{0 \leq \frac{-x_1}{x_1+1} + \overline{u}(x_1^2) \quad \forall x_1 \geq 0, \overline{u} \geq 0\right\} \Leftrightarrow \left\{\frac{1}{x_1(x_1+1)} \leq \overline{u} \ \ \forall x_1 \geq 0\right\}$$

hecho que no es verdad.

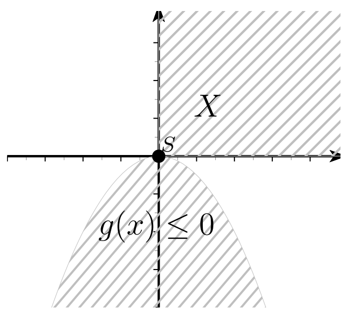


FIGURA 2.1: Región factible

Observación 2.1. Encontrar P.S.K.T o P.S.F.J con $r_0 > 0$ es resolver el problema (PF), pero en la práctica no es fácil resolver el problema de punto de silla.

Condición necesaria

Para la condición necesaria necesitamos convexidad y alguna condición de regularidad. En el Capítulo 1 vimos que en general no hay una regla que nos permita conocer la convexidad de $\frac{f}{h}$ en función de la convexidad de f y h por separado. Si $\frac{f}{h}$ es convexa, podemos usar los teoremas de condiciones de optimalidad generales para programación no lineal (véase Bazaraa, Sherali y Shetty, 2006), pero nos interesa un estudio en función del tipo de convexidad de f y h. Para ello usaremos la transformación de Manas-Schaible.

Teorema 2.2. *Si X es convexo, $f, g, -h$ son convexas en X, ínf$\{f(x) : x \in X\} < 0$ y $\overline{x}$ es solución de (PF). Entonces, existe $(\overline{x}, \overline{r}_0, \overline{r})$ P.S.F.J y ademas $\overline{r}g(\overline{x}) = 0$.*

Demostración.
Usaremos la transformación Manas-Schaible y los resultados vistos en dicho apartado (véase 1.5).
Al ser $\overline{x}$ solución a (PF) tenemos que $(\overline{z}_0, \overline{z})$ es solución a (PMS) donde recordemos que

$$\overline{z}_0 = \tfrac{1}{h(\overline{x})} \qquad \overline{z} = \overline{z}_0\overline{x}$$

Como el problema (PMS), con las condiciones del teorema, es un problema convexo, podemos aplicar el teorema de condición necesaria para el caso no lineal general (véase teorema 5.4.1 de Mangasarian, 1969) para concluir que existe $(\overline{x}, \overline{r}_0, \overline{r}, \overline{s})$ P.S.F.J para (PMS)y por tanto:

$$\overline{r}_0\overline{z}_0 f\left(\tfrac{\overline{z}}{\overline{z}_0}\right) + rg\left(\tfrac{\overline{z}}{\overline{z}_0}\right) + s\left(1 - \overline{z}_0 h\left(\tfrac{\overline{z}}{\overline{z}_0}\right)\right) \leq$$

$$\overline{r}_0\overline{z}_0 f\left(\tfrac{\overline{z}}{\overline{z}_0}\right) + \overline{r}g\left(\tfrac{\overline{z}}{\overline{z}_0}\right) + \overline{s}\left(1 - \overline{z}_0 h\left(\tfrac{\overline{z}}{\overline{z}_0}\right)\right) \leq$$

$$\overline{r}_0 z_0 f\left(\tfrac{z}{z_0}\right) + \overline{r}g\left(\tfrac{z}{z_0}\right) + \overline{s}\left(1 - z_0 h\left(\tfrac{z}{z_0}\right)\right)$$

$$\forall x \in X, \forall(\overline{r}_0, r) \geq 0, (\overline{r}_0, r, s) \neq 0$$

Equivalentemente, aplicando la transformación Manas-Schaible:

$$\overline{r}_0\tfrac{f(\overline{x})}{h(\overline{x})} + rg(\overline{x}) \leq \overline{r}_0\tfrac{f(\overline{x})}{h(\overline{x})} + \overline{r}g(\overline{x}) \leq \overline{r}_0\tfrac{f(x)}{h(x)} + \overline{r}g(x), \quad \forall x \in X, \forall(\overline{r}_0, r) \geq 0, (\overline{r}_0, r) \neq 0$$

Así obtenemos que $(\overline{x}, \overline{r}_0, \overline{r})$ es P.S.F.J, queda ver que $\overline{r}g(\overline{x}) = 0$
Tomando $r = 0$ de la primera desigualdad obtenemos $0 \leq \overline{r}g(\overline{x})$ y por otra parte al ser $\overline{x}$ factible y $\overline{r}$ no negativo $\overline{r}g(\overline{x}) \leq 0$, concluyendo que $\overline{r}g(\overline{x}) = 0$. □

Para poder dar una caracterización del óptimo en función de punto de silla, tenemos que dar una condición necesaria en términos de P.S.K.T y no sólo en términos de P.S.F.J dado que la condición suficiente nos da un P.S.K.T. En la Proposición 2.1 tenemos que bastaría que $\overline{r}_0 > 0$ para tener P.S.K.T, veremos que aquello se obtiene exigiendo alguna condición de regularidad.

Condiciones de regularidad

Aquí veremos algunas condiciones de regularidad, pueden consultarse más condiciones en Cambini y Martein (2009, sección 4.3).

Definición 2.1. *Dado el problema (PF), cualquier condición sobre el conjunto factible que garantice que* $\overline{r}_0 > 0$ *si* $(\overline{x}, \overline{r}_0, \overline{r})$ *es un P.S.F.J, la llamaremos condición de regularidad.*

- Slater:
  $\exists \widetilde{x} \in X : g(\widetilde{x}) < 0$
- Karlin:
  $\nexists p \geq 0, p \neq 0 : pg(x) \geq 0 \quad \forall x \in X$
- Estricta:
  $\exists x_1, x_2$ factibles, $\exists \lambda \in (0,1) : g(\lambda x_1 + (1-\lambda)x_2) < \lambda g(x_1) + (1-\lambda)g(x_2)$

Proposición 2.2.

1. *Slater* $\Leftrightarrow$ *Karlin*
2. *Estricta* $\Rightarrow$ *Slater*

Demostración.

1. Supongamos que no se da la condición de Slater, con lo cual la inecuación $g(x) < 0$ no tiene solución en X. Supongamos g convexa en X convexo, entonces, aplicando Gordan generalizado $\exists p \geq 0, p \neq 0 : pg(x) \leq 0 \quad \forall x \in X$ y por tanto no se tiene la condición de Karlin.
   supongamos que no se da la condición de Karlin, con lo cual $\exists p \geq 0, p \neq 0 : pg(x) \geq 0 \quad \forall x \in X$ con lo cual no se da la condición de Slater.
2. Supongamos que se da la condición Estricta, $\exists x_1, x_2$ factibles, $\exists \lambda \in (0,1) : g(\lambda x_1 + (1-\lambda)x_2) < \lambda g(x_1) + (1-\lambda)g(x_2) \leq 0$, entonces $\widetilde{x} = \lambda x_1 + (1-\lambda)x_2 \in X$ satisface la condición de Slater.

□

Proposición 2.3. *Las condiciones anteriores son de regularidad.*

Demostración.
Sea $(\overline{x}, \overline{r}_0, \overline{r})$ P.S.F.J y supongamos que $\overline{r}_0 = 0$, entonces,
$rg(\overline{x}) \leq \overline{r}g(\overline{x}) \leq \overline{r}g(x) \quad \forall x \in X, \forall r \geq 0, r \neq 0$. Usando Teorema 2.1

$$rg(\overline{x}) \leq 0 \leq \overline{r}g(x) \quad \forall x \in X, \forall r \geq 0, r \neq 0$$

de donde deducimos que no se satisface la condición de Karlin y por Proposición 2.2 todas las condiciones dadas son de regularidad. □

Ahora disponemos de todo lo necesario para enunciar el teorema de la condición necesaria:

Teorema 2.3. *Si X es convexo, $f, g, -h$ son convexas en X, ínf$\{f(x) : x \in X\} < 0$, se satisface alguna condición de regularidad y $\overline{x}$ es solución de (PF). Entonces, existe $(\overline{x}, \overline{u})$ P.S.K.T y ademas $\overline{u}g(\overline{x}) = 0$.*

Demostración.
Basta con usar el Teorema 2.2 y la Proposición 2.1. □

2.2. Condiciones de optimalidad con diferenciabilidad

Veamos las condiciones de optimalidad en el caso de f, h, g sean funciones diferenciables en X. Para ello se usarán los problemas de punto estacionario.

- Punto estacionario de Kuhn-Tucker (P.E.K.T):
  Diremos que $(\overline{x}, \overline{u})$ es un P.E.K.T si se verifica:

$$\bigtriangledown_x \mathcal{K}(\overline{x}, \overline{u}) = 0, \quad \overline{x} \in X, g(\overline{x}) \leq 0, \overline{u} \in \mathbb{R}^m, \overline{u} \geq 0, \overline{u}g(\overline{x}) = 0$$

  Encontrar P.E.K.T equivale a resolver el sistema:

$$\begin{cases} \nabla \frac{f(x)}{h(x)} + \sum\limits_{i=1}^{m} u_i \nabla g_i(x) = 0 \\ x \in X, \quad g(x) \leq 0, \quad u \geq 0 \\ \sum\limits_{i=1}^{m} u_i g_i(x) = 0 \end{cases}$$

- Punto estacionario de Fritz-John (P.E.F.J):
  Diremos que $(\overline{x}, \overline{r}_0, \overline{r})$ es un P.E.F.J si se verifica:

$$\bigtriangledown_x \mathcal{F}(\overline{x}, \overline{r}_0, \overline{r}) = 0, \quad \overline{x} \in X, g(\overline{x}) \leq 0, (\overline{x}, \overline{r}_0, \overline{r}) \geq 0, (\overline{x}, \overline{r}_0, \overline{r}) \neq 0, \overline{r}g(\overline{x}) = 0$$

  Encontrar P.E.F.J equivale a resolver el sistema:

$$\begin{cases} r_0 \nabla \frac{f(x)}{h(x)} + \sum\limits_{i=1}^{m} r_i \nabla g_i(x) = 0 \\ x \in X, \quad g(x) \leq 0, \quad (r_0, r) \geq 0, \quad (r_0, r) \neq 0 \\ \sum\limits_{i=1}^{m} r_i g_i(x) = 0 \end{cases}$$

Observación 2.2. La Proposición 2.1 se verifica también para el caso de punto estacionario.

A continuación veremos algunas definiciones y resultados generales válidos para el caso de optimización no lineal en general, y que podemos encontrar en cualquier libro de optimización no lineal, como por ejemplo Bazaraa, Sherali y Shetty (2006)

Definición 2.2 (Dirección de mejora). *Llamaremos dirección de mejora en el punto $\overline{x}$ a cualquier vector no nulo d que satisfaga la condición $\nabla \frac{f(\overline{x})}{h(\overline{x})} d < 0$, equivalentemente $\nabla f(\overline{x})d < \frac{f(\overline{x})}{h(\overline{x})} \nabla h(\overline{x})d$*

Definición 2.3 (Dirección factible). *Llamaremos dirección factible en el punto $\overline{x}$ a cualquier vector no nulo d que satisfaga la condición $\overline{x} + \lambda d \in \mathcal{S}$ para $0 < \lambda < \lambda_0$*

Definición 2.4 (Dirección tangente). *Llamaremos dirección tangente en el punto $\overline{x}$ a cualquier vector no nulo de la forma $d = \lim\limits_{k \to \infty} \lambda_k (x_k - \overline{x})$, donde $\{x_k\} \subset X, \{x_k\} \to \overline{x}$*

Definición 2.5. *Definimos los conjuntos:*
F_0 *conjunto de las direcciones de mejora en* $\overline{x}$
D *conjunto de las direcciones factibles en* $\overline{x}$
T *conjunto de las direcciones tangentes en* $\overline{x}$
$I = \{i : g_i(\overline{x}) = 0, \overline{x} \in X\}$ *conjunto de las restricciones activas*
$G_0 = \{d : \nabla g_i(\overline{x})d < 0 \ \ \forall i \in I\}$
$G = \{d : \nabla g_i(\overline{x})d \leq 0 \ \ \forall i \in I\}$

Proposición 2.4. *Si* $\overline{x}$ *es mínimo local, entonces,* $F_0 \cap D = \emptyset$

Demostración.
Si $d \in F_0 \cap D$, entonces $\nabla \frac{f(\overline{x})}{h(\overline{x})} d < 0$, y por otra parte, para cierto $0 < \lambda < \lambda_0$ suficientemente pequeño $\overline{x} + \lambda d \in X$, entonces:

$$\frac{f(\overline{x}+\lambda d)}{h(\overline{x}+\lambda d)} = \frac{f(\overline{x})}{h(\overline{x})} + \lambda \underbrace{\nabla \frac{f(\overline{x})}{h(\overline{x})} d}_{<0} + \alpha(\overline{x}, \lambda d)\|\lambda d\| \text{ con } \alpha(\overline{x}, \lambda d) \underset{\lambda \to 0}{\longrightarrow} 0$$

De donde $\frac{f(\overline{x}+\lambda d)}{h(\overline{x}+\lambda d)} < \frac{f(\overline{x})}{h(\overline{x})}$ para cierto $0 < \lambda$ suficientemente pequeño, lo que contradice la hipótesis. □

Proposición 2.5. *Si* $\overline{x}$ *es mínimo local, entonces,* $F_0 \cap T = \emptyset$

Demostración.
Sea $d \in T, \quad d = \text{lím}\lambda_k(x_k - \overline{x}), \quad \lambda_k > 0, \{x_k\} \to \overline{x}$
Siendo f/g diferenciable en $X \Rightarrow \frac{f(x_k)}{h(x_k)} = \frac{f(\overline{x})}{h(\overline{x})} + \nabla \frac{f(\overline{x})}{h(\overline{x})}(x_k - \overline{x}) + \alpha(\overline{x}, x_k - \overline{x})\|x_k - \overline{x}\|$
Fijado $\epsilon > 0$, existe k suficientemente grande tales que $x_x \in B_\epsilon(\overline{x})$ donde $\frac{f(\overline{x})}{h(\overline{x})} \leq \frac{f(x_k)}{h(x_k)} \Rightarrow \nabla \frac{f(\overline{x})}{h(\overline{x})}(x_k - \overline{x}) + \alpha(\overline{x}, x_k - \overline{x})\|x_k - \overline{x}\| \geq 0$, multiplicando por λ_k y haciendo $k \to \infty$ tenemos $\nabla \frac{f(\overline{x})}{h(\overline{x})} \text{lím}\lambda_k(x_k - \overline{x}) \geq 0$, luego $d \notin F_0$ □

Proposición 2.6. *Si* $\overline{x}$ *es mínimo local,* g_i *diferenciables en* $\overline{x}$ $\forall i \in I$ *y continuas en* $\overline{x}$ *con* $\overline{x} \in int(X)$*, entonces,* $F_0 \cap G_0 = \emptyset$

Demostración.
Tomamos $d \in G_0$ y veremos que $d \notin F_0$. Como $\overline{x} \in \text{int}(X) \Rightarrow \overline{x} + \lambda d \in X$ para $0 < \lambda < \lambda_0$. Por otra parte como $d \in G_0$ tenemos que $\nabla g_i(\overline{x})d < 0 \ \ \forall i \in I \quad \Rightarrow g_i(\overline{x} + \lambda d) < g_i(\overline{x}) = 0$ para $0 < \lambda < \lambda_1$, y siendo g_i continuas en $\overline{x}$, para $i \notin I$ tenemos que $g_i(\overline{x}) < 0$, de donde $g_i(\overline{x} + \lambda d)$ para $0 < \lambda < \lambda_2$. Tomando $0 < \lambda \leq \text{mín}\{\lambda, \lambda_1, \lambda_2\}$ tenemos que $\overline{x} + \lambda d \in \mathcal{S}$, luego $d \in D$. Por Proposición 2.4 $F_0 \cap G_0 = \emptyset$ □

Condición suficiente

Nuevamente nos encontramos ante el problema de que el cociente de dos funciones convexas no es por lo general una función convexa, y para la condición suficiente necesitamos que f/h sea pseudoconvexa en un subconjunto abierto de X. Una condición suficiente para la pseudoconvexidad es la convexidad y sí tenemos una transformación del problema (PF) a otro equivalente (PMS) convexo. Así pues, al igual que hicimos en la condición necesaria sin diferenciabilidad, aplicaremos la transformación Manas-Schaible. Otra posibilidad sería usar la tabla 1.5 para obtener pseudoconvexidad para la función objetivo y aplicar entonces la condición suficiente para la optimización no lineal general.

De forma análoga que hicimos para el problema (PF), podemos establecer funciones lagrangianas y problemas de punto estacionario para (PMS).

Teorema 2.4. *Si X es convexo y abierto, $f, -h$ convexas y diferenciables en X,ínf$\{f(x) : x \in X\} < 0$, g_i cuasiconvexas y diferenciables en X, y si $(\overline{x}, \overline{u})$ es solución al problema de P.E.K.T, entonces, $\overline{x}$ es solución a (PF).*

Demostración.
Por una parte, con las hipótesis del teorema obtenemos que el problema (PMS) es un problema donde W es convexo, la función objetivo pseudoconvexa y las funciones g_i cuasiconvexas y diferenciables. Por otra parte, $(\overline{z}_0, \overline{z}, \overline{u}, \overline{v})$ es P.E.K.T para el problema (PMS). Entonces podemos usar la condición suficiente para el caso de problemas no lineales, convexos y con diferenciabilidad (véase teorema 7.2.1 Mangasarian, 1969) para concluir que $(\overline{z}_0, \overline{z})$ es solución a (PMS) y por tanto $\overline{x}$ es solución a (PF) □

Condición necesaria

Lema 2.1. *Si $\overline{x}$ es mínimo local, f, h, g_i diferenciables en $\overline{x}$ $\forall i \in I$, continuas en $\overline{x}$ $\forall i \notin I$, $\overline{x} \in$ int(X). Entonces, $\exists \overline{r}_0 \geq 0$, $\exists \overline{r}_i \geq 0$ $i \in I$, no todos nulos, tales que $\overline{r}_0 \nabla \frac{f(\overline{x})}{h(\overline{x})} + \sum_{i \in I} \overline{r}_i \nabla g_i(\overline{x}) = 0$*

Demostración.
Aplicando Proposición 2.6 tenemos que $F_0 \cap G_0 = \emptyset$ entonces el sistema:

$$\begin{cases} \nabla \frac{f(\overline{x})}{h(\overline{x})} d < 0 \\ \nabla g_i(\overline{x}) d < 0 \quad i \in I \end{cases}$$

No tiene solución. Aplicando Gordan, $\exists \overline{r}_0 \geq 0$, $\exists \overline{r}_i \geq 0$ $i \in I$ (no todos nulos) tales que:

$$\overline{r}_0 \nabla \frac{f(\overline{x})}{h(\overline{x})} + \sum_{i \in I} \overline{r}_i \nabla g_i(\overline{x}) = 0$$

□

Teorema 2.5. *Si $\overline{x}$ es mínimo local, f, h, g_i diferenciables en $\overline{x}$, $i \in I$, g_i continuas en $\overline{x}$, $\overline{x} \in$ int(X). Entonces, $\exists (\overline{r}_0, \overline{r}) \geq 0, (\overline{r}_0, \overline{r}) \neq 0$ tales que $(\overline{x}, \overline{r}_0, \overline{r})$ es P.S.F.J*

Demostración.
Aplicando Lema 2.1, $\exists \overline{r}_0 \geq 0$, $\exists \overline{r}_i \geq 0, i \in I$ no todos nulos tales que

$$\overline{r}_0 \nabla \frac{f(\overline{x})}{h(\overline{x})} + \sum_{i \in I} \overline{r}_i \nabla g_i(\overline{x}) = 0$$

Tomamos $\overline{r}_i = 0$ para $i \notin I$ □

Lema 2.2. *Si $\overline{x}$ es mínimo local, $f, -h$ convexas en X, f, h, g_i diferenciables en $\overline{x}$ $\forall i \in I$, g_i continuas en $\overline{x}$, $\overline{x} \in$ int(X), $\{\nabla g_i(\overline{x}) : i \in I\}$ son linealmente independientes. Entonces, $\exists \overline{u}_i \geq 0, i \in I$ tales que $\nabla \frac{f(\overline{x})}{h(\overline{x})} + \sum\limits_{i \in I} \overline{u}_i \nabla g_i(\overline{x}) = 0$*

Demostración.
Por Lema 2.1 $\exists (\overline{r}_0, \overline{r}) \geq 0, (\overline{r}_0, \overline{r}) \neq 0$, tales que

$$\overline{r}_0 \nabla \frac{f(\overline{x})}{h(\overline{x})} + \sum_{i \in I} \overline{r}_i \nabla g_i(\overline{x}) = 0$$

Ahora tomamos $\overline{u}_i = \frac{\overline{r}_i}{\overline{r}_0}$, pues la condición de independencia lineal de los vectores gradientes es una condición de regularidad, veámoslo:
Si suponemos que $\overline{r}_0 = 0 \Rightarrow \sum\limits_{i \in I} \overline{r}_i \nabla g_i(\overline{x}) = 0$ siendo $\overline{r}_i$ no todos nulos contradice la independencia lineal. □

Teorema 2.6. *Si se dan las condiciones del lema anterior, entonces, $(\overline{x}, \overline{u})$ es un P.E.K.T*

Demostración.
Al igual que hicimos en la prueba de Teorema 2.5 basta definir como $\overline{u}_i = 0$ para $i \notin I$ □

Observación 2.3. Tenemos de este modo caracterizada la existencia de óptimo para (PF) en términos de P.E.K.T en caso con diferenciabilidad y P.S.K.T en caso sin diferenciabilidad. Se trata pues de una generalización de los resultados respectivos para programación no lineal. La dificultad añadida se encontraba en la convexidad de la función objetivo en términos de la convexidad de f y h, para solventar dicho problema, se ha hecho uso de la transformación de Manas-Schaible.

Ejemplo 1.

$$\text{mín}\left\{\frac{(x_1-3)^2+(x_2-2)^2}{x_1+1} : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\right\}$$

Entonces:
$X = \mathbb{R}^2$ convexo.
$f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2$, convexa y diferenciable.
$h(x_1, x_2) = x_1 + 1$, lineal.
$g_1(x_1, x_2) = x_1^2 + x_2^2 - 5$, convexa y diferenciable.
$g_2(x_1, x_2) = x_1 + 2x_2 - 4$, lineal.
Se cumplen por tanto todas las hipótesis para la condición necesaria y suficiente, buscamos por tanto $(\overline{x}, \overline{u})$ P.E.K.T, resolviendo el sistema:

$$\begin{cases} \frac{x_1^2+2x_1-x_2^2+4x_2-19}{x_1^2+2x_1+1} + 2u_1x_1 + 2u_2x_2 = 0 \\ \frac{2x_2-4}{x_1+1} + u_1 + 2u_2 = 0 \\ x_1 \geq 0, x_2 \geq 0 \\ x_1^2 + x_2^2 - 5 \leq 0 \\ x_1 + 2x_2 - 4 \leq 0 \\ u_1 \geq 0, u_2 \geq 0 \\ u_1(x_1^2 + x_2^2 - 5) + u_2(x_1 + 2x_2 - 4) = 0 \end{cases}$$

Resolver dicho sistema, como ya se mencionó anteriormente, no suele ser sencillo. Pero si observamos la función objetivo nos damos cuenta que el mínimo se alcanza para x_1 grande cercana a 3 y x_2 cercana a dos, entonces podemos probar si buscando el punto de corte entre g_1 y g_2 podemos simplificar el sistema y encontrar $\overline{u}$

$$\begin{cases} x_1 \geq 0, x_2 \geq 0 \\ x_1^2 + x_2^2 - 5 = 0 \\ x_1 + 2x_2 - 4 = 0 \end{cases}$$

que tiene por solución $\overline{x} = (2, 1)$
Y nos queda el siguiente sistema en función de u_1, u_2

$$\begin{cases} u_1 \geq 0, u_2 \geq 0 \\ 4u_1 + 2u_2 - \frac{8}{9} = 0 \\ u_1 + 2u_2 - \frac{2}{3} = 0 \end{cases}$$

con solución $\overline{u} = \left(\frac{2}{27}, \frac{8}{27}\right)$
Concluyendo por tanto que $((2, 1), (2/27, 8/27))$ es un P.E.K.T $\Rightarrow \overline{x}$ es el mínimo global y el valor óptimo es $2/3$.

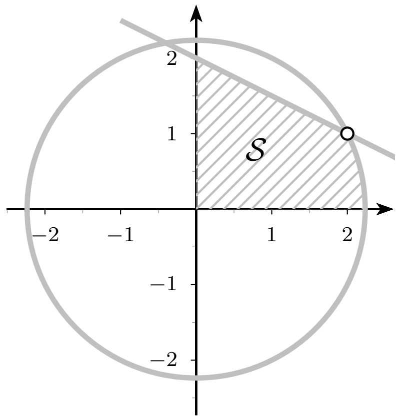


FIGURA 2.2: Región factible.

2.3. Relación entre Optimización Fraccionaria y Paramétrica

En esta sección veremos una relación entre la programación fraccionaria y la paramétrica, que proporciona un soporte teórico a algunos algoritmos iterativos de resolución generales, que veremos en el siguiente capítulo.
El contenido aquí mostrado es una adaptación, al caso de minimizar, del articulo Dinkelbach (1967). También se trata el tema en Minasian (1997, pág 148-154).
Autores como Isbell y Marlow (1956), Schaible (1976c), Dinkelbach (1967) o Jeflea (2003) aplican dicha relación en los algoritmos que exponen en las referencias anteriores.

Sea $\mathcal{S} \subset \mathbb{R}^n$ compacto. Sean $f, h : \mathcal{S} \to \mathbb{R}$ continuas con $h(x) > 0$ en $\mathcal{S}$.
Consideramos los problemas:

$$\text{(PF)} \qquad \text{mín}\left\{\frac{f(x)}{h(x)} : x \in \mathcal{S}\right\}$$

$$(PP)_q \quad \text{mín}\{f(x) - qh(x) : x \in \mathcal{S}\}$$

Lema 2.3. *Definimos $F(q) :=$ mín$\{f(x) - qh(x) : x \in \mathcal{S}\}$, entonces $F(q)$ cumple:*

1. *$F(q)$ es cóncava en $\mathbb{R}$*
2. *$F(q)$ es continua en $\mathbb{R}$*
3. *$F(q)$ es estrictamente decreciente en $\mathbb{R}$*
4. *$F(q) = 0$ tiene solución única*

Demostración.

1. Sean $q, p \in \mathbb{R}, \lambda \in [0,1]$
   $F(\lambda q + (1-\lambda)p) = \text{mín}\,\{f(x) - (\lambda q + (1-\lambda)p)h(x) : x \in \mathcal{S}\} =$
   $= \text{mín}\,\{\lambda f(x) + (1-\lambda)f(x) - \lambda q h(x) - (1-\lambda)ph(x) : x \in \mathcal{S}\} =$
   $= \text{mín}\,\{\lambda(f(x) - qh(x)) + (1-\lambda)(f(x) - ph(x)) : x \in \mathcal{S}\} \geq$
   $\geq \text{mín}\,\{\lambda(f(x) - qh(x))\} + \text{mín}\,\{(1-\lambda)(f(x) - ph(x)) : x \in \mathcal{S}\} =$
   $= \lambda\text{mín}\,\{f(x) - qh(x)\} + (1-\lambda)\text{mín}\,\{f(x) - ph(x) : x \in \mathcal{S}\} = \lambda F(q) + (1-\lambda)F(p)$

2. Definimos $G(q,x) : \mathbb{R}\times\mathcal{S} \to \mathbb{R}$, como $G(q,x) := f(x) - qh(x)$. $F(q) = \underset{x}{\text{mín}}\,\{G(q,x) : x \in \mathcal{S}\}$
   $G(q,x)$ es continua por ser f, h continuas.
   Sean $a, b \in \mathbb{R}$ y consideramos los abiertos $(-\infty, a), (b, \infty)$ que constituyen una sub-base de la topología usual de $\mathbb{R}$. Nuestra labor es ver que $F((-\infty,a))^{-1}$ y $F((b,\infty))^{-1}$ son abiertos de $\mathbb{R}$
   Comencemos con $F((-\infty,a))^{-1}$
   Sea $\pi_1 : \mathbb{R} \times \mathcal{S} \to \mathbb{R}$ proyección, es decir, $\pi_1(q,x) = q$. Es una aplicación continua y abierta.
   $G((-\infty,a))^{-1} = \{(q,x) : G(q,x) < a\} = \{(q,x) : F(q) < a\}$, aplicando la proyección, $\pi_1 \circ G((-\infty,a))^{-1} = \{q : F(q) < a\} = F((-\infty,a))^{-1}$ luego $F((-\infty,a))^{-1}$ es abierto.
   Veamos $F((b,\infty))^{-1}$.
   Para ver que $F((b,\infty))^{-1}$ es abierto, veremos que su complementario es cerrado.
   $\left(F((b,\infty))^{-1}\right)^C = \{q : F(q) \leq b\} = \{q : \exists x$ tales que $f(x) - qh(x) \leq b\} = \pi_1 \circ G((-\infty,b])^{-1}$. Al ser G continua, la preimagen de un cerrado es cerrado, y al ser $\mathcal{S}$ compacto, la proyección $\pi : \mathbb{R} \times \mathcal{S} \to \mathbb{R}$ es cerrada, luego $\left(F((b,\infty))^{-1}\right)^C$ es cerrado como queríamos ver.

3. Sean $q_1 > q_2$
   $F(q_2) = \text{mín}\,\{f(x) - q_2h(x) : x \in \mathcal{S}\} = f(x_2) - q_2h(x_2) > f(x_2) - q_1h(x_2) \geq$
   $\text{mín}\,\{f(x) - q_1h(x) : x \in \mathcal{S}\} = F(q_1)$

4. Al ser F continua usamos Bolzano, pues
$\lim\limits_{q\to-\infty} F(q) = \lim\limits_{q\to-\infty} \min\{f(x) - qh(x) : x \in \mathcal{S}\} = \lim\limits_{q\to-\infty} f(\overline{x}_q) - qh(\overline{x}_q) \geq$
$\lim\limits_{q\to-\infty} f(x_f) - qh(x_h) = \infty$,
donde $f(x_f) = \min\{f(x) : x \in \mathcal{S}\}, \quad h(x_h) = \min\{h(x) : x \in \mathcal{S}\}$
$\lim\limits_{q\to\infty} F(q) = \lim\limits_{q\to\infty} \min\{f(x) - qh(x) : x \in \mathcal{S}\} \leq \lim\limits_{q\to\infty} f(x) - qh(x) = -\infty \quad \forall x \in \mathcal{S}$

Con lo cual $F(q) = 0$ tiene solución, usando que F es estrictamente decreciente, la solución es única.

□

Teorema 2.7. $\overline{q} = \frac{f(\overline{x})}{h(\overline{x})} = \textit{mín}\left\{\frac{f(x)}{h(x)} : x \in \mathcal{S}\right\} \Leftrightarrow F(\overline{q}) = 0$

Demostración.
$\Rightarrow$] Sea $\overline{x}$ solución a (PF), entonces, $\frac{f(\overline{x})}{h(\overline{x})} \leq \frac{f(x)}{h(x)} \quad \forall x \in \mathcal{S}$
$\overline{q} \leq \frac{f(x)}{h(x)} \quad \forall x \in \mathcal{S} \Leftrightarrow f(x) - \overline{q}h(x) \geq 0 \quad \forall x \in \mathcal{S}$
Por otra parte, $\overline{q} = \frac{f(\overline{x})}{h(\overline{x})} \Leftrightarrow f(\overline{x}) - \overline{q}h(\overline{x}) = 0$, luego
$f(x) - \overline{q}h(x) \geq f(\overline{x}) - \overline{q}h(\overline{x}) \quad \forall x \in \mathcal{S} \Rightarrow F(\overline{q}) = \min\{f(x) - \overline{q}h(x) : x \in \mathcal{S}\} = f(\overline{x}) - \overline{q}h(\overline{x}) = 0$

$\Leftarrow$] Ahora supongamos que $F(\overline{q}) = 0$, es decir, $\min\{f(x) - \overline{q}h(x) : x \in \mathcal{S}\} = f(\overline{x}) - \overline{q}h(\overline{x}) = 0 \leq f(x) - \overline{q}h(x) \quad \forall x \in \mathcal{S}$
Queremos ver que $\frac{f(\overline{x})}{h(\overline{x})} \leq \frac{f(x)}{h(x)} \quad \forall x \in \mathcal{S} \Leftrightarrow \overline{q} \leq \frac{f(x)}{h(x)} \Leftrightarrow f(x) - \overline{q}h(x) \geq 0 \quad \forall x \in \mathcal{S}$ hecho que se cumple por hipótesis. □

Entonces, a raíz del teorema anterior, resolver (PF) equivale a hallar el cero de $F(q)$ que sabemos que existe y es único en caso f, h continuas en $\mathcal{S}$ compacto. Si $\overline{q}$ es el cero de $F(\overline{q}) = \min\{f(x) - \overline{q}h(x)\} = 0$ y dicho mínimo se alcanza para $\overline{x}$, entonces, $\overline{x}$ es solución a (PF) y el valor objetivo es $\overline{q}$.
Lo usual es usar métodos numéricos para calcular el cero de $F(q)$ y en cada iteración resolver un problema de optimización no lineal sin fracciones. Véase el ejemplo siguiente.
Ejemplo 2.
Volvemos a usar el ejemplo de la sección anterior:

$$\min\left\{\frac{(x_1-3)^2+(x_2-2)^2}{x_1+1} : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\right\}$$

Que ya sabemos que la solución es $\overline{x} = (2, 1)$, veamos que llegamos a dicha solución buscando el cero de $F(q) = \min\{f(x) - qh(x) : x \in \mathcal{S}\}$
Antes de ver algoritmos realizamos un primer intento de resolución. Usamos el método de la secante para buscar cero de la función:

$$q_{n+1} = q_n - \frac{q_n - q_{n-1}}{F(q_n) - F(q_{n-1})} F(q_n)$$

Tomamos
$q_0 = 0 \Rightarrow F(q_0) = F(0) = \min\{f(x) : x \in \mathcal{S}\} =$

$= \text{mín}\,\{(x_1-3)^2+(x_2-2)^2 : x_1^2+x_2^2 \leq 5, x_1+2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0 \in \mathcal{S}\}$
Que trata de un problema de optimización no lineal a la cual podemos aplicar cualquier algoritmo de resolución existentes para caso no lineal (véase por ejemplo Bazaraa, Sherali y Shetty, 2006). Nosotros aquí seguiremos la linea del ejemplo 2.2 viendo que $((2,1),(1/3,2/3))$ es un P.E.K.T y por tanto la solución es $\overline{x}=(2,1)$ con objetivo $F(0)=2$
$q_1 = 1 \Rightarrow F(q_1) = F(1) = \text{mín}\,\{f(x)-h(x) : x \in \mathcal{S}\} =$
$\text{mín}\,\{(x_1-3)^2+(x_2-2)^2 - x_1 - 1 : x_1^2+x_2^2 \leq 5, x_1+2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0 \in \mathcal{S}\}$
Obteniendo que $((2,1),(2/3,1/3))$ es un P.E.K.T de donde $(2,1)$ es el óptimo y el objetivo $F(1)=-1$
$q_2 = q_1 - \frac{q_1-q_0}{F(q_1)-F(q_0)}F(q_1) = 1 - \frac{-1}{3}(-1) = \frac{2}{3}$
$F\left(\frac{2}{3}\right) = \text{mín}\left\{f(x) - \frac{2}{3}h(x) : x \in \mathcal{S}\right\} =$
$\text{mín}\left\{(x_1-3)^2-(x_2-2)^2 - \frac{2x_1+2}{3} : x_1^2+x_2^2 \leq 5, x_1+2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0 \in \mathcal{S}\right\}$
Donde $((2,1),(5/9,4/9))$ es un P.E.K.T y el objetivo es $F\left(\frac{2}{3}\right)=0$ de donde concluimos que la solución $\overline{x}=(2,1)$ con $\overline{q}=\frac{2}{3}=\frac{f(\overline{x})}{h(\overline{x})}$

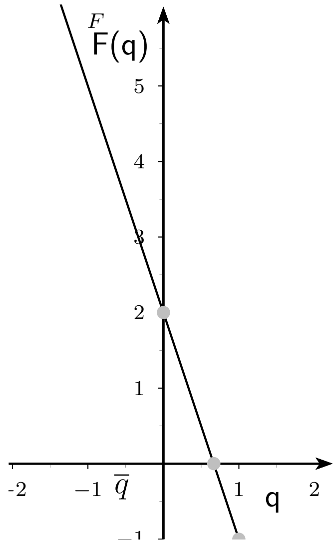


FIGURA 2.3: F(q) interpolada con un conjunto de 9 puntos y spline cúbico
Obteniendo que se aproxima muy bien por la recta F(q)=-3q+2

2.4. Optimización Fraccionaria sobre poliedros

En Optimización lineal donde las restricciones son lineales de la forma $Ax \leq b$, tenemos algoritmo de resolución bastante usado, se trata del método Símplex. Dicho algoritmo busca el óptimo, si este se alcanza, en los vértices del poliedro. En optimización no lineal se tienen resultados sobre la función objetivo, que garantizan, si se alcanza el óptimo, que éste se halla en algún vértice del poliedro. El objetivo de esta sección es particularizar dichos resultados al caso de optimización fraccionaria cuando la región factible es un poliedro.

Nuestro problema (PF) será modelado como mín $\left\{\frac{f(x)}{h(x)} : Ax = b, x \geq 0\right\}$.

Con $\mathcal{S} = \{x : Ax = b, x \geq 0\}$

Definición 2.6 (Poliedro). *Un poliedro es un conjunto de puntos comunes a uno o más semiespacios cerrados, es decir, puntos $x \in \mathbb{R}^n$ tales $Ax \leq b$ con $A \in \mu_{m\times n}(\mathbb{R}), b \in \mathbb{R}^m$.*

Un poliedro F es un subconjunto de $\mathbb{R}^n$ cerrado y convexo.

Definición 2.7 (Punto extremo). *Un punto $\overline{x} \in F$ se dice que es un punto extremo de F si no existen puntos $x_1, x_2 \in F, \lambda \in (0,1)$ tales que $\overline{x} = \lambda x_1 + (1-\lambda)x_2$.*

Teorema 2.8 (Existencia y caracterización de puntos extremos).
Sea $F = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$:

1. *Si $F \neq \emptyset$, entonces, tiene al menos un punto extremo.*
2. *$\overline{x} \in F$ es un punto extremo $\Leftrightarrow$ $A = [B, N]$ donde B es una submatriz invertible con $B^{-1}b \geq 0$ y en cuyo caso $\overline{x} = (\overline{x}_B, \overline{x}_N)$ donde $\overline{x}_B = B^{-1}b, \overline{x}_N = 0$*

Demostración.
Puede encontrarse en cualquier libro sobre optimización lineal. Por ejemplo, véase Teorema 2.1 de Bazaraa, Jarvis y Sherali (2005) para (1) y Bazaraa, Jarvis y Sherali (2005, pág 105-106) para (2). □

Corolario 2.1. *El número de puntos extremos es finito, a lo sumo $\binom{n}{m}$.*

Definición 2.8 (dirección). *Llamamos dirección en F a cualquier vector $d \neq 0$ tales que $x + \lambda d \in F\ \forall x \in F, \forall \lambda \geq 0$*

Definición 2.9 (dirección extrema). *Llamamos dirección extrema en F a cualquier dirección d en F que no se pueda poner como combinación lineal estricta de otras dos direcciones distintas en F, es decir, d es dirección extrema si es dirección y ademas $\nexists d_1, d_2$ direcciones en F tales que $d = \alpha d_1 + \beta d_2$ con $\alpha, \beta > 0$*

La demostración de los siguientes teoremas puede encontrarse en Bazaraa, Jarvis y Sherali (2005).

Teorema 2.9 (Existencia y caracterización de direcciones extremas).
Sea $F = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$:

1. *F es no acotado si y solo si tiene al menos una dirección extrema.*
2. *$\overline{d}$ es una dirección extrema en F $\Leftrightarrow$ $A = [B, N]$ donde B es una submatriz invertible con $B^{-1}a_j \leq 0$ para cualquier columna $a_j \in N$ y en cuyo caso $\overline{d} = \alpha \begin{pmatrix} -B^{-1}a_j \\ e_j \end{pmatrix}, \alpha > 0$*

Teorema 2.10 (Teorema de representación). *Sea $F = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ poliedro. Sean $\{x_1, x_2, \cdots, x_r\}$ puntos extremos de F y sean $\{d_1, d_2, \cdots, d_s\}$ direcciones extremas(si existen) de F. Entonces todo punto de F puede escribirse como:*

$$x = \sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j d_j \qquad \lambda_i \geq 0 \; \forall i, \sum_{i=1}^{r} \lambda_i = 1, \mu_j \geq 0 \; \forall j$$

Definición 2.10. *Diremos que $\overline{x}$ es un vértice mínimo global (VmG) si $\overline{x}$ es un vértice y* $\frac{f(\overline{x})}{h(\overline{x})} \leq \frac{f(x_i)}{h(x_i)} \quad \forall i = 1, \cdots, r.$

Definición 2.11. *Diremos que $\overline{x}$ es un vértice mínimo local (VmL) si $\overline{x}$ es un vértice y* $\frac{f(\overline{x})}{h(\overline{x})} \leq \frac{f(x_i)}{h(x_i)} \quad \forall x_i$ *vértice adyacente.*

La cuestión que nos ocupa ahora es : ¿cuándo un VmG es un mínimo global mG?.
Los siguientes dos importantes resultados son generales, válidos para cualquier función objetivo bien definida en $\mathcal{S}$ poliedro. Y que pretendemos particularizar para el caso de f/h.

Teorema 2.11. *Sea* $\mathcal{S} = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{r} \lambda_i x_i, \lambda_i \geq 0 \; \forall i, \sum_{i=1}^{r} \lambda_i = 1 \right\}$ *, es decir, $\mathcal{S}$ poliedro acotado. Sea f cuasicóncava en S. Entonces, todo vértice mínimo global es mínimo global.*

Demostración.
Al ser f cuasicóncava en $\mathcal{S}$, $\forall x \in \mathcal{S}$, $f(x) \geq \text{mín}\{f(x_i), i = 1, \cdots, r\} = f(\overline{x}_i)$, donde $\overline{x}_i$ es VmG. □

Teorema 2.12. *Sea* $\mathcal{S} = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j d_j, \lambda_i \geq 0 \; \forall i, \sum_{i=1}^{r} \lambda_i = 1, \mu_j \geq 0 \; \forall j \right\}$ *poliedro no acotado. Sea f explícitamente cuasicóncava en $\mathcal{S}$ y se alcanza el mínimo. Entonces, todo VmG es mG.*

Demostración.
Sea $\overline{\mathcal{S}} = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{r} \lambda_i x_i, \lambda_i \geq 0 \; \forall i, \sum_{i=1}^{r} \lambda_i = 1 \right\}$. Sea $\overline{x}$ VmG.
Al ser f explícitamente cuasicóncava, es en particular cuasicóncava y por el resultado anterior $f(\overline{x}) \leq f(x) \;\; \forall x \in \overline{\mathcal{S}}$.
Si el mínimo de f se alcanza en $\overline{\mathcal{S}}$, hemos terminado, sería el mínimo global.
Si el mínimo de F se alcanza en $\mathcal{S} \setminus \overline{\mathcal{S}}$, $\;\exists \widetilde{x}$ donde se alcanza el mínimo.

$$\widetilde{x} = \underbrace{\sum_{i=1}^{r} \lambda_i x_i}_{x_0 \in \overline{\mathcal{S}}} + \sum_{j=1}^{s} \mu_j d_j \quad \text{con } \mu_j \text{ no todos nulos}$$

Consideramos $\widehat{x} = x_0 + 2 \sum_{j=1}^{s} \mu_j d_j \quad \in \mathcal{S}$
Tenemos que $\widetilde{x} = \frac{1}{2} x_0 + \frac{1}{2} \widehat{x}$
Supongamos que $f(\widetilde{x}) < f(\overline{x})$.
Por ser f explícitamente cuasicóncava es en particular estrictamente cuasicóncava y por ende $f(\widetilde{x}) > \text{mín}\{f(x_0), f(\widehat{x})\}$ donde se necesita por definición que $f(x_0) \neq f(\widehat{x})$.
Si $f(x_0) \neq f(\widehat{x})$, $f(\widetilde{x})$ no es el mínimo global de f lo que es contradictorio, y si $f(x_0) = f(\widehat{x})$, tendríamos que $f(\widetilde{x}) \geq \text{mín}\{f(x_0), f(\widehat{x})\} = f(x_0)$ que tampoco es un caso posible pues $x_0 \in \overline{\mathcal{S}}$

y estamos suponiendo que $f(\widetilde{x}) < f(\overline{x})$ contradicción con que $\overline{x}$ es VmG.
Concluyendo que $f(\overline{x}) \leq f(\widetilde{x}) \quad \forall \widetilde{x} \in \mathcal{S}$. Y por tanto VmG es mG como se quería ver.

□

Teorema 2.13. *Consideramos el problema mín* $\left\{\frac{f(x)}{h(x)} : Ax = b, x \geq 0\right\}$ $h(x) > 0$*, con* $\mathcal{S} = \{x : Ax = b, x \geq 0\}$*. Supongamos que se alcanza el mínimo y se tiene alguna de las siguientes situaciones:*

1. $f(x)$ *no positiva en* $\mathcal{S}$ *y* $h(x)$ *cóncava en* $\mathcal{S}$*.*
2. $f(x)$ *cóncava en* $\mathcal{S}$*, no negativa en* $\mathcal{S}$ *y* $h(x)$ *convexa en* $\mathcal{S}$*.*
3. $f(x)$ *cóncava en* $\mathcal{S}$ *y* $h(x)$ *lineal.*
4. $f(x)$ *y* $h(x)$ *son lineales en* $\mathcal{S}$*.*

Entonces todo VmG es mG.

Demostración.
En casos anteriores tenemos que f/h es una función explícitamente cuasicóncava (véase la tabla 1.4) y por tanto le es aplicable los teoremas anteriores. □

En el caso de tener el problema (PF) sobre un poliedro $\mathcal{S}$ y se satisface alguna condición del Teorema 2.13, una posible forma de resolver el problema es hallar todos los vértices del poliedro y evaluar la función objetivo, tomando como solución aquel vértice que proporciona menor valor objetivo. En caso de $\mathcal{S}$ no acotado, es necesario garantizar también que se alcanza el mínimo para asegurar que entonces éste se alcanzará en un vértice.
Entonces, la cuestión ahora es preguntarse cómo podemos obtener todos los vértices de un poliedro. El Teorema 2.8 nos da una caracterización de los puntos extremos en función de bases factibles y una manera de obtener los vértices, poco eficiente, sería considerar todas las posibles bases y de estas tomar las factibles, es decir, si tenemos $A \in \mu_{m\times n}(\mathbb{R})$ y $\mathcal{S} = \{x : Ax = b, x \geq 0\}$, tenemos $\binom{n}{m}$ posibles bases, de las cuales tomamos las factibles que son las que cumplen $B^{-1}b \geq 0$.
Notemos que la forma anterior puede servir cuando n y m son pequeños, pues el número de posibilidades y vértices aumenta considerablemente si crecen n y m. Diversos algoritmos han sidos propuestos para tal propósito, como por ejemplo el algoritmo de Manas y Nedoma (1968), el de Balinski (1961) así como el de Altherr (1975).
Ejemplo 3.

$$\text{Mín}\left\{\frac{3log(x_1)+8log(x_2)}{2x_1+7} : x_1 + x_2 \leq 10, x_1 \geq 1, x_2 \geq 2\right\}$$

La función $f(x) = 3log(x_1) + 8log(x_2)$ es suma de cóncavas por coeficientes positivos y por tanto es cóncava en $\mathcal{S}$.
La función $h(x)$ es lineal.
Entonces podemos aplicar el punto 3 del Teorema 2.13 y el óptimo se encuentra en algún vértice

del poliedro $\mathcal{S} = \{x_1 + x_2 \leq 10, x_1 \geq 1, x_2 \geq 2\}$ que es acotado y por tanto el mínimo se alcanza (véase Figura 2.4).
Nótese que al usar únicamente 2 variables, un vértice es la intersección de dos rectas, luego tenemos $\binom{3}{2} = 3$ posibles vértices.
Añadiendo x_3, x_4, x_5 variables de holgura, tenemos $\binom{5}{3} = 10$ posibles bases de las cuales a lo sumo 3 son bases básicas factibles, que caracterizan los vértices.
En este ejemplo se podrían calcular los vértices como punto de intersección de rectas resolviendo los tres posibles sistemas lineales de dos incógnitas, pero para mostrar un ejemplo, usaremos el algoritmo de Manas-Nedoma.

Usando la técnica de variable artificial obtenemos la siguiente base factible:

	x_1	x_2	x_3	x_4	x_5	$\bar{b}$
x_3	0	0	1	1	1	7
x_1	1	0	0	-1	0	1
x_2	0	1	0	0	-1	2

Entonces $R_1 = \{(3,1,2)\}$ y $W_1 = \{(4,1,2),(5,1,2)\}$
Tomamos entonces x_4, x_1, x_2 como nueva base:

	x_1	x_2	x_3	x_4	x_5	$\bar{b}$
x_4	0	0	1	1	1	7
x_1	1	0	1	0	1	8
x_2	0	1	0	0	-1	2

Así llegamos a que $R_2 = \{(3,1,2),(4,1,2)\}$ y $W_2 = \{(5,1,2)\}$
Y tomaríamos $(5,1,2)$ como nueva base:

	x_1	x_2	x_3	x_4	x_5	$\bar{b}$
x_5	0	0	1	1	1	7
x_1	1	0	0	-1	0	1
x_2	0	1	1	1	0	9

$R_3 = \{(3,1,2),(4,1,2),(5,1,2)\}$ y $W_3 = \emptyset$ y por tanto pararía el algoritmo con R_3 conjunto de bases que corresponden a los vértices, obviando las variables de holgura, $v_1 = (1,2), v_2 = (8,2)$ y $v_3 = (1,9)$
Evaluando la función objetivo en esos puntos: $\frac{f(v_1)}{h(v_1)} \approx 0.6161$, $\frac{f(v_2)}{h(v_2)} \approx 0.5123$ y $\frac{f(v_3)}{h(v_3)} \approx 1.9530$.
Concluyendo así que el mínimo es $\overline{x} = (8,2)$ y el valor objetivo 0.5123.

Ejemplo 4.

$$\text{mín} \left\{ \frac{-4e^{-2x_1} - 5e^{-3x_2}}{log(x_1+x_2)+4} : x_1 + x_2 \geq 3, 2x_1 - x_2 \geq 4, x_1 \geq 0, x_2 \geq 0 \right\}$$

$f(x) = -(4e^{-2x_1} + 5e^{-3x_2})$ no positiva en todo el espacio, no sólo en la región factible.
$h(x) = log(x_1 + x_2) + 4$ es una función cóncava en $\mathcal{S}$.
Y por tanto podemos aplicar el punto 1 del Teorema 2.13.
Añadiendo x_3, x_4 variables de holgura y usando la técnica de variable artificial tenemos la siguiente base factible:

	x_1	x_2	x_3	x_4	$\overline{b}$
x_1	1	1	-1	0	3
x_4	0	-3	-2	1	2

De donde $R_1 = \{(1,4)\}$ y $W_1 = \{(1,2)\}$ y por Teorema 2.9 tenemos que $d_1 = (1,0)$ es dirección extrema.
Pivotamos en $a_{2,2}$ para considerar la base $(1,2)$ que indica W_1

	x_1	x_2	x_3	x_4	$\overline{b}$
x_1	1	0	$-1/3$	$-1/3$	$7/3$
x_2	0	1	$-2/3$	$1/3$	$2/3$

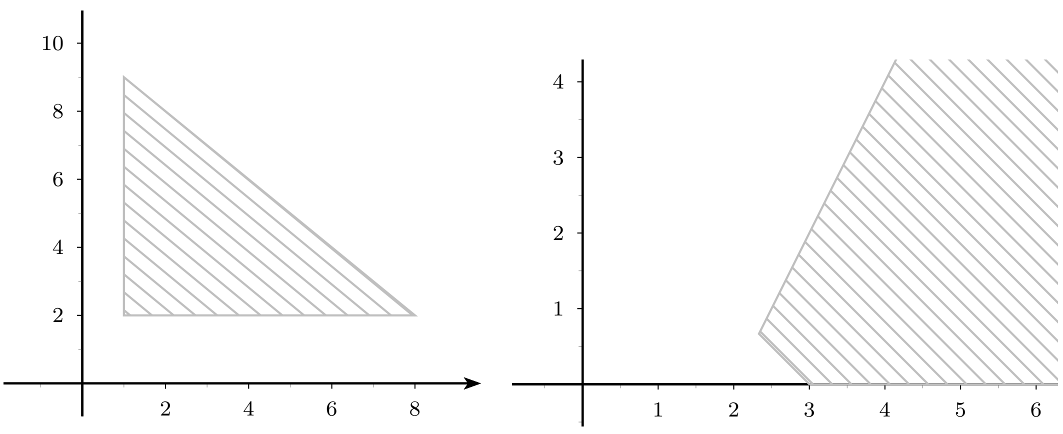
Obteniendo $R_1 = \{(1,4),(1,2)\}$ y $W_1 = \emptyset$. Y otra dirección extrema $d_2 = (1/3, 2/3)$.
Por tanto los puntos extremos, obviando las variables de holgura, son $v_1 = (3,0)$ y $v_2 = (7/3, 2/3)$.
En este ejemplo hemos encontrado direcciones extremas y por tanto el poliedro es no acotado. Debemos por tanto garantizar que el mínimo se alcanza para poder aplicar el teorema.
Una dirección cualquiera en el poliedro es de la forma $t(1,0) + z(1,2)$ con $t, z > 0$. Entonces nos interesa ver que la función objetivo está acotada cuando t y z tienden a infinito.

$$\lim_{(t,z)\to\infty} \frac{-4e^{-2(t+z)} - 5e^{-6z}}{log(t+3z)+4} = 0$$

Pues al ser t y z estrictamente positivas se tiene que:

$$\left| \frac{-4e^{-2(t+z)} - 5e^{-6z}}{log(t+3z)+4} \right| \leq \frac{9}{\mathbf{log}(t+3z)+4} \underset{t,z\to\infty}{\longrightarrow} 0$$

Concluyendo así que la función objetivo está acotada en $\mathcal{S}$ y por tanto se alcanza el mínimo.
$f(v_1) \approx -0.9826$.
$f(v_2) \approx -0.1401$.
Y por tanto la solución es $\overline{x} = (3,0)$ y el valor objetivo -0.9826.



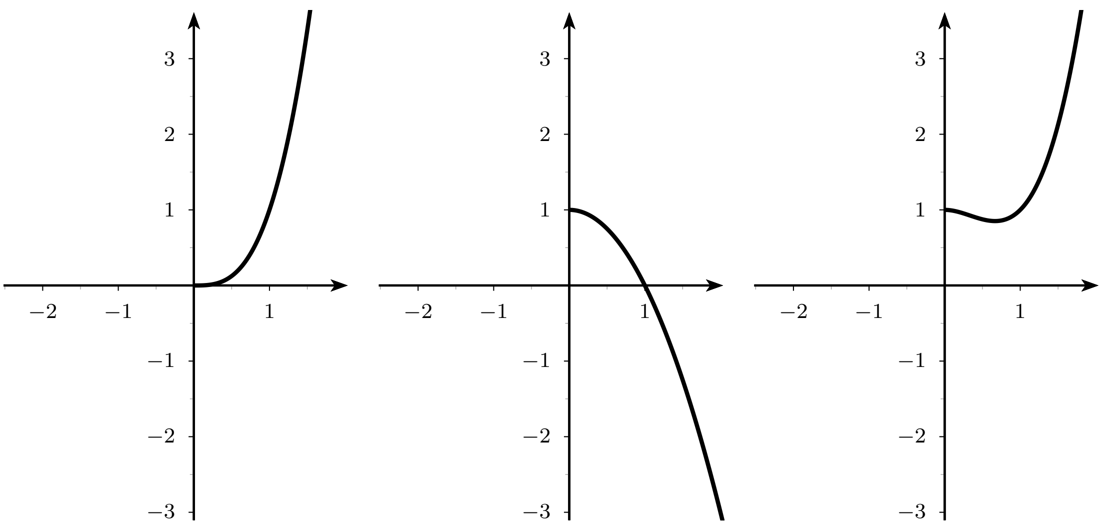
(a) región factible ejemplo 1.

(b) región factible ejemplo 2.

FIGURA 2.4: Región factible.

Algunas observaciones

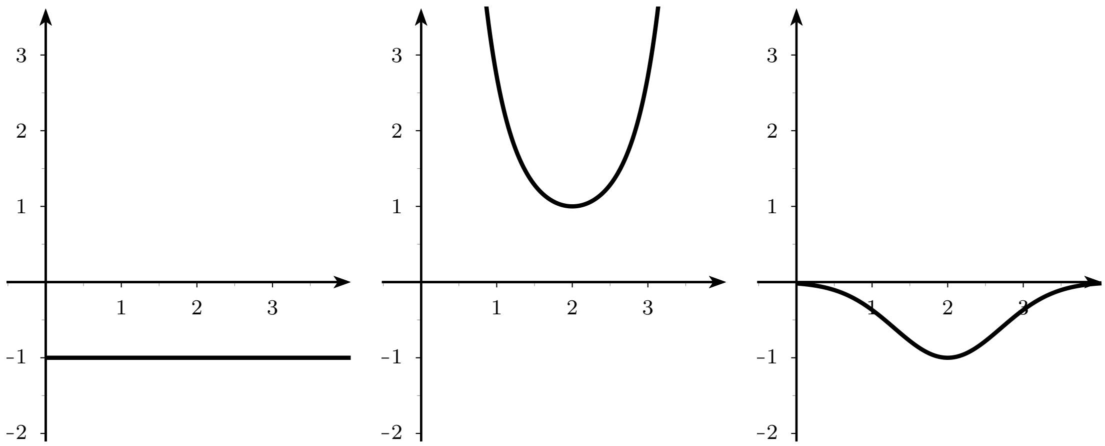
En el teorema central de este apartado hemos necesitado que la función objetivo f/h sea cuasicóncava en el poliedro para garantizar que si el mínimo se alcanza, éste se alcanzará en un vértice del poliedro. Pero las hipótesis que hemos impuesto sobre f y h respecto de tipo convexidad y signo, implicaban que f/h sea explícitamente cuasicóncava.
Nos podemos preguntar si podemos rebajar las hipótesis, por ejemplo ¿bastaría que f sea cuasicóncava y h cuasiconvexa? ó ¿si h no es lineal y f no conserva el signo en $\mathcal{S}$, se podría garantizar que VmG es mG?
La primera cuestión surge por que si fuera cierto que suma de cuasicóncavas diese una función cuasicóncava y f no negativa en $\mathcal{S}$ y supongamos que se alcanza el mínimo, entonces usando la relación entre Optimización Fraccionaria y Paramétrica, el cero de $F(q)$ se obtiene para $\overline{q} > 0$, luego, usando continuidad de f y h existe un subconjunto donde $f(x)$ es cuasicóncava y $-qh(x)$ también es cuasicóncava y tendríamos que $f(x) - qh(x)$ es cuasicóncava, y por tanto $\min\{f(x) - qh(x) : x \in \mathcal{S}\}$ alcanzaría el mínimo para cada $q > 0$ en un vértice y para $\overline{q}$ también en un vértice y tendríamos entonces que f/h alcanza el mínimo en un vértice. Pero por lo general suma de cuasicóncavas no es cuasicóncava, veamos un ejemplo:
Tómese $f(x) = x^3$ en $[-1, 1]$, es una función creciente y por tanto cuasicóncava.
Tómese $h(x) = 1 - x^2$ en $[-1, 1]$, es una función cóncava y por tanto cuasicóncava.
Si consideramos la suma $z(x) = x^3 - x^2 + 1$ no es cuasicóncava en $[-1, 1]$, pues si tomamos los puntos $x_1 = 0$ y $x_2 = 1$ tenemos que $z(0) = z(1) = 1$ y cualquier punto $x_\lambda \in (x_1 : x_2)$ corresponde a $z(x_\lambda) < 1$.



(a) $f(x) = x^3$ (b) $h(x) = 1 - x^2$ (c) $z(x) = x^3 - x^2 + 1$

FIGURA 2.5: Suma de cuasicóncavas no es en general cuasicóncava.

En cuanto a la segunda cuestión, si el cociente entre cóncava y convexa diese una función cuasicóncava, tendríamos entonces hipótesis menos restrictivas en cuanto al cambio de signo de f en el poliedro. Pero en general no es cierto, veamos un ejemplo:
$f(x) = -1$ es una función lineal y por tanto cóncava.
$h(x) = e^{(x-2)^2}$ es una traslación de 2 unidades a la derecha de la función e^{x^2} que es una función convexa.
Consideramos $z(x) = \frac{f(x)}{h(x)} = -e^{-(x-2)^2}$ que no es cuasicóncava, pues basta tomar los puntos $x_1 = 1$ y $x_2 = 3$ y cualquier punto del segmento que los une $x_\lambda \in (x_1 : x_2)$ y observar que $z(x_\lambda) <$ mín$\{z(x_1), z(x_2))\}$.



(a) $f(x) = -1$ (b) $h(x) = e^{(x-2)^2}$ (c) $z(x) = -e^{-(x-2)^2}$

FIGURA 2.6: Cóncava/convexa no es en general cuasicóncava.

2.5. Dualidad Lagrangiana

Dado un problema de Optimización, que llamamos "primal", buscamos asociarle otro problema que llamamos "dual". Y la relación que existe entre ambos es usada para buscar el óptimo de dichos problemas. Los llamados teoremas de dualidad débil y fuerte muestran dicha relación. Veamos primero las definiciones y algunas propiedades de la Dualidad Lagrangiana para el caso no lineal general.
Consideramos $X \subset \mathbb{R}^n, f : X \to \mathbb{R}, g : X \to \mathbb{R}^m, l : X \to \mathbb{R}^k$

Problema primal (P) Mín$\{f(x) : g(x) \leq 0, l(x) = 0, x \in X\}$

Problema dual (D) Máx$\{\theta(u,v) = \operatorname*{Inf}_{x \in X}\{f(x) + ug(x) + vl(x)\} : u \in \mathbb{R}^m, u \geq 0, v \in \mathbb{R}^k\}$

Citamos los dos teoremas fundamentales que relacionan dichos problemas:
Teorema débil.[Bazaraa, Sherali y Shetty (2006, Teorema 6.2.1)] Sea x solución factible al problema (P), sea (u, v) solución factible al problema (D), entonces:

$$f(x) \geq \theta(u, v)$$

Teorema fuerte.[Bazaraa, Sherali y Shetty (2006, Teorema 6.2.4)] Sea X convexo, f, g convexos, l lineal. Supongamos que se satisface alguna condición de regularidad. Entonces,

$$\text{ínf}\{f(x) : g(x) \leq 0, l(x) = 0, x \in X\} = \sup\{\theta(u,v) : u \in \mathbb{R}^m, u \geq 0, v \in \mathbb{R}^k\}$$

Además si $\overline{x}$ es la solución alcanzada para el primal y $(\overline{u}, \overline{v})$ la solución alcanzada para el dual, se cumplen $\overline{u}g(\overline{x}) = 0$.
Entonces, si un problema tiene solución finita, entonces el otro problema también tiene solución finita y ademas los valores objetivos coinciden. Dicha relación es bastante útil en la resolución de diversos problemas donde el dual presenta mejores propiedades que el primal desde un punto de vista algorítmico.
La función θ, considerando X compacto, f, g, l continuas en X, resulta ser cóncava Bazaraa, Sherali y Shetty (2006, Teorema 6.3.1) y diferenciable Bazaraa, Sherali y Shetty (2006, Teorema 6.3.3).
La definición de dualidad en optimización no lineal no tiene una aplicación directa al caso de optimización fraccionaria, y los teoremas de dualidad tanto la débil como la fuerte fallan, veamos un ejemplo extraído de Minasian (1997, pág 163):

$$\text{(P)} \qquad \text{mín}\left\{\tfrac{1}{x} : x > 0, 1 \leq x \leq 2\right\}$$

$$\text{(D)} \qquad \max_{u}\left\{\inf_{x}\left\{\tfrac{1}{x} + u_1(1-x) + u_2(x-2) : x > 0\right\} : u_1, u_2 \geq 0\right\}$$

Donde el problema primal (P) tiene solución finita en $\overline{x} = 2$ y valor objetivo $1/2$, sin embargo, el problema dual (D) es no acotado.
El hecho anterior obliga a desarrollar nuevas definiciones del problema dual para el caso fraccionario, quizás la forma más natural de hacerlo es transformar el problema fraccionario a un problema no lineal no fraccionaria al cual aplicar las definiciones de dualidad. En dicha linea la transformación de Manas-Schaible puede aplicarse.
El contenido aquí desarrollado sigue de las referencias Schaible (1976a), Schaible (1976b),Kydland (1972), Minasian (1997, capítulo 5) y Bazaraa, Sherali y Shetty (2006, capítlo 6).

Caso sin diferenciabilidad

Consideramos nuevamente el problema de la optimización fraccionaria:
mín $\left\{\frac{f(x)}{h(x)} : g(x) \leq 0, x \in X\right\}$, donde $X \subset \mathbb{R}^n$ no vacío, $f, h : X \to \mathbb{R}$ funciones reales de variable real, $g : X \to \mathbb{R}^m$ función vectorial de variable real. Supongamos que $h(x) > 0 \ \ \forall x \in X$
.
Como una aplicación directa de definición de problema dual no resulta útil, usaremos la transformación de Manas-Schaible:

$$z_0 = \tfrac{1}{h(x)} \qquad z = xz_0$$

$$\text{(PMS)} \qquad \text{mín}\left\{z_0 f(z/z_0) : g(z/z0) \leq 0, z_0 h(z/z_0) = 1, (z_0, z) \in W\right\}$$

Recordemos que si X es convexo, $f, -h, g$ convexas en X y ínf$\{f(x) : x \in X\} < 0$ o h lineal, tenemos que (PMS) es un problema convexo.
Para aplicar la definición de dualidad Lagrangiana al problema (PF) usaremos el problema no lineal (PMS) con la siguiente transformación equivalente:

$$\text{(PMS)'} \qquad \text{mín}\{z_0 f(z/z_0) : z_0 g(z/z_0) \leq 0, z_0 h(z/z_0) = 1, (z_0, z) \in W\} \tag{2.2}$$

Aplicando ahora la definición de dualidad Lagrangiana a (PMS)' tenemos:

$$\text{(D-PMS)} \quad \underset{(u,v)\geq 0}{máx} \left\{ \underset{(z_0,z)\in W}{ínf} \{z_0 f(z/z_0) + u z_0 g(z/z_0) + v(z_0 h(z/z_0) - 1)\} \right\} \tag{2.3}$$

Deshaciendo la transformación

$$\text{(D-PF)} \qquad \underset{u\geq 0}{máx} \left\{ \underset{x\in X}{ínf} \left\{ \frac{f(x) + ug(x)}{h(x)} \right\} \right\} \tag{2.4}$$

En adelante definimos $\theta(u) = \underset{x\in X}{ínf} \left\{ \frac{f(x)+ug(x)}{h(x)} \right\}$
Donde $u \in \mathbb{R}^m, u \geq 0$ y el producto $ug(x)$ es un producto escalar, donde abusamos de notación y no usamos traspuesto.
Volvamos ahora al ejemplo de la introducción:

$$\text{(P)} \qquad \text{mín} \left\{ \tfrac{1}{x} : x > 0, 1 \leq x \leq 2 \right\}$$

$$\text{(D)} \qquad \underset{u}{\text{máx}} \left\{ \underset{x}{\text{inf}} \left\{ \tfrac{1+u_1(1-x)+u_2(x-2)}{x} : x > 0 \right\} u_1, u_2 \geq 0 \right\} =$$

$$= \underset{u}{\text{máx}} \left\{ \underset{x}{\text{inf}} \left\{ \tfrac{1}{x}(1 + u_1 - 2u_2) + u_2 - u_1 : x > 0 \right\}, u_1, u_2 \geq 0 \right\} =$$

$$= \underset{u}{\text{máx}} \left\{ \tfrac{1}{x}(1 + u_1 - 2u_2) + u_2 - u_1 : \tfrac{-1}{x^2}(1 + u_1 - 2u_2) = 0, x > 0, u_1, u_2 \geq 0 \right\}$$

Obteniendo así que $u_2 = \frac{1}{2}u_1 + \frac{1}{2}$ y que el problema dual es equivalente a:

$$\text{máx}\{u_2 - u_1 : u_1, u_2 \geq 0\} = \text{máx} \left\{ -\tfrac{1}{2}u_1 + \tfrac{1}{2} : u_1 \geq 0 \right\} = \tfrac{1}{2} \text{ ,con solución } \overline{u} = (0, 1/2)$$

Nótese que los valores objetivos coinciden en el primal y en el dual.

Teorema 2.14 (Dualidad débil). *Sea x un punto factible del problema (PF) y sea u factible de (D-PF). Entonces:*
$\frac{f(x)}{h(x)} \geq$ *ínf*$\left\{ \frac{f(y)+ug(y)}{h(y)} : y \in X \right\}$
Demostración.

$$\text{ínf} \left\{ \frac{f(y)+ug(y)}{h(y)} : y \in X \right\} \leq \frac{f(y)+ug(y)}{h(y)} = \frac{f(y)}{h(y)} + u\frac{g(y)}{h(y)} \leq \frac{f(y)}{h(y)} \qquad \forall y \in X \cap \{g(y) \leq 0\} \qquad \square$$

Corolario 2.2. *ínf*$\left\{ \frac{f(x)}{h(x)} : g(x) \leq 0, x \in X \right\} \geq$ *sup*$\{\theta(u) : u \geq 0\}$
Es decir, el valor objetivo de cualquier solución factible del dual supone una cota inferior al valor objetivo de cualquier solución factible al problema primal.

Corolario 2.3. *$\frac{f(\overline{x})}{h(\overline{x})} = \theta(\overline{u}) \Rightarrow \overline{x}, \overline{u}$ resuelven el problema primal y dual respectivamente.*

Corolario 2.4. *Sí ínf$\left\{\frac{f(x)}{h(x)} : g(x) \leq 0, x \in X\right\} = -\infty \Rightarrow \theta(u) = -\infty \quad \forall u \geq 0$*

Corolario 2.5. *Sí sup$\{\theta(u) : u \geq 0\} = +\infty \Rightarrow$ (PF) es no acotado.*

Teorema 2.15 (Dualidad fuerte). *Sea X convexo, $f, -h, g$ convexas en X,ínf$\{f(x) : x \in X\} < 0$ o h lineal y supongamos que se da alguna condición de regularidad. Entonces:*

$$\text{ínf}\left\{\frac{f(x)}{h(x)} : g(x) \leq 0, x \in X\right\} = \text{sup}\{\theta(u) : u \geq 0\}$$

Además, si el ínfimo se alcanza en $\overline{x}$ y es finito, entonces, el supremo se alcanza en $\overline{u}$ y es finito, y se cumple que $\overline{u}g(\overline{x}) = 0$

Demostración.
Consideramos la transformación de Manas-Schaible 2.2
ínf$\{z_0 f(z/z_0) : z_0 g(z/z_0) \leq 0, z_0 h(z/z_0) = 1, (z_0, z) \in W\}$
Que bajo las condiciones del teorema es un problema convexo.
Sea $\gamma =$ ínf$\{z_0 f(z/z_0) : z_0 g(z/z_0) \leq 0, z_0 h(z/z_0) = 1, (z_0, z) \in W\}$ y supongamos que $\gamma < \infty$. Si $\gamma = -\infty$, por el Corolario 2.4 sup$\{\theta(u) : u \geq 0\} = -\infty$ y el teorema sería cierto.
Supongamos que γ es finito y consideramos el siguiente sistema:

$$\begin{cases} z_0 f(z/z_0) - \gamma < 0 \\ z_0 g(z/z_0) \leq 0 \\ z_0 h(z/z_0) - 1 = 0 \\ (z_0, z) \in W \end{cases}$$

No tiene solución por la definición de γ. Utilizando Gordan, existen (u_0, u, v) no todos nulos con $(u_0, u) \geq 0$ que podemos suponer que $(u_0, u) > 0$, pues en caso contrario las desigualdades que usaremos quedarán degeneradas en trivialidades, tales que:

$$u_0(z_0 f(z/z_0) - \gamma) + u z_0 g(z/z_0) + v(z_0 h(z/z_0) - 1) \geq 0 \quad \forall (z_0, z) \in W. \tag{2.5}$$

Veamos primero que $u_0 > 0$. Supongamos que no es así y que $u_0 = 0$, usando que se cumple alguna condición de regularidad, por ejemplo Slater, existe $(\widehat{z_0}, \widehat{z}) \in W$ tales que $\widehat{z_0} g(\widehat{z}/\widehat{z_0}) < 0$ y $\widehat{z_0} h(\widehat{z}/\widehat{z_0}) = 1$. Sustituyendo en 2.5 se tiene:
$u \underbrace{\widehat{z_0} g(\widehat{z}/\widehat{z_0})}_{<0} + v \underbrace{(\widehat{z_0} h(\widehat{z}/\widehat{z_0}) - 1)}_{=0} \geq 0 \Rightarrow u \widehat{z_0} g(\widehat{z}/\widehat{z_0}) \geq 0 \Rightarrow u = 0$
Y tendríamos $u_0 = u = 0$ que no es posible, luego $u_0 > 0$ y podemos dividir entonces por u_0:

$$z_0 f(z/z_0) + \overline{u} z_0 g(z/z_0) + \overline{v}(z_0 h(z/z_0) - 1) \geq \gamma \quad \forall (z_0, z) \in W, \text{ donde } \overline{u} = \tfrac{u}{u_0} \text{ y } \overline{v} = \tfrac{v}{u_0}.$$

Deshaciendo la transformación

$$\frac{f(x)+\overline{u}g(x)}{h(x)} \geq \gamma \quad \forall x \in X \Rightarrow \theta(\overline{u}) \geq \gamma \tag{2.6}$$

Y por otra parte, por Corolario 2.2 $\gamma \geq \sup\{\theta(u) : u \geq 0\} \geq \theta(\overline{u})$
Concluyendo así que $\theta(\overline{u}) = \gamma$ y por tanto $\overline{u}$ resuelve el problema dual.
Para completar la prueba, supongamos que $\overline{x}$ es un óptimo del problema primal, esto es, $g(\overline{x}) \leq 0, \overline{x} \in X$ con $\frac{f(\overline{x})}{h(\overline{x})} = \gamma$ y sustituyendo $\overline{x}$ en 2.6:
$\frac{f(\overline{x})}{h(\overline{x})} + \overline{u}\frac{g(\overline{x})}{h(\overline{x})} = \frac{f(\overline{x})}{h(\overline{x})} \Rightarrow \overline{u}g(\overline{x}) = 0$

□

Caso con diferenciabilidad

Consideramos el problema (PF) detallado en el apartado anterior, añadiendo que X es abierto y las funciones f, h, g son diferenciables en X. Veremos una nueva definición del dual también deducida de la transformación de Manas-Schaible, con lo cual necesitaremos las hipótesis de las proposiciones 1.18 o 1.19. También puede definirse el dual usando la relación con el problema paramétrico como puede hallarse en Minasian (1997).
En el caso del teorema de dualidad débil, tenemos el mismo teorema que el caso general sin diferenciabilidad.
Hemos visto en la Ecuación 2.2 que el problema (PF) es equivalente a:

$$\text{(PMS)'} \qquad \text{mín}\{z_0 f(z/z_0) : z_0 g(z/z_0) \leq 0, z_0 h(z/z_0) = 1, z_0 \geq 0, (z_0, z) \in W\}$$

Si h es lineal podemos usar (PMS)' sin más modificaciones pues el problema sería convexo pese a la restricción $z_0 h(z/z_0) = 1$, en caso de ser h no lineal, necesitaríamos, tal y como se vio en Sección 1.5, que ínf$\{f(x) : x \in X\} < 0$, y en cuyo caso, el problema es equivalente a:

$$\text{(PMS)'} \qquad \text{mín}\{z_0 f(z/z_0) : z_0 g(z/z_0) \leq 0, z_0 h(z/z_0) \leq 1, z_0 \geq 0, (z_0, z) \in W\}$$

El ejemplo que queremos mostrar y resolver por el dual es el ejemplo ya visto en Ejemplo 1, en dicho ejemplo se tiene que ínf$\{f(x) : x \in X\} \not< 0$, con lo cual, usando el mismo razonamiento hecho al final de la Sección 1.5, se tiene que el problema (PF) es equivalente a:

$$\text{(PMS)'} \qquad \text{mín}\{z_0 f(z/z_0) : z_0 g(z/z_0) \leq 0, -z_0 h(z/z_0) \leq -1, z_0 \geq 0, (z_0, z) \in W\}$$

El dual de dicho problema es:

$$\text{(D-PMS')} \max_{u,v,w \geq 0} \left\{ \underset{(z_0,z)\in W}{inf} \{z_0 f(z/z_0) + z_0 u g(z/z_0) + v(-z_0 h(z/z_0) + 1) + w(-z_0)\} \right\} \tag{2.7}$$

A raíz del resultado Proposición 1.13, en el ínfimo se anula el gradiente de la función objetivo. Calculando la parcial respecto se z_0 y agrupando términos obtenemos la restricción:

$$f(z/z_0) + ug(z/z_0) - vh(z/z_0) - \frac{z}{z_0}\left(\nabla f(z/z_0) + u\nabla g(z/z_0) - v\nabla h(z/z_0)\right) - w = 0 \tag{2.8}$$

Calculando la parcial respecto se z y agrupando términos obtenemos la restricción:

$$\nabla f(z/z_0) + u\nabla g(z/z_0) - v\nabla h(z/z_0) = 0 \tag{2.9}$$

Usando 2.9 es 2.8, obtenemos las siguientes restricciones:
$f(z/z_0) + ug(z/z_0) - vh(z/z_0) = w \geq 0$
$\nabla f(z/z_0) + u\nabla g(z/z_0) - v\nabla h(z/z_0) = 0$
Entonces, el problema dual queda así:

$$\text{(D-PMS')} \qquad \begin{array}{ll} \text{máx} & v \\ \text{s.a} & \\ & \nabla f(z/z_0) + u\nabla g(z/z_0) - v\nabla h(z/z_0) = 0 \\ & f(z/z_0) + ug(z/z_0) - vh(z/z_0) \geq 0 \\ & u, v \geq 0 \end{array} \tag{2.10}$$

Deshaciendo la transformación, obtenemos el dual de (PF):

$$\text{(D-PF)} \qquad \begin{array}{ll} \text{máx} & v \\ \text{s.a} & \\ & \nabla f(x) + u\nabla g(x) - v\nabla h(x) = 0 \\ & f(x) + ug(x) - vh(x) \geq 0 \\ & u, v \geq 0 \end{array} \tag{2.11}$$

En caso de ínf$\{f(x) : x \in X\} < 0$ el dual tiene la misma definición salvo que $v \leq 0$. Con lo cual, podemos tener una definición para ambos casos, dejando v libre en signo.

Teorema 2.16 (Dualidad fuerte). *Si $\overline{x}$ es solución óptima de (PF), entonces, existen $(\overline{u}, \overline{v})$ tales que $(\overline{x}, \overline{u}, \overline{v})$ es solución óptima de 2.11, y se cumple que $\overline{v} = \frac{f(\overline{x})}{h(\overline{x})}$*

Demostración.
Basta aplicar el teorema de dualidad fuerte para el caso de problema no lineal general al problema equivalente de Transformación de Manas-Schaible. □

Ejemplo 5. Si retomamos el ejemplo 1:

$$\text{mín}\left\{\frac{(x_1-3)^2+(x_2-2)^2}{x_1+1} : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\right\}$$

El dual sería:
máx v
s.a

$$\begin{array}{l} 2(x_1 - 3) - v + 2u_1x_1 + u_2 = 0 \\ x_2 - 2 + u_1x_2 + u_2 = 0 \\ (x_1 - 3)^2 + (x_2 - 2)^2 - v(x_1 + 1) + u_1(x_1^2 + x_2^2 - 5) + u_2(x_1 + 2x_2 - 4) \geq 0 \\ x_1, x_2, u_1, u_2, v \geq 0 \end{array}$$

Que podríamos resolver con cualquier algoritmo de optimización no lineal obteniendo como objetivo $\overline{v} = 2/3$, $\overline{x} = (2, 1)$, $\overline{u} = (11/20, 9/20)$.

Capítulo 3

Métodos de Optimización Fraccionaria

En la literatura podemos encontrar diversos algoritmos de resolución de problemas de Optimización Fraccionaria, la mayoría de estos se enfocan a problemas particulares, como son los algoritmos de Charnes y Cooper (1962), Isbell y Marlow (1956), Martos (1964), Gilmore y Gomory (1963), Ghadle y Pawar (2015) para el caso de Optimizacion Fraccionaria Lineal. Para el caso lineal entera, tenemos Rajendra (1993), C.R y Tibekar (1980), entre otros. Benson (2006) propuso algoritmo para el caso cuadrático.
Para el caso más general, el caso de Optimización Fraccionaria no Lineal, la idea central es aplicar algoritmos de Optimización no Lineal general bien directamente al problema fraccionario ó aplicando alguna transformación que permita llevar el problema a otro equivalente no fraccionario. El algoritmo central, base de otros algoritmos como el de C.R y Tibekar (1980) ya citado, es el algoritmo de Dinkelbach (1967), que se basa en la relación con el problema paramétrico.
Este tipo de algoritmos, que pasan de Optimización Fraccionaria no Lineal a Optimización no Lineal son iterativos, en cada iteración se resuelve un problema no lineal, y a su vez, los algoritmos generales para optimización no lineal son iterativos, con lo cual el éxito de estos algoritmos está muy sujeto a cómo son los subproblemas no lineales en cada iteración.
El algoritmo de Dinkelbach transforma un problema fraccionario no lineal en resolución de subproblemas de optimización no lineal, en caso de Optimización Fraccionaria Lineal, los subproblemas son lineales y en caso de Optimización Cuadrática, los subproblemas también son cuadráticos pero ya la función objetivo no es cociente de dos. Por tanto es de amplia aplicación dependiendo siempre de los subproblemas y su complejidad.

En este capítulo daremos algunos algoritmos de los citados, junto con ejemplos.

Estructura del capítulo.
Sección 3.1. Se usa la relación entre Optimización Fraccionaria y Paramétrica para introducir el algoritmo de Dinkilbach junto con el método de la secante como métodos de resolución.
Sección 3.2. Se recuerda la definición del problema dual, que deja de ser un problema de Optimización Fraccionaria y por tanto se puede resolver con los métodos existentes en Optimización no Lineal general.
Sección 3.3. Se recuerda una posible metodología de resolución en caso de que la región factible sea un poliedro y la función objetivo sea cuasicóncava o explícitamente cuasicóncava.

Sección 3.4. Esta sección esta dedicada a la Optimización Fraccionaria Lineal, se dan resultados teóricos así como algoritmos de resolución.

3.1. Métodos paramétricos

Consideramos el problema de Optimización Fraccionaria:

$$\text{(PF)} \qquad \text{mín}\left\{\frac{f(x)}{h(x)} : g(x) \leq 0, x \in X\right\} \tag{3.1}$$

Donde $X \subset \mathbb{R}^n, f, h : X \to \mathbb{R}, \quad g : X \to \mathbb{R}^m$. Consideramos $h(x) > 0$ en todo X
Tomaremos $\mathcal{S} = \{x \in \mathbb{R}^n : g(x) \leq 0, x \in X\}$.
Para garantizar existencia de mínimo global, supongamos $\mathcal{S}$ compacto, f, h continuas.
Consideramos el problema paramétrico:

$$(PP)_q \qquad F(q) = \text{mín}\{f(x) - qh(x) : x \in \mathcal{S}\} \tag{3.2}$$

La función $F : \mathbb{R} \to \mathbb{R}$ está bien definida por las hipótesis consideradas sobre $\mathcal{S}, f, h$.

Teorema 3.1 (Teorema de Jagannathan). *Sea $\overline{x} \in \mathcal{S}, \overline{x}$ es óptimo de (PF) $\Leftrightarrow$ $\overline{x}$ es óptimo de mín$\left\{f(x) - \frac{f(\overline{x})}{h(\overline{x})}h(x) : x \in \mathcal{S}\right\}$*

Demostración.
Por el Teorema 2.7, $\overline{x}$ es óptimo de (PF) $\Leftrightarrow$ $F\left(\frac{f(\overline{x})}{h(\overline{x})}\right) = 0 \Leftrightarrow$ mín $\left\{f(x) - \frac{f(\overline{x})}{h(\overline{x})}h(x) : x \in \mathcal{S}\right\} = 0$
Dicho teorema es el mismo que Teorema 2.7, salvo que nos dice quien es el óptimo del problema mín$\left\{f(x) - \frac{f(\overline{x})}{h(\overline{x})}h(x) : x \in \mathcal{S}\right\}$. □

Algoritmo de Dinkilbach.

El algoritmo de Dinkilbach se basa en buscar el cero de $F(q)$
Paso 1. Tomar $x_1 \in \mathcal{S}$ y definir $q_1 := \frac{f(x_1)}{h(x_1)}$
Hacer k=1.
Paso 2. (Subproblema) resolver $F(q_k) = \text{mín}\{f(x) - q_k h(x) : x \in \mathcal{S}\}$
Sea x_{k+1} una solución óptima.
Paso 3. Si $F(q_k) = 0$ PARAR, x_k es el óptimo,
si no, hacer $q_{k+1} = \frac{f(x_{k+1})}{h(x_{k+1})}$, $k = k + 1$ y volver al paso 2.

Convergencia

Sea $\{x_k\}$ una secuencia de puntos en $\mathcal{S}$, denotamos por

$$q_k = \frac{f(x_k)}{h(x_k)}$$
$$\varphi_k(x) = f(x) - q_k h(x)$$

Lema 3.1. *Si $\varphi_k(x) < 0$ para algún $x \in \mathcal{S}$, entonces $\frac{f(x)}{h(x)} < \frac{f(x_k)}{h(x_k)}$*

Demostración.
$\varphi_k(x) < 0 \Leftrightarrow f(x) - \frac{f(x_k)}{h(x_k)} h(x) < 0 \Leftrightarrow \frac{f(x)}{h(x)} < \frac{f(x_k)}{h(x_k)}$ □

Lema 3.2. *Sea $x_k \in \mathcal{S}$ y sea x_{k+1} óptimo de $F(q_k)$. Si x_k es óptimo de $F(q_k)$, entonces es óptimo de (PF), en otro caso, $\frac{f(x_{k+1})}{h(x_{k+1})} < \frac{f(x_k)}{h(x_k)}$*

Demostración.
Si x_k es óptimo de $F(q_k)$ por Teorema 3.1 es óptimo de (PF). En otro caso, al ser x_{k+1} óptimo de $F(q_k)$ tenemos que

$$\varphi_k(x_{k+1}) < \varphi_k(x_k) = 0$$

Por Lema 3.1 $\frac{f(x_{k+1})}{h(x_{k+1})} < \frac{f(x_k)}{h(x_k)}$. □

Concluimos de los resultados anteriores que el algoritmo de Dinkilbach mejora en cada iteración y que, al tener el óptimo x_k es necesaria la comprobación del paso 3, haciendo una iteración más para ver que $F(q_k) = 0$.

Antes de probar el teorema que nos garantiza la convergencia del algoritmo, necesitamos dar algunos conceptos de algorítmica en Optimización, que se pueden ampliar en Bazaraa, Sherali y Shetty (2006, capítulo 7).

Definición 3.1 (Aplicación algorítmica). *El proceso mediante el cual a partir de un punto x_k se obtiene un nuevo punto x_{k+1} siguiendo unas reglas de un algoritmo se denomina iteración y puede describirse mediante un aplicación algorítmica A. Esta aplicación puede ser de punto a conjunto $(x_{k+1} \in A(x_k))$ o de punto a punto $(x_{k+1} = A(x_k))$.*

Definición 3.2 (Conjunto de soluciones). *Dado un algoritmo, llamamos conjunto de soluciones, Ω, al conjunto de puntos factibles que cumplen la condición de parada del algoritmo.*

En la optimización no lineal general, los algoritmos son iterativos y en muchas ocasiones no se da el óptimo global con exactitud, sino aproximado según cierta tolerancia prefijada. Los puntos factibles que cumplen dicha condición de tolerancia, forman el conjunto de soluciones.

Definición 3.3 (Aplicación algorítmica convergente). *Una aplicación algorítmica A se dice que es convergente si partiendo de cualquier punto inicial $x_1 \in \mathcal{S}$, el límite de cualquier subsucesión convergente de la sucesión de puntos $\{x_k\}$ generada por el algoritmo pertenece a Ω.*

Definición 3.4 (Aplicación algorítmica cerrada). *Sean $\mathcal{S}_1, \mathcal{S}_2$ conjuntos cerrados. Una aplicación $A : \mathcal{S}_1 \to \mathcal{S}_2$ de punto a conjunto se dice que es cerrada en $x \in \mathcal{S}_1$ si para cada* *sucesión $\{x_k\}$ de puntos de $\mathcal{S}_1$ convergente a x y cada sucesión $\{y_k\}$ de puntos de $A(x_k)$ que converge a un punto y, se cumple que $y \in A(x)$.*

Teorema 3.2 (convergencia). *Sea $\mathcal{S} \in \mathbb{R}^n$ conjunto cerrado y $\Omega \in \mathcal{S}$ el conjunto de soluciones. Sea $A : \mathcal{S} \to \mathcal{S}$ una aplicación algorítmica punto a conjunto. Supongamos que la sucesión $\{x_k\}$ generada iterativamente por el algoritmo está contenida en un subconjunto compacto de $\mathcal{S}$ y supongamos que existe una función continua α, denominada función de descenso, tal que $\alpha(y) < \alpha(x)$ si $x \notin \Omega$ e $y \in A(x)$. Si la aplicación A es cerrada sobre el complementario de Ω, entonces o el algoritmo finaliza en un número finito de iteraciones con un punto en Ω o genera una sucesión infinita $\{x_k\}$ tal que:*

1. *Cada subsucesion convergente de $\{x_k\}$ tiene límite en Ω*

2. *$\{\alpha(x_k)\}$ converge a $\alpha(x)$ para algún $x \in \Omega$*

Demostración.
Véase Bazaraa, Sherali y Shetty (2006, Teorema 7.2.3). □

Podemos identificar al algoritmo de Dinkilbach como una aplicación algorítmica punto a conjunto como sigue:

$$D(x_k) = \left\{ \overline{x} \in \mathcal{S} : \overline{x} \text{ óptimo de mín} \left\{ f(x) - \tfrac{f(x_k)}{h(x_k)} h(x) : x \in \mathcal{S} \right\} \right\}$$

Por Lema 3.2 podemos considerar como aplicación de descenso a f/h que es continua por hipótesis.

Teorema 3.3 (Convergencia Dinkilbach). *El algoritmo de Dinkilbach converge en un número finito de iteraciones o genera una sucesión infinita de puntos $\{x_k\}$ que tiene límite en Ω*

Demostración.
Bastaría ver que la aplicación D es cerrada sobre $\mathcal{S} - \Omega$ y aplicar el Teorema 3.2.
Sean $\{x_k\}$ y $\{y_k\}$ secuencia de puntos que satisfacen:
$x_k \in \mathcal{S}$ y $\lim_k x_k = \overline{x} \in \mathcal{S} - \Omega$
$y_k \in D(x_k)$ y $\lim_k y_k = \overline{y}$
Queremos ver que $\overline{y} \in D(\overline{x})$.
Usando que $y_k \in \mathcal{S}$ y $\mathcal{S}$ es compacto, en particular cerrado, $\overline{y} \in \mathcal{S}$.
Definimos $\Gamma(x, y) = f(y) - \frac{f(x)}{h(x)} h(y)$. Sea $\widehat{y} \in D(\overline{x})$, entonces $\widehat{y}$ es solución de
$\text{mín}\left\{ f(x) - \frac{f(\overline{x})}{h(\overline{x})} h(x) : x \in \mathcal{S} \right\}$ y por tanto

$$f(\widehat{y}) - \tfrac{f(\overline{x})}{h(\overline{x})} h(\widehat{y}) \leq f(\overline{y}) - \tfrac{f(\overline{x})}{h(\overline{x})} h(\overline{y})$$
$$\Gamma(\overline{x}, \widehat{y}) \leq \Gamma(\overline{x}, \overline{y})$$

Por hipótesis $y_k \in D(x_k)$ y por tanto y_k resuelve $\text{mín}\left\{ f(x) - \frac{f(\overline{x})}{\overline{x}} h(x) : x \in \mathcal{S} \right\}$, de donde

$$\Gamma(x_k, y_k) \leq \Gamma(x_k, \widehat{y})$$

$\Gamma(x, y)$ es una función continua en $\mathcal{S} \times \mathcal{S}$ al ser f, h continuas, y en particular es continua en $(\overline{x}, \overline{y})$, tomando límite en la expresión anterior y usando continuidad

$$\lim_k \Gamma(x_k, y_k) \leq \lim_k \Gamma(x_k, \widehat{y})$$
$$\Gamma(\overline{x}, \overline{y}) \leq \Gamma(\overline{x}, \widehat{y})$$

luego

$$\Gamma(\overline{x}, \overline{y}) = \Gamma(\overline{x}, \widehat{y})$$

y por tanto $\overline{y} \in D(\overline{x})$ como se quería probar. □

En el ejemplo 2 hemos usado el método de la secante para aproximar el cero de $F(q)$ y resolver el problema de Optimización Fraccionaria. Dicho algoritmo tiene un orden de convergencia $\frac{1+\sqrt{5}}{2}$ y converge "casi" siempre si no se toma un punto inicial muy alejado del cero de la función. También recordamos que la convergencia depende de los subroblemas y su complejidad.
Un algoritmo de orden 2 de convergencia es el método de Newton, pero dicho método requiere de la derivada de $F(q)$. Suponiendo que f y h son diferenciables en el interior de $\mathcal{S}$ y aplicando el teorema de la función implícita a la función $G(q, x) = \nabla f(x) - q\nabla h(x)$, deducimos que $F(q)$ es derivable pero sin obtener explícitamente la derivada, con lo cual no es aplicable Newton al problema.
Lo que sí que podemos aplicar para agilizar el algoritmo es truncamiento de soluciones o interpolación para acercarse al cero de $F(q)$.
Llegados a este punto, disponemos de dos mecanismos para aproximar el cero de $F(q)$, bien el algoritmo de Dinkilbach o el método de la secante. Posteriormente haremos una comparativa computacional de ambos sobre el mismo problema, pero de momento veamos el algoritmo de Dinkilbach en acción aplicado al ejemplo 2 ya resuelto por el método de la secante.

Ejemplo 6.

$$\min\left\{\frac{(x_1-3)^2+(x_2-2)^2}{x_1+1} : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\right\}$$

$\mathcal{S} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\}$ es compacto. f, h son continuas. En los subproblemas usaremos solver BARON Sahinidis (2017) para optimización no lineal.

Iteración 1.
Comenzamos con el punto factible $(0, 0)$
$q_1 = \frac{f(0,0)}{h(0,0)} = 13$
$F(13) = \min\{(x_1 - 3)^2 + (x_2 - 2)^2 - 13(x_1 + 1) : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq$

$0\} = -38.41658669$ con minimizador $(2.1881, 0.460653)$

Iteración 2.
$q_2 = \frac{f(2.1881,0.460653)}{h(2.1881,0.460653)} = 0.95002$
$F(0.95002) = \text{mín}\{(x_1-3)^2 + (x_2-2)^2 - 0.95002(x_1+1) : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\} = -0.85006$ con minimizador $(2, 1)$

Iteración 3.
$q_3 = \frac{f(2,1)}{h(2,1)} = 2/3$
$F(2/3) = \text{mín}\{(x_1-3)^2 + (x_2-2)^2 - \frac{2}{3}(x_1+1) : x_1^2 + x_2^2 \leq 5, x_1 + 2x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\} = 4.440892099e - 16$ con lo cual $(2, 1)$ es el minimizador y el óptimo $2/3$.

3.2. Métodos duales.

Hemos visto en la Sección 2.5 la definición del dual para el problema (PF) bajo ciertas hipótesis. Es interesante el caso de f, h, g diferenciables en X abierto y convexo, y además $f, -h, g$ convexos en X, pues en tal caso el dual se definía como:

$$
\begin{array}{lll}
\text{(D-PF)} & \text{máx} & v \\
 & \text{s.a} & \\
 & & \nabla f(x) + u\nabla g(x) - v\nabla h(x) = 0 \\
 & & f(x) + ug(x) - vh(x) \geq 0 \\
 & & u \geq 0
\end{array}
$$

Que en caso de $\text{mín}\{f(x) : g(x) \leq 0, x \in X\} < 0$ o $\text{mín}\{f(x) : g(x) \leq 0, x \in X\} \geq 0$ añadimos la restricción $v \leq 0$ o $v \geq 0$ respectivamente.
Dicho problema es un problema no lineal, que en ocasiones puede resultar más sencillo de resolver que el primal. En ocasiones puede resultar la restricción válida $\det(\nabla f(x), \nabla h(x)) = 0$ de ayuda al delimitar más la región factible.
Sea $\overline{v}$ solución óptima de (D-PF), entonces $\overline{v} = \frac{f(\overline{x})}{h(\overline{x})}$ para cierto $\overline{x}$ factible que resulta ser el minimizador de (PF).
Como ejemplo véase el Ejemplo 5 de la Sección 2.5.

3.3. Métodos simpliciales.

Consideramos el problema

$$
\text{(PF)} \quad \text{mín}\left\{\frac{f(x)}{h(x)} : Ax = b, x \geq 0\right\} \tag{3.3}
$$

Con $\mathcal{S} = \{x : Ax = b, x \geq 0\}$ poliedro.
En tal caso, por los resultado Teorema 2.11 y Teorema 2.12, si f/h es cuasicóncava o explícitamente cuasicóncava, según el poliedro es acotado o no respectivamente, el minimizador del problema es un vértice del poliedro.
Para conocer si la función objetivo cumple las hipótesis necesarias, consultamos la tabla 1.4.
Un procedimiento a proseguir entonces, es localizar todos los vértices del poliedro, usando por ejemplo el algoritmo de Manas y Nedoma (1968) y evaluar la función objetivo en dichos puntos.
Usando técnicas de gradiente y dirección de mejora, es posible obtener el óptimo sin evaluar, en la mayoría de los casos, todos los vértices, ésta idea la veremos en el algoritmo de Gilmore y Gomory en la siguiente sección. También veremos el algoritmo de Charnes y Cooper, ambos algoritmos están basados en el algoritmo Símplex para la Optimización Lineal.

3.4. Optimización Fraccionaria Lineal

En esta sección vamos a considerar el problema de Optimización Fraccionaria Lineal, es decir, minimizar el cociente de dos funciones lineales sujeto a restricciones lineales, más en concreto:

$$\text{(PFL)} \qquad \min\left\{ f(x) = \frac{p'x + \alpha}{q'x + \beta} : Ax = b, x \geq 0 \right\} \tag{3.4}$$

donde $p, q \in \mathbb{R}^n,\ A \in \mu_{m\times n}(\mathbb{R}),\ \alpha, \beta \in \mathbb{R}$.
En el caso lineal, puede hallarse una bibliográfica bastante amplia, aquí seguiremos el texto de Bazaraa, Sherali y Shetty (2006, Sección 11.4).

Proposición 3.1. *Sea $f(x) = \frac{p'x+\alpha}{q'x+\beta}$, sea $\mathcal{S}$ convexo y supongamos que $q'x + \beta \neq 0$ en $\mathcal{S}$. Entonces, f es pseudoconvexa y pseudocóncava en $\mathcal{S}$.*

Demostración.
Podemos suponer que $q'x + \beta > 0$ en $\mathcal{S}$, en caso contrario $\exists x_1, x_2 \in \mathcal{S}$ tales que $q'x_1 + \beta < 0$ y $q'x_2 + \beta > 0$, por convexidad de $\mathcal{S}$ existiría un punto $x_\lambda \in \mathcal{S}$ tales que $q'x_\lambda + \beta = 0$, lo que contradice la hipótesis.
Veamos primero que f es pseudoconvexa.
Tomamos $x_1, x_2 \in \mathcal{S}$ tales que $\nabla f(x_1)'(x_2 - x_1) \geq 0$, queremos ver que $f(x_2) \geq f(x_1)$.

$$\nabla f(x_1) = \tfrac{(q'x_1+\beta)p-(p'x_1+\alpha)q}{(q'x_1+\beta)^2},$$

como $\nabla f(x_1)'(x_2 - x_1) \geq 0$ y $(q'x_1 + \beta)^2 > 0$, tenemos que

$$\begin{gathered}
[p'(q'x_1 + \beta) - q'(p'x_1 + \alpha)]\,(x_2 - x_1) \geq 0 \Leftrightarrow \\
p'(q'x_1 + \beta)x_2 - q'(p'x_1 + \alpha)x_2 - p'(q'x_1 + \beta)x_1 + q'(p'x_1 + \alpha)x_1 \geq 0 \Leftrightarrow \\
(p'x_2 + \alpha)(q'x_1 + \beta) - (q'x_2 + \beta)(p'x_1 + \alpha) \geq 0 \Leftrightarrow \\
(p'x_2 + \alpha)(q'x_1 + \beta) \geq (q'x_2 + \beta)(p'x_1 + \alpha)
\end{gathered}$$

siendo $(q'x_1 + \alpha)(q'x_2 + \beta) > 0$, dividiendo

$$\frac{p'x_2+\alpha}{q'x_2+\beta} \geq \frac{p'x_1+\alpha}{q'x_1+\beta}$$

Para el caso pseudocóncava, tomamos $x_1, x_2 \in \mathcal{S}$ tales que $\nabla f(x_1)'(x_2 - x_1) \leq 0$ y de forma análoga que el caso anterior se llega a que $f(x_2) \leq f(x_1)$. □

Como consecuencias inmediatas del resultado anterior, tenemos que:

- f es explícitamente cuasiconvexa, explícitamente cuasicóncava, estrictamente cuasiconvexa, estrictamente cuasicóncava, cuasiconvexa y cuasicóncava (véase Figura 1.2).
- Todo mínimo local es global, por ser estrictamente cuasiconvexa (véase Proposición 1.6).
- Se satisfacen las hipótesis para las condiciones de Kuhn-Tucker (véase Sección 2.2).
- Si $\mathcal{S}$ es cerrado, entonces es un poliedro y el mínimo global, si se alcanza, este se alcanza en un vértice (véase Teorema 2.12).

Lo citado para problema de minimizar es adaptable a problema de maximizar.

Usando el Método Símplex-Convexo de Zangwill (1967) se deriva el algoritmo de Gilmore y Gomory

Supongamos que tenemos un vértice x caracterizado por una base factible B con $x_B = B^{-1}b > 0,\ x_N = 0$.

El método Símplex-Convexo, en cada iteración aumenta o disminuye una variable no básica, como las variables no básicas están a valor cero y el problema dispone de las restricciones de no negatividad de las variables, tenemos que sólo puede aumentar. Esto da la idea para obtener la dirección de movimiento, consideremos el gradiente reducido

$$r' = \nabla_B f(x)' B^{-1} A - \nabla f(x)'$$
$$r'_N = \nabla_B f(x)' B^{-1} N - \nabla_N f(x)'.$$

Por el teorema 10.5.3 Bazaraa, Sherali y Shetty (2006), se tiene que si $r_N \leq 0$, entonces, (x, r) es P.E.K.T y como se cumplen las hipótesis necesarias, x sería el vértice óptimo.

Sea $r_j = \text{máx}\{r_i : r_i \leq 0, i \in N\}$. La variable no básica x_j aumenta y alguna variable básica es modificada manteniendo factibilidad de la tabla símplex. Esto último se consigue moviéndose a lo largo de una dirección de mejora,$\nabla f(x) d < 0$, hasta el siguiente vértice.

Lema 3.3. *Sea* $f(x) = \frac{p'x+\alpha}{q'x+\beta}$, $\mathcal{S}$ *convexo. Supongamos que* $q'x + \beta \neq 0$ *en* $\mathcal{S}$. *Sea* $x \in \mathcal{S},\ \ d \neq 0$ *tales que* $\nabla f(x)' d < 0$. *Entonces,* $f(x + \lambda d)$ *es función decreciente en* λ.

Demostración.

En las condiciones anteriores d es dirección de mejora de f en x y por tanto $f(x + \lambda d) < f(x), \lambda > 0$ □

Con lo cual, estando en un vértice, esto es encontrar una base básica factible, y teniendo alguna dirección de mejora, el algoritmo avanza por dicha dirección hasta el vértice siguiente y la función decrece por el lema anterior. Cómo sólo existe un número finito de vértices, el algoritmo para en un paso finito, a lo sumo el número total de vértices. Cuando r_n es no positivo, tenemos un P.E.K.T y por tanto se alcanzó el óptimo, y no es necesario continuar actualizando las tablas del símplex en los vértices aún no considerados.

Consideramos el problema min$\left\{ f(x) = \frac{p'x+\alpha}{q'x+\beta} : Ax = b, x \geq 0 \right\}$ y notamos por $\mathcal{S} = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$. Consideramos $\mathcal{S}$ acotado y $q'x + \beta > 0$ en $\mathcal{S}$.

Algorítmo Gilmore y Gomory (1963).

Paso 1. Considerar x_1 vértice factible con la tabla del símplex asociada. Puede aplicarse la técnica de restricción artificial para obtener una base B factible.
Hacer k=1.
Paso 2. Considerar $A = (B, N)$ y calcular el vector

$$r'_N = \nabla_B f(x)' B^{-1} N - \nabla_N f(x)'.$$

Si $r_N \leq 0$ PARAR, x_k es P.E.K.T y por tanto óptimo. Si no, ir a paso 3.
Paso 3. Sea $r_j = \text{máx}\{r_i : r_i \leq 0, i \in N\}$. La variable no básica x_j es seleccionada para entrar en la base, mientras que la variable básica x_i que sale, es seleccionada según el criterio de mínima razón

$$\frac{\bar{b}_i}{y_{ij}} = \text{mín}\left\{ \frac{\bar{b}_i}{y_{ij}} : y_{ij} > 0, i \in B \right\}$$

donde $\bar{b} = B^{-1}b, \quad y_j = B^{-1}a_j$ y a_j es la columna j de A.

Paso 4. Actualizar tabla, hacer k=k+1 e ir a paso 2.

Observación 3.1. En paso 3, si $y_{ij} \leq 0 \ \forall i \Rightarrow d = \begin{pmatrix} -B^{-}a_j \\ e_j \end{pmatrix}$ es dirección factible y el poliedro $\mathcal{S}$ es no acotado, hecho que no se considera por hipótesis. En tal caso hay que garantizar que la función objetivo es acotada, pues puede que no se alcance el mínimo.

Ejemplo 7.

$$\text{mín}\left\{ \frac{-2x_1+x_2+2}{x_1+3x_2+4} : -x_1 + x_2 \leq 4, 2x_1 + 2x_2 \leq 14, x_2 \leq 6, x_1, x_2 \geq 0 \right\}$$

La región factible es acotada y el denominado no se anula en dicha región, por tanto f/h es acotada y el mínimo se alcanza.

$$\nabla f(x_1, x_2) = \left(\frac{-7x_2-10}{(x_1+3x_2+4)^2}, \frac{-7x_1-2}{(x_1+3x_2+4)^2} \right)$$

Añadimos variables de holgura x_3, x_4, x_5

$$\begin{aligned} -x_1 + x_2 + x_3 &= 4 \\ 2x_1 + 2x_2 + x_4 &= 14 \\ x_2 + x_5 &= 6 \end{aligned}$$

y tenemos como punto factible $x_1 = (0, 0, 4, 14, 6)'$ con tabla

Iteración 1.

$\nabla f(x_1)$	$-10/16$	$-2/16$	0	0	0	
	x_1	x_2	x_3	x_4	x_5	$\overline{b}$
0 x_3	-1	0	1	0	0	4
0 x_4	2	1	0	1	0	14
0 x_5	0	1	0	0	1	6
r_j	$10/16$	$2/16$	0	0	0	

$r_j =$ más$\{10/16, 2/16\} = 10/16 \Rightarrow$ entra x_1

$\frac{\overline{b}_i}{y_{i1}} =$ mín$\{14/2\} \Rightarrow$ sale x_4

Iteración 2.

$\nabla f(x_2)$	$-10/121$	$47/121$	0	0	0	
	x_1	x_2	x_3	x_4	x_5	$\overline{b}$
0 x_3	0	$3/2$	1	$1/2$	0	17
$-10/121$ x_1	1	$1/2$	0	$1/2$	0	7
0 x_5	0	1	0	0	1	6
r_j	0	$-47/121$	0	$-5/121$	0	

Como $r_j \leq 0 \Rightarrow \overline{x} = (7, 0)$ es el óptimo.

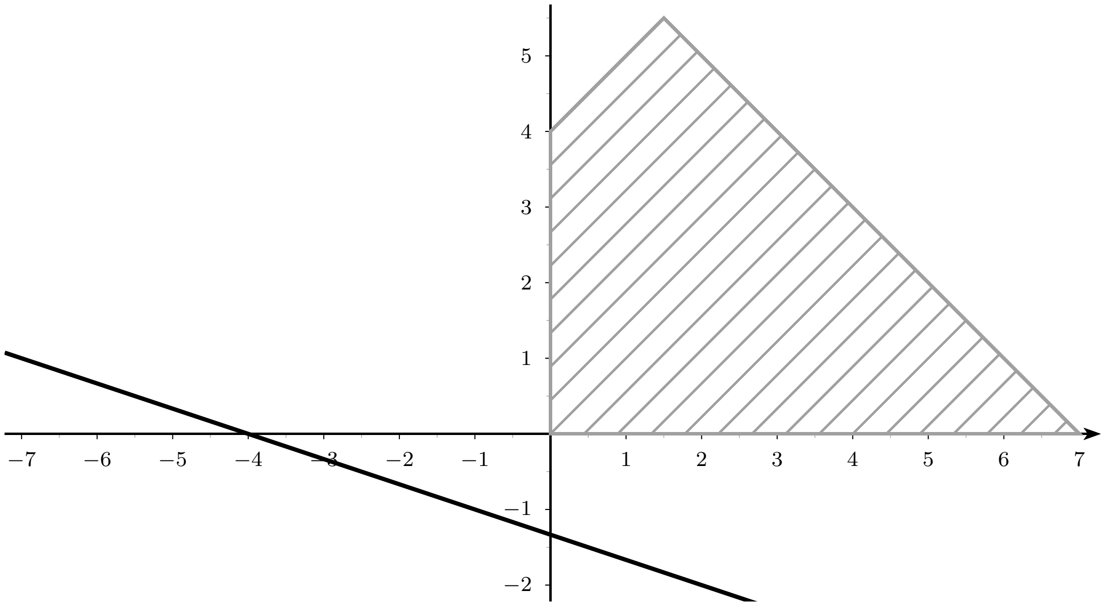


FIGURA 3.1: Región factible (en gris), $x_1 + 3x_2 + 4 = 0$ (en negro)

Algorítmo Charnes y Cooper (1962).

Supongamos el problema

$$\text{mín}\left\{f(x) = \frac{p'x+\alpha}{q'x+\beta} : Ax = b, x \geq 0\right\}$$

Consideramos $\mathcal{S} = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ compacto y $q'x + \beta > 0$ en $\mathcal{S}$.
Consideramos la transformación de Manas-Schaible

$$z_0 = \frac{1}{q'x+\beta}, \qquad z = xz_0$$

el problema sería equivalente a

$$\text{mín}\{p'z + \alpha z_0 : Az - bz_0 \leq 0, q'z + \beta z_0 = 1, z_0, z \geq 0\}$$

que es un problema lineal al cual podemos aplicar cualquier algoritmo de optimización lineal como es el dual-Símplex.
Al considerar $\mathcal{S}$ compacto y función objetivo continua, se alcanza mínimo finito. Por otra parte, la restricción $q'z+\beta z_0 = 1$ hace que el punto $(0, 0)$ no es factible, luego, el punto óptimo $(\overline{z}, \overline{z}_0)$ será con $\overline{z_0} > 0$ y por tanto, por la correspondencia dada en la transformación, el óptimo del problema fraccionario sería $\overline{x} = \overline{z}/\overline{z}_0$.

Ejemplo 8.
Consideramos el Ejemplo 7 ya resuelto por el algoritmo de Gilmore y Gomory, aplicando le ahora el algoritmo de Charnes y Cooper.

$$\text{mín}\left\{\tfrac{-2x_1+x_2+2}{x_1+3x_2+4} : -x_1+x_2 \le 4, 2x_1+2x_2 \le 14, x_2 \le 6, x_1, x_2 \ge 0\right\}$$

que se transforma en:

$$\begin{array}{ll}
\text{min} & -2z_1 + z_2 + 2z_0 \\
\text{s.a} & -z_1 + z_2 - 4z_0 \le 0 \\
& 2z_1 + 2z_2 - 14z_0 \le 0 \\
& z_2 - 6z_0 \le 0 \\
& z_1 + 3z_2 + 4z_0 = 1 \\
& z_1, z_2, z_0 \ge 0
\end{array}$$

Utilizando el algoritmo Dual-Símplex, obtenemos en dos iteraciones la solución de: $z_1 = 63/99, z_2 = 0$ y $z_0 = 9/99$, con lo cual el óptimo es $\overline{x} = (z_1/z_0, z_2/z_0) = (7, 0)$.

Capítulo 4

Implementación del algoritmo de Dinkilbach en Python

En este capítulo mostraremos una posible implementación del algoritmo de Dinkilbach y del método de la secante, para buscar el cero de $F(q)$. Para ello usaremos el lenguaje Python y las librerías Numpy [1] y Scipy[2] que nos permiten resolver los subproblemas. También consideraremos una aplicación directa del método de resolución de problemas no lineales escogido de la Librería Scipy sobre el problema, visto como un problema de Optimización no Lineal en vez de fraccionario. El objetivo de esto es por una parte ver la eficiencia de Dinkilbach sobre el método de la Secante y compararlo con una aplicación directa del optimizador sobre el problema considerado.

Consideramos dichos algoritmos paramétricos por ser bastante generales, aplicables a Optimización Fraccionaria Lineal, no Lineal, Entera y Mixta (aunque de estos dos últimos no hayamos hablado en el texto, como referencia bibliográfica de ello véase Rodenas, Ldpez y Verastegui (1999)), y por su importancia teórica en el caso del algoritmo de Dinkilbach dentro de la teoría de Optimización Fraccionaria. Usar el lenguaje Python nos permite usar el método SLSQP, uno de los algoritmos más exitosos para resolver numéricamente un problema de Optimización no Lineal, de forma gratuita. Dicho algoritmo se encuentra implementado en diversos paquetes de Optimización como KNITRO, NPSOL o como librería de lenguajes como MATLAB, OCTAVE, PYTHON entre otros.
Los algoritmos mostrados a continuación están hechos para el problema de ejemplo 2, pero pueden adaptarse a cualquier problema que cumpla las hipótesis teóricas modificando las definiciones de funciones y restricciones al principio, de hecho en el código fuente 4.3 vemos como modificamos el código para usarlo en otro ejemplo distinto y englobando los tres algoritmos en el mismo programa fuente.

[1] Librería de rutinas algebraicas http://www.numpy.org/
[2] Librería con diversas rutinas, una de ellas optimización no lineal, https://www.scipy.org/

A continuación implementamos el algoritmo de Dinkilbach, importamos las librerías necesarias fundamentalmente Numpy y Scipy. Definimos la función objetivo y restricciones como funciones en python y en la función F(q) llamamos al optimizador. Posteriormente realizamos los pasos del algoritmo de Dinkilbach hasta obtener un valor de F(q) próximo a cero según la tolerancia que hayamos establecido.

```
# -*- coding: cp1252 -*-

#librerias
import numpy as np
from scipy.optimize import minimize
from time import time

x_k = np.array([0.0,0.0])      #punto inicial.
k=0
q_k=0.0
F_k = 100.0
epsilon=0.1e-14

#definimos las funciones
def f(x):
    return (x[0]-3.0)**2 + (x[1]-2.0)**2
def h(x):
    return x[0]+1.0
def g1(x):
    return x[0]**2+x[1]**2-5.0
def g2(x):
    return x[0]+2*x[1]-4.0

#definimos qk
def q(x):
    return 1.0*f(x)/h(x)
q_k = q(x_k)

#resolvemos F(q_k)
def F(q):
    def fun(x):
        return f(x)-q*h(x)
    def restriccion1(x):
        return -g1(x)
    def restriccion2(x):
        return -g2(x)
    b=(0.0,100)
    c=(b,b) #cotas
    res1 = {'type':'ineq', 'fun': restriccion1}
    res2 = {'type':'ineq', 'fun': restriccion2}
    r   = [res1,res2] #restricciones
    sol =
    minimize(fun,x_k,method='SLSQP',bounds=c,constraints=r)
    global x_k
    x_k = sol.x
    global F_k
    F_k = sol.fun

def Dinkilbach(tol):
    while(abs(F_k)>tol):
        global q_k
        F(q_k)
        q_k=q(x_k)
        global k
        k=k+1
t0=time()
Dinkilbach(epsilon)
tf = time()-t0
#salida por pantalla
print "solución=",x_k
print "objetivo=",q_k
print "iteraciones=",k
print "tiempo",tf
```

Código 4.1:

A continuación hacemos la implementación del algoritmo de la Secante.

```
# -*- coding: cp1252 -*-

#librerias
import numpy as np
from scipy.optimize import minimize
from time import time

x_k = np.array([0.0,0.0])  #punto inicial.
k=0
q_0=0.0
q_1=1.0
q_n=0.0
F_k = 100.0
epsilon=0.1e-14
#definimos las funciones
def f(x):
    return (x[0]-3.0)**2 + (x[1]-2.0)**2
def h(x):
    return x[0]+1.0
def g1(x):
    return x[0]**2+x[1]**2-5.0
def g2(x):
    return x[0]+2*x[1]-4.0

#definimos qk
def q(x):
    return 1.0*f(x)/h(x)

#resolvemos F(q_k)
def F(q):
    def fun(x):
        return f(x)-q*h(x)
    def restriccion1(x):
        return -g1(x)
    def restriccion2(x):
        return -g2(x)
    b=(0.0,100)
    c=(b,b) #cotas
    res1 = {'type':'ineq', 'fun': restriccion1}
    res2 = {'type':'ineq', 'fun': restriccion2}
    r    = [res1,res2] #restricciones
    sol  =
    minimize(fun,x_k,method='SLSQP',bounds=c,constraints=r)
    global x_k
    x_k = sol.x
    global F_k
    F_k = sol.fun
    return F_k

def Secante(q_0,q_1):
    return q_1-((q_1-q_0)/(F(q_1)-F(q_0)))*F(q_1)
t0=time()
while(abs(F_k)>epsilon):
    global q_0
    global q_1
    global k
    aux = q_1
    q_1 = Secante(q_0,q_1)
```

```
    q_0 = aux
    k = k+1
tf = time()-t0

#salida por pantalla
print "solución=␣",x_k
print "objetivo=␣",q(x_k)
print "iteraciones=",k
print "tiempo",tf
```

Código 4.2:

Aplicados al problema del ejemplo 2, Dinkilbach realiza tres iteraciones en un tiempo de 0.00399994850159 segundos y el método de la secante realiza dos iteraciones en un tiempo de 0.00400018692017 segundos. Ambos algoritmos actúan de forma similar sobre el problema. Vamos a aplicarlos sobre el siguiente problema más complejo.

$$
\begin{aligned}
\min \quad & \frac{\sum\limits_{i=1}^{n} c_i x_i^2 + \sum\limits_{\substack{i=1\\ i\,\%2=1}}^{n} x_i x_{i+2} - \sum\limits_{\substack{i=1\\ i\,\%2=0}}^{n} x_i x_{i+2}}{\log(1+\sum\limits_{i=1}^{20} t_i x_i)} \\
\text{s.a} \quad & \\
& \sum_{i=1}^{n} x_i = 40 \\
& \sum_{i=1}^{n} x_i^2 = 240 \\
& \sum_{\substack{i=1\\ i\,\%2=0}}^{n} x_i \leq 15 \\
& \sum_{\substack{i=1\\ i\,\%2=1}}^{n} x_i \geq 15 \\
& 0 \leq x_i \leq 100 \quad \forall i
\end{aligned}
\tag{4.1}
$$

A continuación implementamos el algoritmo de Dinkilbach y de la Secante además de una aplicación directa de SLSQP sobre el problema modelo considerado.

```
# -*- coding: cp1252 -*-

#librerias
import numpy as np
from scipy.optimize import minimize
from time import time
import random

n = 200
x_k = np.array([2.0]*n)    #punto inicial.
c = np.array([1.0]*n)
t = np.array([1.0]*n)
k=0
q_k=0.0
F_k = 100.0
epsilon=0.1e-14
for i in range(n):
    c[i] = random.random()
    t[i] = random.random()
#definimos las funciones
def f(x):
    return np.dot(c,x**2)+
    sum(x[i]*x[i+2] for i in range(n-2) if i%2==0)-
    sum(x[i]*x[i+2] for i in range(n-2) if i%2!=0)
def h(x):
    return np.log(1+np.dot(t,x))
#definimos las restricciones
def g1(x):
    return np.dot(np.ones(n),x)-40
def g2(x):
    return np.dot(np.ones(n),x**2)-240
def g3(x):
    return sum(x[i] for i in range(n) if i%2==0)-15
def g4(x):
    return -(sum(x[i] for i in range(n) if i%2!=0)-15)
#definimos qk
def q(x):
    return 1.0*f(x)/h(x)
q_k = q(x_k)

#resolvemos F(q_k)
def F(q):
    def fun(x):
        return f(x)-q*h(x)
    def restriccion1(x):
        return -g1(x)
    def restriccion2(x):
        return -g2(x)
    def restriccion3(x):
        return -g3(x)
```

```
    def restriccion4(x):
        return -g4(x)
    b=(0.0,100)
    c=([b]*n) #cotas
    res1 = {'type':'eq', 'fun': restriccion1}
    res2 = {'type':'eq', 'fun': restriccion2}
    res3 = {'type':'ineq', 'fun': restriccion3}
    res4 = {'type':'ineq', 'fun': restriccion4}

    r    = [res1,res2,res3,res4] #restricciones
    sol  =
    minimize(fun,x_k,method='SLSQP',bounds=c,constraints=r)
    global x_k
    x_k = sol.x
    global F_k
    F_k = sol.fun
    return F_k
def Dinkilbach(tol):
    while(abs(F_k)>tol):
        global q_k
        F(q_k)
        q_k=q(x_k)
        global k
        k=k+1
t0=time()
Dinkilbach(epsilon)
tf = time()-t0
#salida por pantalla para Dinkilbach
print "La solución es X= ",x_k
print "Valor objetivo q= ",q_k
print "iteraciones del algoritmo Dinkilbach = ",k
print "En un timempo de ",tf
print "=========================================================="


#reciclar para aplicar secante.
x_k = np.array([2.0]*n)
k=0
q_k=0.0
F_k = 100.0
q_0=0.0
q_1=1.0
q_n=0.0
def Secante(q_0,q_1):
    return q_1-((q_1-q_0)/(F(q_1)-F(q_0)))*F(q_1)
t0_0=time()
while(abs(F_k)>epsilon):
    global q_0
    global q_1
    global k
    aux = q_1
    q_1 = Secante(q_0,q_1)
    q_0 = aux
    k = k+1
tf_1 = time()-t0_0

#salida por pantalla
print "La solución es X= ",x_k
print "Valor objetivo q= ",q(x_k)
print "iteraciones del algoritmo secante = ",k
print "En un timempo de ",tf_1
print "=========================================================="


#reciclar para aplicar directo.
x_k = np.array([2.0]*n)
def Fun(x):
    def funcion(x):
        return 1.0*f(x)/h(x)
    def restriccion1(x):
        return -g1(x)
    def restriccion2(x):
        return -g2(x)
    def restriccion3(x):
        return -g3(x)
    def restriccion4(x):
        return -g4(x)
    b=(0.0,100)
    c=([b]*n) #cotas
    res1 = {'type':'eq', 'fun': restriccion1}
    res2 = {'type':'eq', 'fun': restriccion2}
    res3 = {'type':'ineq', 'fun': restriccion3}
    res4 = {'type':'ineq', 'fun': restriccion4}

    r    = [res1,res2,res3,res4] #restricciones
    solucion =
    minimize(funcion,x_k,method='SLSQP',bounds=c,constraints=r)
    global x_k
    x_k = solucion.x
    global F_k
    F_k = solucion.fun
    return F_k

t0_00 = time()
Fun(x_k)
t0_11 = time()-t0_00

#salida por pantalla
print "La solución es X= ",x_k
print "Valor objetivo = ",q(x_k)
print "En un timempo de ",t0_11
```

Código 4.3:

Nota: En el código fuente, en donde se llama al optimizador SLSQP con la sentencia minimize(argumentos) podemos añadir opciones para obtener más información del optimizador, como por ejemplo añadir "options={'disp': True,'maxiter': 1000, 'ftol': epsilon}", con *disp* en true obtenemos por pantalla el número de evaluaciones de la función y del gradiente que realiza el algoritmo internamente además del número de iteraciones y mensajes de error, con *maxiter* establecemos el máximo número de iteraciones que por defecto se establecen en 100, y finalmente la opción *ftol* es para establecer la tolerancia. Más información en la documentación del paquete: https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.optimize.fmin_slsqp.html

Los coeficientes c_i y t_i se han obtenido de forma aleatoria, valores reales en $[0, 1)$ pero en cada ejecución se han tomado los mismos valores para los tres métodos.

En las siguientes tablas se muestran los tiempos de ejecución del algoritmo de Dinkilbach, Secante y aplicación directa de SLSQP. Así podemos realizar comparativas modificando el número de variables (modificando el valor de n en el código fuente). Es posible añadir o modificar restricciones definiendo las como funciones.

	n=10				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	0.1559	4	0.3689	5	0.1370
2	0.1239	3	0.3320	5	0.0629
3	0.0939	6	0.2799	6	0.0469
4	0.1089	4	0.2649	5	0.0780
5	0.1619	4	0.3430	4	0.0460
6	0.1410	5	0.2809	5	0.0469
7	0.1410	5	0.3589	5	0.0470
8	0.1870	5	0.3750	5	0.0629
9	0.1089	4	0.3149	4	0.0629
10	0.1089	4	0.3120	5	0.0620
Media	0.1440	4.4	0.3230	4.9	0.0652

	n=20				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	0.2650	4	0.7079	5	0.0970
2	0.3899	4	0.9830	6	0.2339
3	0.3429	4	0.8789	6	0.1089
4	0.5150	4	1.2820	5	0.5899
5	0.3620	4	0.9039	5	0.1250
6	0.4030	6	0.8169	5	0.0780
7	0.4430	4	1.0130	6	0.1089
8	0.3270	4	0.7490	5	0.0780
9	0.2969	4	0.7020	5	0.1089
10	0.3140	5	0.8250	5	0.1800
Media	0.3658	4.3	0.8861	5.3	0.1708

	n=40				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	0.9840	5	1.7200	6	0.4040
2	1.5340	4	1.3120	5	0.9609
3	1.2419	4	2.0980	5	0.9379
4	1.1729	4	1.4370	5	1.5659
5	1.1019	5	3.3019	6	0.7130
6	1.1970	4	2.2749	6	0.5639
7	0.9580	6	2.2810	6	0.1990
8	1.0169	5	1.5869	5	0.7179
9	1.7639	6	3.6280	7	0.8530
10	1.0789	4	2.1230	6	0.3069
Media	1.2048	4.7	2.1759	5.7	0.7222

	n=200				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	11.507	5	19.310	7	5.5199
2	12.578	5	35.717	5	4.1180
3	13.022	4	17.773	6	3.3229
4	12.193	5	25.367	8	2.8480
5	12.003	5	18.970	6	4.0010
6	12.451	5	22.509	6	3.2850
7	11.520	5	20.021	7	5.1029
8	10.055	4	15.218	5	3.8969
9	11.267	5	20.010	5	3.4909
10	6.693	5	30.769	7	3.6229
Media	11.329	4.8	22.566	6.2	3.9114

De los resultados obtenidos concluimos que el Algoritmo de Dinkilbach es más rápido que el algoritmo de la Secante, pero menos rápido que una aplicación directa del método usado para los subproblemas(al menos en el problema modelo considerado). No obstante el trabajo realizado por Rodenas, Ldpez y Verastegui (1999) donde se implementan algoritmos para Optimización Fraccionaria Lineal, Lineal entera y no Lineal usando FORTRAN, muestran que el algoritmo es bastante eficiente para los casos de Optimización Fraccionaria Entera y mixta. Podemos ver un ejemplo de aplicación dentro de la teoría de información en el artículo Zhou y Gao (2017).

Con lo cual un posible trabajo futuro sería usar Dinkilbach sobre una muestra de problemas de Optimización Fraccionaria Entera-Mixta y tener entonces más resultados computacionales para sacar conclusiones viendo distintos tipos de problemas y no sólo el problema anterior.
De hecho, realizamos una segunda prueba sobre un problema donde f convexa no negativa (tomamos forma cuadrática semidefinida positiva) y h cóncava positiva para obtener f/h explícitamente cuasiconvexa y poder ver si llegamos a la conclusión de que Dinkilbach es eficaz en ese tipo de problemas como se afirma en la literatura. Conseguimos f y h con dichas propiedades haciendo una pequeña modificación en el problema anterior 4.1.

$$
\begin{aligned}
\text{min} \quad & \frac{\sum\limits_{i=1}^{n} c_i x_i^2 + \sum\limits_{\substack{i=1 \\ i\,\%2=1}}^{n} x_i x_{i+2} + \sum\limits_{\substack{i=1 \\ i\,\%2=0}}^{n} x_i x_{i+2}}{\log(1+\sum\limits_{i=1}^{20} t_i x_i)} \\
\text{s.a} \quad & \\
& \sum_{i=1}^{n} x_i = 40 \\
& \sum_{i=1}^{n} x_i^2 = 240 \\
& \sum_{\substack{i=1 \\ i\,\%2=0}}^{n} x_i \leq 15 \\
& \sum_{\substack{i=1 \\ i\,\%2=1}}^{n} x_i \geq 15 \qquad 0 \leq x_i \leq 100 \quad \forall i
\end{aligned}
\tag{4.2}
$$

Donde los resultados son:

	n=10				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	0.1059	4	0.3029	5	0.0580
2	0.0939	4	0.2960	5	0.0629
3	0.1239	3	0.3820	5	0.0780
4	0.125	4	0.312	5	0.0469
5	0.078	4	0.25	5	0.047
6	0.1089	5	0.39	5	0.032
7	0.157	4	0.519	5	0.078
8	0.1849	4	0.401	5	0.0779
9	0.0739	3	0.2409	5	0.054
10	0.1459	4	0.378	5	0.0929
Media	0.119	3.9	0.348	5	0.0628

	n=20				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	0.2639	3	0.3359	4	0.3369
2	0.296	3	0.23	4	0.305
3	0.1399	3	0.81	6	0.279
4	0.972	4	1.2219	5	0.907
5	0.3	4	0.7209	5	0.1719
6	0.5739	4	0.5380	6	0.6349
7	0.38	4	1.085	5	0.633
8	0.203	4	0.9709	6	0.788
9	0.296	4	0.9539	5	0.653
10	0.276	4	0.644	5	0.5829
Media	0.412	3.7	0.75105	5.1	0.529

	n=40				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	1.058	4	1.933	5	1.168
2	1.051	4	2.917	5	2.012
3	1.296	4	1.652	5	1.881
4	1.358	5	2.722	7	1.664
5	1.128	4	1.686	5	1.732
6	0.801	4	3.128	5	1.665
7	1.821	4	2.311	5	1.285
8	1.197	4	3.907	4	1.91
9	1.026	4	2.053	4	1.52
10	0.82	4	2.135	4	1.651
Media	1.155	4.1	2.444	4.9	1.645

	n=200				
	Dinkilbach		Secante		SLSQP
	t	k	t	k	t
1	21.832	5	63.845	4	22.725
2	22.899	5	36.404	4	26.013
3	18.843	5	24.083	4	21.06
4	16.956	4	34.452	4	20.052
5	26.177	5	47.891	5	25.88
6	19.87	5	27.555	4	19.722
7	15.654	4	23.819	4	18.193
8	16.311	4	25.238	4	20.309
9	25.581	5	32.385	4	27.111
10	15.701	4	43.052	4	17.326
Media	19.982	4.6	35.919	4.1	21.64

Observamos que efectivamente Dinkilbach obtiene mejores resultados que una aplicación directa al aumentar número de variables en problemas convexo-diferenciables o como es el caso explícitamente cuasiconvexa y diferenciable. Esto se debe en parte a que se realiza un menor número de evaluaciones del gradiente en cada iteración en los subproblemas de Dinkilbach que una aplicación directa de SLSQP (Esto se puede observar con la opción *disp* en true.)

Conclusiones.

Comenzábamos dicho trabajo con dos objetivos, uno el de mostrar contenido teórico y métodos y otro, el de implementar el algoritmo más general considerado en la teoría mostrada.
En cuanto al primer objetivo, concluimos que en el caso lineal se obtienen resultados satisfactorios al ser f/h pseudoconvexa y pseudocóncava lo que hace que se tengan las hipótesis suficientes para las condiciones de optimalidad, para los teoremas de dualidad y se tienen algoritmos eficaces de resolución. Igualmente se obtienen propiedades buenas en el caso f no lineal pero h lineal, pues en tal caso la transformación de Manas-Schaible transforma el problema en uno equivalente convexo y le es aplicable por tanto los resultados vistos.
En el caso f y h ambos no lineales, comenzamos a necesitar , aparte de convexidad, restricciones sobre el signo de las funciones lo que supone en muchas ocasiones restricciones problemáticas desde un punto de vista teórico al aplicar resultados , y por tanto, el investigador en muchas ocasiones se verá obligado a hacer un estudio personalizado al problema que trata para poder garantizar el funcionamiento de los métodos de resolución.

En cuanto al segundo objetivo, el algoritmo más general mostrado, el algoritmo de Dinkilbach, resulta ser eficaz aplicable a diversos tipos de problemas, bien en nuestros problemas 4.1 fue más exitosa una aplicación directa de SLSQP, en los problemas 4.2 Dinkilbach obtuvo mejores resultados al tratarse de problemas con función objetivo cuasiconvexa y diferenciable. Además, Dinkilbach es bastante usado hoy en día en Optimización Fraccionaria descantando sobre todo en Optimización Fraccionaria Entera como podemos leer en Rodenas, Ldpez y Verastegui (1999), y convexo-diferenciable como hemos visto en 4.2 y que podemos ampliar dicha conclusión en You, Castro y Grossmann (2009).
Concluyendo finalmente, que la Optimización Fraccionaria Lineal es fácilmente trasladada a un problema lineal y por tanto eficazmente resoluble. Para el caso no lineal, optimización mixta y para problemas diferenciables y convexos el algoritmo de Dinkilbach es una alternativa eficaz para resolver el problema.

Bibliografía

Aggarwal, S. P. e I. C. Sharma (1970). «Maximization of the transmission rate of a discrete, constant channel». En: *Unternehmensforschung* 14.1, págs. 152-155.

Altherr, W. (1975). «An Algorithm for Enumerating All Vertices of a Convex Polyhedron». En: *Springer-Verlag* 15, págs. 181-193.

Balinski, M. L. (1961). «An Algorithm for finding all vertices of convex polyhedral sets». En: *Journal of the Society for Industrial and Applied Mathematics* 9.1, 72–88.

Bazaraa, Mokhtar S., Hanif D. Sherali y C. M. Shetty (2006). *Nonlinear Programming: Theory and Algorithms*. 3rd. New Jersey: John Wiley & Sons, Inc. ISBN: 978-0-471-48600-8.

Bazaraa, M.S., J.J. Jarvis y H.D. Sherali (2005). *Programación lineal y flujo en redes*. 2rd. México: Limusa. ISBN: 968-18-4867-5.

Benson, Harold P. (2006). «Fractional programming with convex quadratic forms and functions». En: *European Journal of Operational Research* 173.2, págs. 351-369.

Boyd, S.P. y L. Vandenberghe (2004). *Convex Optimization*. Berichte über verteilte messysteme. Cambridge University Press. ISBN: 9780521833783.

Cambini, Alberto y Laura Martein (2009). *Generalized Convexity and Optimization. Theory and Applications*. Heidelberg: Springer. ISBN: 978-3-540-70875-9.

Charles, V. (2007). *E-model for Transportation Problem of Linear Stochastic Fractional Programming*. Inf. téc. SDM Institute for Management Development.

Charnes, A. y W. W. Cooper (1962). «Programming with linear fractional functionals». En: *Naval Research Logistics* 9.3-4, págs. 181-186.

C.R, Seshan y V.G. Tibekar (1980). «Algorithms for integer fractional programming». En: *Indian Institute of Science* 62, págs. 9-16.

C.Singh (1981). «Optimality Conditions in Fractional Programming». En: *Journal of Optimization Theory and Applications* 33.2, págs. 287-294.

Dinkelbach, Werner (1967). «On Nonlinear Fractional Programming». En: *Management Science* 13.7, págs. 492-498.

Frenk, J.B.G. y S. Schaible (2004). *Fractional Programming*. Inf. téc. Erasmus Research Institute of Management.

Ghadle, Kirtiwant P. y Tanaji S. Pawar (2015). «An Alternative Method For Solving Linear Fractional Programming Problems». En: *International Journal of Recent Scientific Research* 6.6, págs. 4418-4420.

Gilmore, P. C. y R. E. Gomory (1963). «A linear Programming Approach to the Cutting Stock problem-Part II». En: *Operations Research* 11.6, págs. 863-888.

Isbell, J. R. y W. H. Marlow (1956). «Attrition games». En: *Naval Research Logistics Quarterly* 3.1-2, págs. 71-94.

Jeflea, Antoneta (2003). «A parametric study for solving nonlinear fractional problems». En: *Analele Stiintifice ale Universitatii Ovidius Constanta* 11.2, 87–92.

Kydland, Finn (1972). «Duality in Fractional Programming». En: *Naval Research Logistics Quarterly* 19, págs. 691-697.

Manas, Miroslav y Josef Nedoma (1968). «Finding All Verticesof a Convex Polyhedron». En: *Numerische Mathematik* 12, págs. 226-229.

Mangasarian, Olvi L. (1969). *Nonlinear Programming*. New York: McGraw Hill. ISBN: 0-89871-341-2.

— (1970). «Convexity, Pseudo-Cconvexity snd Quasi-Cconvexity of Composite Functions». En: *Cahiers du Centre d'Etudes de Recherche Operatio* 12, págs. 114-122.

Martos, Bela (1964). «Hyperbolic programming». En: *Naval Research Logistics* 11.2, págs. 135-155.

Meister, Bernd y Werner Oettli (1967). «On the Capacity of a Discrete, Constant Channel». En: *Information and Control* 11.3, págs. 341-351.

Menger, Karl (1998). *Ergebnisse eines Mathematischen Kolloquiums*. Springer. ISBN: 978-3-7091-7330-5.

Minasian, I.M. Stancu (1997). *Fractional Programming. Theory, Methods and Applications*. Vol. 409. Bucharest: Kluwer Academic Publishers. ISBN: 978-94-01 0-6504-7.

R. A. Fisher Sc.D., F.R.S. (1936). «The use of multiple measurements in taxonomic problems». En: *Annals of human genetics* 7.2, págs. 179-188.

Rajendra, V. (1993). «On integer fractional linear programming». En: *Operational Research Society of India* 30, págs. 174-176.

Reza, Fazlollah M. (1961). *An Introduction to Information Theory*. New York: McGraw-I lill. ISBN: 13-9780486682105.

Rodenas, Ricardo G., M. Luz Ldpez y Doroteo Verastegui (1999). «Extensions of Dinkelbach's Algorithm for Solving Non-linear Fractional Programming Problems». En: *Sociedad de Estadistica e Investigaci6n Operativa* 7.1, págs. 33-70.

Sahinidis, N.V. (2017). *BARON 14.3.1 Global Optimization of Mixed-Integer Nonlinear Programs*. http://www.minlp.com/downloads/docs/baron%20manual.pdf.

Schaible, Siegfried (1976a). «Duality in Fractional Programming: A Unified Approach». En: *Operations Research* 24.3, págs. 452-461.

— (1976b). «Fractional Programming. I, Duality». En: *Management Science* 22.8, págs. 858-867.

— (1976c). «Fractional Programming. II,On Dinkelbach's Algorithm». En: *Management Science* 22.8, págs. 868-873.

Vera, Gabriel (2008). *Lecciones de Análisis Matemático II*. Inf. téc. Universidad de Murcia.

You, Fengqi, Pedro M. Castro e Ignacio E. Grossmann (2009). *Dinkelbach's Algorithm as an Efficient Method for Solving a Class of MINLP Models for Large-Scale Cyclic Scheduling Problems*. Inf. téc. Carnegie Institute of Technology.

Zangwill, Willard I. (1967). «The Convex Simplex Method». En: *Management Science* 14.3, págs. 221-238.

Zhou, Zhenyu y Caixia Gao (2017). «Energy-Efficient Stable Matching for Resource Allocation in Energy Harvesting-Based Device-to-Device Communications». En: *IEEE Access* 5, págs. 15184-15196.

Ziemba, W. T. y R. G. Vickson (1975). *Stochastic Optimization Models in Finance*. Vol. 409. New York: Academic Press, INC. ISBN: 0-12-780850-7.