UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"

Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas

"Método Escalarizado de Punto Proximal Logarítmo Cuadrático para Programación Multiobjetivo"

Trabajo Especial de Grado presentado por

Br.Ana M. Quintero G.

como requisito final
para obtener el título de Licenciado
en Ciencias Matemáticas
Área de Conocimiento: Optimización.
Tutor: Ms.C. Clavel Quintana.

Barquisimeto, Venezuela. Julio de 2014

Universidad Centroccidental
“Lisandro Alvarado”
Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas

ACTA
TRABAJO ESPECIAL DE GRADO

Los suscritos miembros del Jurado designado por el Jefe del Departamento de Matemáticas del Decanato de Ciencias y Tecnología de la Universidad Centroccidental “Lisandro Alvarado”, para examinar y dictar el veredicto sobre el Trabajo Especial de Grado titulado:

“Método Escalarizado de Punto Proximal Logarítmo Cuadrático para Programación Multiobjetivo”

presentado por el ciudadano Br.Ana M. Quintero G. titular de la Cédula de Identidad No. 18844280, con el propósito de cumplir con el requisito académico final para el otorgamiento del título de Licenciado en Ciencias Matemáticas.

Luego de realizada la Defensa y en los términos que imponen los Lineamientos para el Trabajo Especial de Grado de la Licenciatura en Ciencias Matemáticas, se procedió a discutirlo con el interesado habiéndose emitido el veredicto que a continuación se expresa:

[1] ____________________

Con una calificación de ____________________ puntos.

En fe de lo expuesto firmamos la presente Acta en la Ciudad de Barquisimeto a los _____ días del mes de ____________________ de __________.

TUTOR	FIRMA
PRINCIPAL	FIRMA
PRINCIPAL	FIRMA

OBSERVACIONES:

____

____

____

[1] Aprobado ó Reprobado

A Dios Todopoderoso y a todos aquellos que hicieron posible el logro de mi trabajo, en especial a mis padres Anibal y Simona y mi hermana Ana Margarita.

AGRADECIMIENTOS

Primeramente a Dios Todo poderoso por darme la vida y por estar conmigo en todo momento...
A mis padres Anibal y Simona por su apoyo, comprensión. Todo lo que soy es gracias a ustedes. Los amo.
A mi hermana Ana Margarita por su apoyo incondicional.
A mi abuela Petra de Quintero por ser mi confidente.
A mis tias Beatriz y Doris por brindarme su apoyo en los inicios de mi carrera.
A Andreina Guevara por ser mi amiga incondicional y por demostrarme que con esfuerzo se puede lograr lo que se propone.
A Lilibeth por ser mi compañera de estudios en estos ultimos semestres y por brindarme su apoyo.
A Dionny por ser mi amigo y por estar presente en los momentos dificiles.
A mi tutora Clavel Quintana, por su grandiosa colobaración en la elaboración de este trabajo sin ella esto no fuese sido posible, la admiro.

ÍNDICE

Resumen

Nuestro trabajo está basado en la propuesta hecha por Oliveira [17] en su artículo en el año 2011.

Se desarrolla un método de punto proximal para resolver problemas de programación multiobjetivo basado en la escalarización para funciones. Se construye una familia de representaciones escalares convexa estricta de una función vectorial F convexa de $\mathbb{R}^n$ a $\mathbb{R}^m$ con respecto al orden lexicográfico en $\mathbb{R}^m$ y se añade la variante de Auslender, donde las variables sin restricciones en el dominio de F son introducidos en el término cuadrático. Las variables en la escalarización no negativa se colocan en el término logarítmico.

Se demuestra que la sucesión $\{(x^k, z^k)\}$ del problema escalarizado es acotado y converge a una solución débil del problema de optimización multiobjetivo.

Introducción

Se desarrolla la teoría propuesta en el articulo de Oliveira [17] donde se considera la clase de problemas que minimizan a un conjunto de funciones objetivo llamado programación multiobjetivo. Muchas de las aplicaciones importantes que se encuentran en la literatura tienen esta estructura, la cual se relaciona con los problemas de toma de decisiones. Hay una clase mas general de problemas que contiene programación multiobjetivo, conocido como optimización de vector, [11]. Por otro lado, los métodos desarrollados para esta clase de problemas se puede clasificar en dos tipos: Los métodos de escalarización basado en el producto interno $\langle F(x), Z\rangle$, con $Z \in \mathbb{R}^m_+ \backslash \{0\}$ y las extensiones de los logaritmos no lineales a la optimización de vectores. Algunas técnicas de optimización global se discuten en [5].

El método de punto proximal es una clase particular de algoritmo que se ha extendido a optimización multiobjetivo. El primero en esta linea fue el método multiobjetivo bundle presentado por Kiwiel 1990 [13] y [9] donde el método de punto proximal multiobjetivo para representación escalar $\langle F(x), Z\rangle$ se sustituye el núcleo cuadrático por distancia de bregman en espacios de dimensión finita [3]. En el 2005 [10] presenta un algoritmo proximal con una regularización cuadrática en la forma vectorial.

Oliveira [17] propone un método de punto proximal donde se sustituye el núcleo cuadrático por una variante de la función logarítmica cuadrática de Auslender como la regularización.

Se introduce las variables del dominio de la función objetivo F en el término cuadrático y las variables en la escalarización se introducen en el término logarítmico.

Este método se desarrolla solamente para problemas de optimización multiobjetivo.

CAPÍTULO 1

PRELIMINARES

1.1. El concepto de regularización

La idea de regularización surge en relación con los problemas mal planteados. Dado un problema de la forma

$$\mathcal{L}(f) = 0, \tag{1.1}$$

donde f es un elemento de $X \subset \mathbb{R}^n$ (usualmente una función intervalo) y $\mathcal{L} : X \longrightarrow X$ es un operador (usualmente diferencial), (1.1) se dice que es mal planteado cuando no tiene solución, o tiene mas de una solución, o tiene solución única, pero esta solución depende del operador $\mathcal{L}$. La idea es reemplazar $\mathcal{L}$ por un operador regularizado $\mathcal{L}+\lambda\mathcal{M}$ (con $\lambda \in \mathcal{R}$, $\mathcal{M} : X \longrightarrow X$), cuando $\mathcal{M}$ es tal que el problema

$$f_\lambda = (\mathcal{L} + \lambda\mathcal{M})(f) = 0, \tag{1.2}$$

es bien planteado, para cualquier $\lambda > 0$. En tal caso (1.2) tiene única solución f_λ, y con un λ que se aproxima a 0, f_λ provee alguna clase de aproximación de una solución de (1.1).
Si se aplica este concepto para problemas de optimización tomando $X = \mathcal{R}^n$ y $\mathcal{L} = \nabla f$ cuando f es una función convexa ($f : R^n \longrightarrow R$) en el caso (1.1)

$$\nabla f(x) = 0. \tag{1.3}$$

O equivalentemente

$$min_{x\in\mathbb{R}^n} f(x) = 0 \tag{1.4}$$

Se asume que f esta acotada por debajo y tomando $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ estrictamente convexa y coerciva (es decir $lim_{\|x\|\longrightarrow\infty} f(x) = +\infty$). El problema (1.3) no tiene solución o tiene más de una solución para el problema regularizado

$$min_{x\in\mathbb{R}^n} f(x) + \lambda g(x) \tag{1.5}$$

tiene solución única para cada $\lambda > 0$, para el minimo y $f + \lambda g$ es coerciva (usando el hecho que f es acotada por debajo) así el problema se reduce a un conjunto compacto, garantizando la existencia de soluciones y también estrictamente convexa, implicando que la solución de (1.5) es única $x(\lambda)$ y bajo hipótesis razonable (incluyendo existencia de soluciones de (1.4)) puede ser que provee que $lim_{\lambda\longrightarrow 0^+} X(\lambda)$ existe y solución (1.4). El problema en esta regularización es una aproximación aunque $f + \lambda g$ es estrictamente convexa y coerciva para algún $\lambda > 0$ por muy pequeño, para cada λ muy pequeño esta función f es ordenada numéricamente es casi mal planteado, en otras palabras si el sistema $\nabla f(x) = 0$ es mal condicionado entonces el sistema $(\nabla f + \lambda \nabla g)(x) = 0$ es mal condicionado cuando λ se aproxima a 0, a pesar del hecho que tiene solución única para todo $\lambda > 0$.

1.2. El algoritmo de punto proximal para optimización en $\mathbb{R}^n$

Con el fin de superar la dificultad que se acaba de mencionar, sería deseable desarrollar un enfoque de regularización que no requiere el parametro de regularización λ para acercarse a 0 (es decir con una función con constante λ). El algoritmo de punto proximal alcanza tal objetivo.
Se genera una sucesión $\{x^k\} \subset \mathbb{R}^n$ de la siguiente manera:

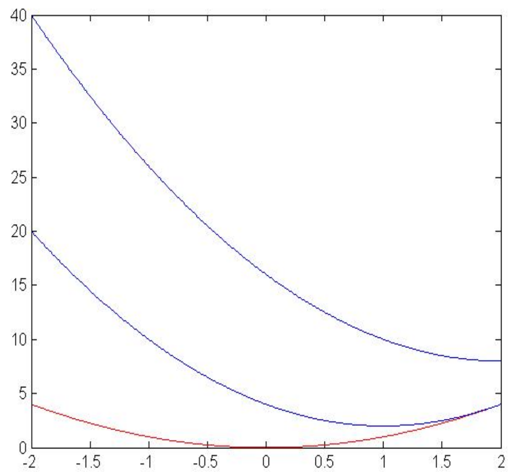
$$x^0 \in \mathbb{R}^n, \tag{1.6}$$

$$x^{k+1} = argmin_{x\in\mathbb{R}^n}\{f(x) + \lambda_k\|x - x^k\|^2\}, \tag{1.7}$$

donde λ_k es un número real que satisface

$$0 < \lambda_k < \overline{\lambda}. \tag{1.8}$$

Para algún $\overline{\lambda} > 0$ (que incluye el caso de los λ_k constante) y $\|.\|$ es la norma euclideana. Se mostrara a continuación que, bajo ciertas hipótesis razonables la sucesión generada



por (1.6) y (1.7) converge a un minimizador de f.
Un enfoque para esta prueba es la convergencia que pasa a través del concepto de expansión y función, de hecho para un problema mucho más general que (1.3). Se sigue aqui un enfoque basado en el más débil de f Fejér convergencia, que funciona también para las extensiones no cuadráticas.

Definición 1.1. (Ver[10]) Una sucesión $\{y^k\} \in \mathbb{R}^n$ se dice que es fejér convergencia a un conjunto $U \subset \mathbb{R}^n$ con respecto a la distancia euclidea si

$$\|y^{k+1} - u\| \leq \|y^k - u\| \tag{1.9}$$

para todo $k \geq 0$, para todo $u \in U$.
Tenemos los siguientes resultados

Propisicion 1.1. *(Ver[10]) Si $\{y^k\}$ es Fejér convergente a $U \neq \emptyset$ entoces $\{y^k\}$ está acotado si un punto adherente y de $\{y^k\}$ pertenece a U entonces $Y = lim_{k\longrightarrow\infty} y^k$*

Prueba 1.1. *(1.9) implica $\|y^k - u\| \leq \|y^0 - u\|$ para cualquier $u \in U$ se tiene que la sucesión $\{y^k\}$ está contenida en una bola de centro u y radio $\|y^0 - U\|$ por lo tanto*

está acotada.
Para la segunda declaración, sea $\{y^{jk}\}$ *una subsucesión de* $\{y^k\}$ *tal que* $lim_{k\longrightarrow\infty} y^{jk} = y$ *puesto que* $y \in U$ *por (1.9) la sucesión* $\{\|y^k - y\|\}$ *es decreciente y no negativo, y una subsucesión (es decir* $\{\|y^{jk} - y\|\}$*) la cual converge a cero . Entonces toda la sucesión converge a 0, es decir* $0 = lim_{k\longrightarrow\infty}\|y^k - y\| \Longrightarrow y = lim_{k\longrightarrow\infty} y^k$

Sea $u \in U$ con $U \in \mathbb{R}^n$ sea $\{y^k\}$ una sucesión por la proposición 1 se tiene que:
Ahora la convergencia del algoritmo de punto proximal.

Teorema 1.1. *(Ver[10]) Sea* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *convexa y continuamente diferenciable. Suponga que el conjunto de minimizadores* U *de* $f \in \mathbb{R}^n$ *es no-vacío. Entonces la sucesión* $\{y^k\}$ *generada por (1.6) y (1.7) converge a un punto* $x^* \in U$.

Prueba 1.2. *Ver la prueba en [10]*

En el trabajo de Teobulle 1992 (ver[1]), se introduce y estudia las propiedades básicas de las aplicaciones proximal, las cuales se asemejan al punto proximal clásico donde las distancias cuadráticas en (1.7) se reemplazan por la $\varphi-divergencia$ de Csiszár en 1967. Dada una función estrictamente convexa φ definida sobre la recta real, la $\varphi-divergencia$ entre x y y esta definida por:

$$d_\varphi(x, y) = \sum_{j=1}^{m} y_j \varphi(\frac{x_j}{y_j}) \tag{1.10}$$

Más recientemente, las propiedades de convergencia de algoritmos de minimización han sido investigados donde (1.7) el término cuadrático es reemplazado por una $\varphi-divergencia$. Censor y Zenios (1992) propusieron un algoritmo de punto proximal donde el término cuadrático es remplazado por la distancia de Bregman (1967)(ver[10]).
Este algoritmo fue estudiado por [8] en 1993 y alli [6] demuestra que la aplicaci´òn proximal se asemeja al proximal clásico.

CAPÍTULO 2

OPTIMIZACIÓN MULTIOBJETIVO

En las ciencias como en la ingeniería se dan, en muchas ocasiones, problemas que requieran encontrar la optimización simultánea de dos o más función objetivo (optimización multiobjetivo). Habrá que optimizar por tanto una función de la forma $f : S \to T$, donde $S \subset \mathbb{R}^n$ y $T \subset \mathbb{R}^m$. Pero el problema está en que normalmente no existe un elemento de S que produzca un óptimo de forma simultánea para cada uno de los m objetivos que componen f. Esto se deberá a la existencia de conflictos entre objetivos, que harán que la mejora de uno de ellos dé lugar a un empeoramiento de algún otro. Pensemos por ejemplo:

- En el caso de un avión con dos objetivos que fueran su velocidad y el ahorro de combustible: un aumento de la velocidad traería consigo una disminución del ahorro. Habrá que llegar por tanto a una situación de compromiso en la que todos los objetivos sean satisfechos en un grado aceptable, desde el punto de vista de diseño.

- En la fabricación de cierto producto que se requiera minimizar el costo de su producción, reducir la cantidad de horas de sobretiempos y el de reducir al mínimo la merma de fabricación de dicho producto; el cual evidencia que el problema posee más de un objetivo.

A diferencia de los problemas de optimización con un único objetivo, el concepto de óptimo es ahora relativo y será necesario decidir de alguna forma cuál es la mejor solución (o cuáles son las mejores soluciones) al problema presente.

En términos matemáticos, el problema de optimización multiobjetivo restringido puede establecerse como: encontrar un vector $x = (x_1, x_2, ..., x_n)^t$, que satisfaga:

$$\begin{array}{ll} \min & f(x) = (f_1(x), ..., f_m(x)) \\ \text{s.a:} & g(x) = (g_1(x), ..., g_p(x)) \leq 0, \\ & h(x) = (h_1(x), ..., h_q(x)) = 0, \\ & x \in C, \end{array} \tag{MP}$$

donde C es un conjunto convexo, con $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^p$ y $h : \mathbb{R}^n \to \mathbb{R}^q$ funciones localmente Lipschitz.

Si una solución x satisface todas las restricciones se dice que es factible y en caso contrario se dice que no es factible; además el conjunto de todas las soluciones factibles es llamado *Región Factible* $\mathcal{X}$ (Espacio de Decisión).

La teoría que a continuación presentamos en las siguientes secciones fue tomada de [7], el cual nos ayudo a entender los problemas de optimización multiobjetivo.

2.1. Óptimo de Pareto

En contraste con la optimización mono-objetivo, la solución a un problema multiobjetivo es más un concepto que una definición. Por regla general, no hay una única solución global y es necesario determinar un conjunto de puntos que cumplan una definición predeterminada de óptimo. El concepto predominante en la definición de este óptimo es el óptimo en el sentido de Pareto, que se define de la siguiente manera para un problema donde se asume que todas las funciones objetivos son de minimización:

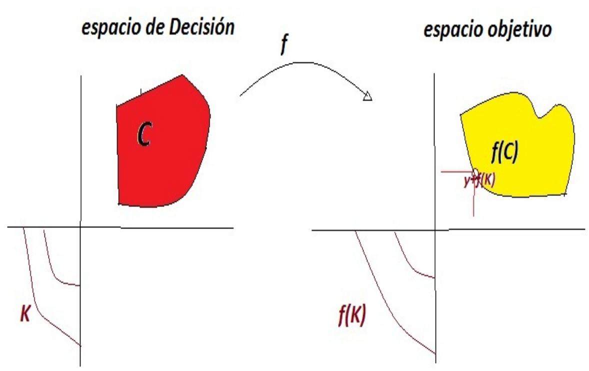
Definición 2.1. (Óptimo de Pareto) (Ver [7])
Un punto $\bar{x} \in X$, es un óptimo en el sentido de Pareto de (MP) si y sólo si no existe otro punto $x \in X$, de manera que $f_i(x) < f_i(\bar{x})$ para algún $i \in M$ y $f_j(x) \leq f_j(\bar{x})$ para

todo $j \in M$.

Con frecuencia, los algoritmos proporcionan soluciones que puede que no sean óptimos de Pareto, pero que pueden satisfacer otros criterios, haciéndolos apropiados para aplicaciones prácticas. Es por ello que se define el óptimo débil de Pareto como sigue:

Definición 2.2. (Óptimo Débil de Pareto) (Ver [7])
Un punto $\bar{x} \in X$, es un óptimo en el sentido de Pareto si y sólo si no existe otro punto $x \in X$, de manera que $f(x) \leq f(\bar{x})$. Es decir, un punto es un óptimo débil de Pareto si no hay otro punto que mejore todas las funciones objetivo simultáneamente.

Por el contrario, un punto es un óptimo de Pareto si no existe otro punto que mejore al menos una de las funciones objetivo sin empeorar cualquier otra de ellas. Los óptimos de Pareto son también óptimos débiles de Pareto, pero el recíproco no es cierto.



2.2. Métodos de solución para problemas MP

Los métodos para resolver un MP pueden ser divididos en cinco conjuntos. Estos son:

1. Métodos Escalares
2. Métodos Interactivos
3. Métodos Difusos
4. Métodos que usan Metaheurística
5. Métodos de Decisiones.

De este conjunto de métodos que permiten solucionar el problema MP vamos a considerar dos de los métodos escalares: el Método de la Suma Ponderada de Pesos y el Método de ϵ-restricciones.

2.2.1. Método de la suma ponderada de pesos

El método de la suma ponderada sugiere escalarizar un conjunto de objetivos, multiplicar cada objetivo por un vector de peso y luego sumarlos, es decir, que un problema multiobjetivo es convertido en un problema de un sólo objetivo de la siguiente manera:

$$\begin{array}{lll} \text{Minimizar-Maximizar} & F(x) = \displaystyle\sum_{i=1}^{m} w_i f_i(x) & \\ \text{s.a.} & g_j(x) \leq 0, & j = 1, 2, \ldots, p, \\ & h_k(x) = 0, & k = 1, 2, ..., q, \\ & x \in C \subset \mathbb{R}^n, & \end{array}$$

donde $x \in \mathbb{R}^n$ y w_i es el peso del i-ésima función objetivo. Usualmente y sin pérdida de generalidad, se escogen pesos fraccionales y diferentes de cero, de manera que se cumpla: $w_i \in (0, 1)$ y $\displaystyle\sum_{i=1}^{m} w_i = 1$.

El peso de un objetivo es usualmente escogido en proporción a la importancia relativa de los objetivos del problema. Si se utiliza un algoritmo de optimización que obtiene resultados exactos y los pesos escogidos son siempre positivos, el método genera soluciones que pertenecen al conjunto Pareto óptimo.

Ejemplo 2.1. Al determinar un número real x que sea mínimo en el problema planteado por Shaffer (1984):

$$f(x) = \begin{cases} f_1(x) = x^2 \\ f_2(x) = (x-2)^2 \end{cases} \qquad con -2 \leq x \leq 4$$

Ver figura 2.1

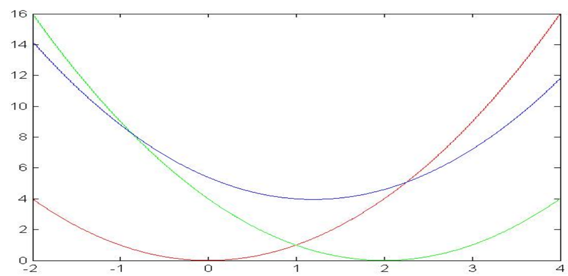


Figura 2.1: Gráfica del problema de Shaffer

2.2.2. Método de ϵ-restricciones

El método consiste en construir un problema de optimización en el que todos los objetivos, excepto uno, se usan como restricciones ($m-1$ restricciones), mientras el sobrante, que puede escogerse aleatoriamente, se usa como función objetivo del problema resultante. La forma general del problema es la siguiente:

$$\begin{array}{lll} \text{Minimizar} & f_r(x) & \\ \text{s.a.} & f_i(x) \leq f_i(x^*) & 1 \leq i \leq M-1 \text{ con } i \neq r \\ & g_j(x) \leq 0 & j = 1, 2, \ldots, p \\ & h_k(x) = 0 & k = 1, 2, ..., q \\ & x \in C. & \end{array}$$

Para encontrar una solución Pareto óptima, se escoge un valor adecuado para $\epsilon_i = f_i(x^*)$ de la i-ésima función objetivo ($i \neq r$). Luego se resuelve el problema de optimización con objetivo único. El procedimiento se repite, con diferentes valores de ϵ_i para hallar nuevas soluciones pertenecientes al frente Pareto óptimo. La dificultad principal que presenta este método reside en la necesidad de un conocimiento del rango apropiado de valores para asignar a ϵ_i para las $(m-1)$ funciones objetivos, lo que no resulta fácil en la práctica.

2.2.3. Escalarización

La teoría que se presenta a continuación fue tomada de [11].
El ortante no negativo de $\mathbb{R}^m$ denotado por $\mathbb{R}^m_+$, cuyo interior está representado por $\mathbb{R}^m_{++}$, desempeña un papel fundamental en la teoría de óptimización de vectores y el orden lexicográfico generado para este cono establece una clase de problemas conocidos como la programación multiobjectivo. Definimos como el orden lexicográfico y denotamos por $\leq$ (a veces por $<$ en otras ocasiones por $\ll$) el orden dado por las siguientes deficiones:

Definición 2.3. (Ver[11])
Sea y, $\overline{y} \in \mathbb{R}^m$ vectores. Tenemos

$$y \leq \overline{y} \Longleftrightarrow y_i \leq \overline{y}_i, i = 1, ..., my < \overline{y} \Longleftrightarrow y_i \leq \overline{y}_i, i = 1, ..., m \tag{2.1}$$

con la desigualdad estricta asegurando por lo menos un indice y

$$y \ll \overline{y} \Longleftrightarrow y_i \leq \overline{y}_i, \forall i = 1, ..., m \tag{2.2}$$

Es fácil ver que $\leq$ satisface los axiomas de la relación de orden parcial en $\mathbb{R}^m$. En un nivel más general cono convexo cerrado apuntado K en $\mathbb{R}^m$ podemos construir una relación de orden parcial $\leq_k$ suponiendo que $y \leq_k \overline{y}$ si $\overline{y} - y \in$ K ($y <_k \overline{y}$ si $\overline{y} - y \in int(k)$)
El problema sin restricciones de optmización multiobjetivo está dado por:

$$minF(x) \tag{2.3}$$

$$x \in \mathbb{R}^m$$

donde F es una función de $\mathbb{R}^n$ a $\mathbb{R}^m$

Definición 2.4. (Ver[11])
Decimos que a $\in \mathbb{R}^n$ es una solución pareto local si existe un disco $B_\delta(a) \subset \mathbb{R}^n$, con $\delta > 0$, de tal manera que no hay ninguna $x \in B_\delta(a)$ que satisface $F(x) < F(a)$
Del mismo modo tenemos la siguiente definición

Definición 2.5. (Ver[11])
a $\in \mathbb{R}^n$ se conoce como solución débil pareto local si existe un disco $B_\delta(a) \subset \mathbb{R}^n$, con $\delta > 0$ tal que no hay ninguna $x \in B_\delta(a)$ que satisface $F(x) \ll F(a)$.
Denotaremos por $argmin\{F(x)/x \in \mathbb{R}^n\}$ y $argmin_w\{F(x)/x \in \mathbb{R}^n\}$ solución local pareto y la solución local débil de pareto que se establece en el problema (1.3). Es fácil ver que $argmin\{F(x)/x \in \mathbb{R}^n\} \subset argmin_w\{F(x)/x \in \mathbb{R}^n\}$. Más detalles sobre el óptimo de pareto y optimización multiobjetivo se puede encontrar en [5].

2.3. Representación escalar

Escalarización es un concepto en la optimización de vector que juega un papel fundamental en el desarrollo de métodos para resolver esta clase de problemas y también se emplea como una herramienta para obtener la convergencia de los algoritmos, tales como, por ejemplo el método de punto proximal presentado por [9] y [3].

Definición 2.6. (Ver[11])
Una función de valor real $f : \mathbb{R}^n \to \mathbb{R}$ se llama representación escalar estricta de un mapa $F : \mathbb{R}^n \to \mathbb{R}^m$ cuando dados $x, \bar{x} \in \mathbb{R}^n$

$$F(x) \leq F(\bar{x})(x) \leq f(\bar{x}),$$

y

$$F(x) \ll F(\bar{x}) \Longrightarrow f(x) < f(\bar{x}).$$

Mas aún, decimos que f es una representación escalar débil de F si

$$F(x) \ll F(\bar{x}) \Longrightarrow f(x) < f(\bar{x}).$$

Es obvio que todas las representaciones escalares estrictas son representaciones escalares debiles la siguiente proposición muestra una forma interesante de conseguir representación escalar para las funciones.

Propisicion 2.1. *(Ver[11])*
Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función. f es una representación escalar estricta de F si y sólo si f es una composición de F con una función $g : F(\mathbb{R}^n) \to \mathbb{R}$ estrictamente creciente.

Prueba 2.1. *Supongamos que g es una función estrictamente creciente. Dados $x, y \in \mathbb{R}^n$*
$(\Longrightarrow)$

$$F(x) \leq F(y) \Longrightarrow g \circ F(x) \leq g \circ F(y)$$

$$F(x) \ll F(y) \Longrightarrow g \circ F(x) < g \circ F(y)$$

Tenemos que $f = g \circ F$ es una representación escalar estricto de F.

$(\Longleftarrow)$ *Ahora supongamos que f es una representación escalar estricta.*
Definamos $g(z) = f(x)$ con $z \in F(\mathbb{R}^n)$ para cualquier $x \in \mathbb{R}^n$ tal que $z = F(x)$
Supongamos que existe $y \in \mathbb{R}^n$ con $F(y) = z$ y $F(x) = z$

$$F(y) \leq F(x) \leq F(y) \tag{2.4}$$

De (2.4) $F(y) \leq F(x) \Longrightarrow f(y) \leq f(x)$
De (2.4) $F(x) \leq F(y) \Longrightarrow f(x) \leq f(y)$
As ´ì $g(F(x)) = f(x) = f(y) = g(F(y))$
g es creciente.
Sean $z, w \in F(\mathbb{R}^n)$ tal que $z \leq w$ asi existen $x, y \in \mathbb{R}^n$ tal que $z = F(x) \leq F(y) = w$ como f es $\mathbb{R} \in \epsilon$

$$\Longrightarrow f(x) \leq f(y) \Rightarrow g(z) \leq g(w)$$

Para las soluciones debilmente pareto.

Propisicion 2.2. *(Ver[11])*
Sea $f : \mathbb{R}^n \to \mathbb{R}$ una representación débil escalar de una aplicación $F : \mathbb{R}^n \to \mathbb{R}^m$ y $argmin\{f(x)/x \in \mathbb{R}^n\}$ el conjunto minimizador local de f, tenemos la inclusión:

$$argmin\{F(x)/x \in \mathbb{R}^n\} \subset argmin_w\{F(x)/x \in \mathbb{R}^n\}.$$

Prueba 2.2. *La demostración es inmediata de la definición (2.5)*

2.4. Representación escalar y convexidad

Definición 2.7. (Ver[11])
Se dice que $F : \mathbb{R}^n \to \mathbb{R}^m$ es una función convexa si y sólo si, para cada $x, y \in \mathbb{R}^n$ y $\lambda \in (0, 1)$,

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y). \tag{2.5}$$

Bajo este supuesto, el problema (2.3) se dice convexo. La desigualdad (2.5) también implica que F es una función convexa si y sólo si, cada componente $F_i : \mathbb{R}^n \to \mathbb{R}$,$i = 1, ..., m$, es una función convexa. La relevancia de la convexidad en la programación multiobjetivo se debe al hecho que cada solución local (débil) pareto es también solución global (débil) pareto sin restricciones o restringida para problemas de optimización multiobjetivo. Este resultado se examina en el teorema 2.2.3 en [13].
La proposición (2.2) establece las condiciones necesarias y suficiente para construir representaciones escalar estricta de F. De acuerdo con la proposición 2.9 en Luc, para obtener la convexidad del problema, se debe elegir una función g convexa de $F(\mathbb{R}^n)$ a $\mathbb{R}$. La función $g_z : \mathbb{R}^m \to \mathbb{R}$ dada por $g_z(y) = \langle y, z\rangle$, con $z \in \mathbb{R}^m_+ \diagdown \{0\}$ fijo, es un ejemplo de la función convexa creciente que se puede componer con F para obtener una representación escalar convexa estricta f de [3], utilizan esta representación para conseguir la convergencia del algoritmo de punto proximal clásico extendido a el vector de optimización [9]
También se presenta un algoritmo punto proximal escalar con distancias de bregman en las mismas línea para vector optmización en espacios de dimensión finita.
$g \circ F(x)$ es convexa solo cuando z es fijo. Si z es una variable incorporada al problema no se puede garantizar la convexidad de la función $f(x, z) = \langle F(x), z\rangle$.
En este trabajo se asume la existencia de una función convexa $f : \mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}$ que satisface las siguientes propiedades:
(P1) f esta acotada por debajo por cualquier $\alpha \in \mathbb{R}$, es decir, $f(x, z) \geq \alpha$, para cada $(x, z) \in \mathbb{R}^n \times \mathbb{R}^m_+$
(P2) f es convexa en $\mathcal{R}^n \times \mathcal{R}^m_+$, es decir teniendo en cuenta $(x_1, z_1), (x_2, z_2) \in \mathcal{R}^n \times \mathbb{R}^m_+$

y $\alpha \in (0,1)$

$$f(\lambda(x_1,z_1)+(1-\lambda)(x_2,z_2)) \leq \lambda f(x_1,z_1)+(1-\lambda)f(x_2,z_2)$$

(P3) f es una representación escalar estricta de F, con respecto a x, es decir,

$$F(x) \leq F(y) \Rightarrow f(x,z) \leq f(y,z)$$

y

$$F(x) \ll F(y) \Rightarrow f(x,z) < f(y,z)$$

para cada $x, y \in \mathbb{R}^n$ y $z \in \mathbb{R}^m_+$

(P4) f es diferenciable, con respecto a z y

$$\frac{\partial}{\partial z} f(x,z) = h(x,z)$$

donde $h(x,z) = (h_1(x,z), ..., h_m(x,z))^T$ es una aplicación continua de $\mathbb{R}^n \times \mathbb{R}^m_+$, es decir, $h_i(x,z) \geq 0$ para todo $i = 1, ..., m$.

La suposición (P1) no es una hipótesis fuerte porque la definición de f implica que no hay un punto $(x,z) \in \mathbb{R}^n \times \mathbb{R}^m_+$ tal que $f(x,z) = +\infty$ se utiliza esta notación en el sentido que no hay un punto $(x,z) \in \mathcal{R}^n \times \mathcal{R}^m_+$ con $lim_k \to +\infty \mid f(x^k,z^k) \mid = +\infty$ para toda sucesión $\{(x^k,z^k)\}_{k\in\mathbb{N}} \subset \mathbb{R}^n \times \mathbb{R}^m_+$ satifaciendo $lim_k \to +\infty(x^k,z^k) = (x,z)$

El conjunto de funciones que sastiface esas propiedades es no vació. Como un ejemplo se tiene $f(x,z) = \sum_{i=1}^m exp(z_i + F_i(x))$ satisface de (P1) a (P4)

Prueba 2.3. *(P1) f esta acotada por debajo por cualquier $\alpha \in \mathbb{R}$, es decir, $f(x,z) \geq \alpha$, para cada $(x,z) \in \mathbb{R}^n \times \mathbb{R}^m_+$*

La exponencial esta acotada por debajo por 0, es decir para cualquier $\alpha \in \mathbb{R}$, $f(x,z) \geq \alpha$ para cada $(x,z) \in \mathbb{R}^n \times \mathbb{R}^m_+$

(P2) f es convexa en $\mathbb{R}^n \times \mathbb{R}^m_+$, es decir teniendo en cuenta $(x_1,z_1),(x_2,z_2) \in \mathbb{R}^n \times \mathbb{R}^m_+$ y $\alpha \in (0,1)$

Ya que $\exp^{z_i+F_i(x)}$ es convexa $\forall i$ y la suma de convexa es convexa se tiene que

$$f(\lambda(x_1,z_1)+(1-\lambda)(x_2,z_2)) \leq \lambda f(x_1,z_1)+(1-\lambda)f(x_2,z_2)$$

(P3) f es una representación escalar estricto de F, con respecto a x, es decir,

$$F(x) \leq F(y) \rightarrow f(x,z) \leq f(y,z)$$

Como, $F_i(x) \leq F_i(y)$ con i=1,...m, $z_i + F_i(x) \leq z_i + F_i(y)$

$\sum_i \exp^{z_i+F_i(x)} \leq \sum_i \exp^{z_i+F_i(y)}$

(P4) f es diferenciable, con respecto a z y

$$\frac{\partial}{\partial z} f(x,z) = h(x,z)$$

$$\frac{\partial}{\partial z} f(x,z) = \frac{\partial}{\partial z} \sum_{i=1}^{m} exp^{z_i+F_i(x)}$$

$$\frac{\partial}{\partial z} f(x,z) = z_1(z_1 + F_1(x))e^{z_1+F_1(x)} + ... + z_m(z_m + F_m(x))e^{z_m+F_m(x)}$$

En este trabajo, se busca $x^* \in argmin\{f_{\overline{z}}(x) = f(x,\overline{z})/x \in \mathbb{R}^n\}$ para cualquier $\overline{z} \in \mathbb{R}^m_+$ Por lo tanto por (P3) y la proposición 3 se concluye que x^* es una solución pareto débil para el problema sin restricciones multiobjetivo.

2.5. Problema Proximal Multiobjetivo

La versión exacta del algoritmo proximal utilizado por [10] en su trabajo (llamado ALG1) requiere dos susesiones exógenas: una sucesión acotada de números reales positivos $\{\alpha_n\}$, y una sucesión $\{e_n\} \subset int(c)$ tal que $\|e_n\| = 1$, para todo n. El método genera una sucesión de la siguiente manera:
Inicialización: elegir $x^0 \in (F)$

Regla de interrupción: dado x_n, si $x_n \in \mathbb{C} - ARGMIN_w\{F(x)/x \in X\}$, entonces $x_{n+p} = x_n$ para todo $p \geq 1$

Paso iterativo: dado x_n, si $x_n \notin \mathbb{C} - ARGMIN_w\{F(x)/x \in X\}$, entonces toma alguna de la siguiente iteración $x_{n+1} \in X$ tal que

$$x_{n+1} \in \mathbb{C} - ARGMIN_w\{F(x) + \tfrac{\alpha_n}{2}\|x - x_n\|^2 e_n/x \in \Omega_n\}$$

con $\Omega := \{x \in X/F(x) \preceq_c F(x_n)\}$ y $dom(F) = \{x \in /F(x) \neq \infty_c\}$ se denomina el dominio efectivo de F

Ahora se hace la siquiente suposición en la función F y la iteración incial x_0:

$$(F(x_0) - C) \cap F(X)(A)$$

El conjunto (A) es C-completo, lo que significa que para toda la sucesión $\{a_n\} \subset X$, con $a_0 = x_0$, tal que $F(a_{n+1}) \preceq_c F(a_n)$ para todo $n \in \mathbb{N}$, no existe $a \in X$ tal que $F(a) \preceq_c F(a_n)$ para todo $n \in \mathbb{N}$.

A continuación se presenta el resultado de convergencia

Teorema 2.1. *(Ver[3]) Sea $F : X \to Y \cup \{\infty_c\}$ una función con la propiedad C-convexa, positiva y semicontinua inferiormente. Bajo la suposición (A), cualquier sucesión $\{x_n\}$ generada por el algoritmo ALG1 es convergente, con respecto a la topología débil de X, para una solución debilmente eficiente de VOP.*

Donde VOP viene dado por:

$$C - MINIMIZE\{F(x)/x \in X\}$$

Prueba 2.4. *Ver la demostración en [3]*

CAPÍTULO 3

MÉTODO DE ESCALARIZACIÓN DEL PUNTO PROXIMAL LOGARÍTMO CUADRÁTICO

En este capítulo se desarrolla los teoremas de [17] para el problema de optimización multiobjetivo con la variante de Auslender [1], mostrando su resultado de convergencia. Sea $F : \mathbb{R}^n \to \mathbb{R}^m$ convexa. Teniendo en cuenta los puntos iniciales $x^0 \in \mathbb{R}^n$, $z^0 \in \mathbb{R}^m_{++}$ y las sucesiones $\beta_k, \mu_k > 0$, k=0,1,..., la escalarización del punto proximal logarítmo cuadrático (LQPS) método que genera las sucesiones, $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ y $\{z^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^m_+$ con las iteraciones x^{k+1} y z^{k+1} se define como la solución del problema (LQPS)

$$min\varphi^k(x,z) = f(x,z) + \beta^k\langle\frac{z}{z^k} - log\frac{z}{z^k} - e, e\rangle + \frac{\mu^k}{2}\|x - x^k\| \tag{3.1}$$

$$x \in \Omega^k, z \in \mathcal{R}^m_+,$$

donde $f : \mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}$ verifica las propiedades (P1) a (P4), $\frac{z}{z^k}$ y $log\frac{z}{z^k}$ son los vectores cuyas componentes están dadas por $\frac{z_i}{z_i^k}$ y $log\frac{z_i}{z_i^k}$ respectivamente $e \in \mathbb{R}^m$ es el vector con todas las componentes igual a 1 y

$$\Omega^k = \{x \in \mathbb{R}^n / F(x) \leq F(x^k)\}.$$

La función objetivo $\Omega^k : \mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}\bigcup\{+\infty\}$, se define en (3.1), es una función integrada con la adición de una variante de la función logaritmica cuadrática de Auslender, donde la regularización logaritmica-cuadrática puede ser visto como una suma de dos funciones separables, un término logarítmico con respecto a la variable z y un término cuadrático con respecto a la variable x.

3.1. Problema bien puesto o bien planteado

Lema 3.1. *(Ver([17])) Sea $F : \mathbb{R}^n \to \mathbb{R}^m$ función convexa y sea $f : \mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}$ una función de verificación de las propiedades (P1) a (P4). Entonces, para cada $k \in \mathbb{N}$, solo hay una solución (x^{k+1}, z^{k+1}) para el problema (LQPS) caracterizado por:*

$$\mu^k(x^k - x^{k+1}) \in \partial_x^{k+1} f(x^{k+1}, z^{k+1}) \tag{3.2}$$

y

$$\frac{1}{z_i^{k+1}} - \frac{1}{z_i^k} = \frac{h_i(x^{k+1}, z^{k+1})}{\beta^k}, i = 1, ..., m$$

Prueba 3.1. *Si F es convexa entonces es continua y convexa en Ω^k. En efecto*

$$\Omega^k = \{x \in \mathbb{R}^n / F(x) \leq F(x^k)\}$$

$x, y \in \Omega^k$ *tal que* $\alpha x + (1 - \alpha)y \in \Omega^k$
$x, y \in \Omega$ $F(x) \leq F(x^k)$ *(1) y* $F(y) \leq F(x^k)$ *(2)*

$$F(\alpha(x) + (1 - \alpha)y) \leq \alpha F(x) + (1 - \alpha)F(y)$$

$$F(\alpha(x) + (1 - \alpha)y) \leq \alpha F(x^k) + (1 - \alpha)F(x^k) \textit{ por (1) y (2)}$$
$$F(\alpha(x) + (1 - \alpha)y) = \alpha F(x^k) + F(x^k) - \alpha F(x^k)$$
$$F(\alpha(x) + (1 - \alpha)y) = F(x^k)$$

$$F(\alpha(x) + (1 - \alpha)y) \leq F(x^k)$$

$\alpha x + (1 - \alpha)y \in \Omega^k$

$$\Omega^k = \{x \in \mathbb{R}^n / F(x) \leq F(x^k)\} = \{x \in \mathbb{R}^n / f_1(x) \leq f_1(x^k) \wedge ... \wedge f_m(x) \leq f_m(x^k)\}$$

$$= \{x \in \mathbb{R}^n / f_i(x) \in (-\infty, f_i(x^k)] \forall i = 1, ..., m\}$$

$$= \bigcap_{i=1}^{m} (-\infty, f_i(x^k)]$$

Así Ω^k *es cerrado. Por lo tanto* $\Omega^k \times \mathbb{R}^m_+$ *es un conjunto convexo cerrado. Tomando* $\frac{z_i}{z_i^k}$*, tenemos que la función* $g : \mathbb{R}_{++} \longrightarrow \mathbb{R}$*, definida por:*

$$g(t) = t - log(t) - 1$$
$$g'(t) = 1 - \frac{1}{t} \Longrightarrow 0 = 1\frac{1}{t}$$
$$\Longrightarrow \frac{1}{t} = 1$$
$$\Longrightarrow 1 = t$$

es estrictamente convexa (pues $g''(t) = \frac{1}{t^2} > 0$*), con su mínimo en* $g(1) = 0$*.*
Se llega a la conclusión que $g(t) \geq 0$*, para cada* $t > 0$*. Esto implica que la función* $\langle \frac{z}{z^k} - log\frac{z}{z^k} - e, e\rangle$ *es estrictamente convexa por ser suma de funciones convexas.*
Por otra parte; Estudiar $\varphi^k(x, z)$

$$f(x,z)+\beta\langle\frac{z}{z^k}-log\frac{z}{z^k}-e,e\rangle+\frac{\mu}{2}\|x-x^k\|^2 = f(x,z)+\beta\langle\frac{z}{z^k}-log\frac{z}{z^k}-e,e\rangle+\frac{\mu}{2}(\|x\|^2-2\langle x,x^k\rangle+\|x^k\|^2)$$
$$\geq \alpha + \frac{\mu}{2}(\|x\| - \|x^k\|)^2 + \beta(\|\frac{z}{z^k} - log\frac{z}{z^k} - e\|_1)$$

donde $\|.\|_1$ *es la norma en* $\mathbb{R}^m$ *definida por* $\|z\|_1 = \sum_{i=1}^m |z_i|$*. La desigualdad es dada por la desigualdad de cauchy-schawrz.*
Ahora, definimos $\|(x, z)\| = \|x\| + \|z\|$ *y supongamos que* $\|(x, z)\| \to +\infty$*. Esto implica que* $\|x\| \to +\infty$ *a* $\|z\| \to +\infty$*. En el primer caso es evidente que* $\varphi^k(x, z) \to +\infty$*.*
Supongamos que $\|z\| \to +\infty$*. Puesto que todas las normas son equivalentes en* $\mathcal{R}^n$*, sin pérdida de generalidad, que* $\|z\|_1 \to +\infty$*. Esto implica que* $|z_l| \to +\infty$ *para algún* $1 \leq l \leq m$ *con* $z_i \geq 0$ *para todo* $i = 1, ..., m$*, se tiene para este índice que* $z_l \to +\infty$*.*
Por otra parte, $\|\frac{z}{z^k} - log\frac{z}{z^k} - e\|_1 = \sum_{i=1}^m \frac{z_i}{z^k} - log\frac{z_i}{z_i^k} - 1$*. Por lo tanto, la parte de la suma asociada a la componente* z_l *satisface* $(\frac{z_l}{z_l^k} - log\frac{z_l}{z_l^k} - 1) \longrightarrow +\infty$*.*
Notese que $g(t)$ *es una función estrictamente creciente para* $t > 1$ *(*$g'(t) > 0$ *para todo* $t > 1$*). Esto es suficiente para tener* $\varphi^k(x, z) \longrightarrow +\infty$ *donde* $\|z\|_1 \longrightarrow +\infty$

$$(\varphi^k(x, z))'_x = \frac{\partial f(x, z)}{\partial x} + \mu^k(x - x^k)$$

$$0 \in \{\frac{\partial f(x,z)}{\partial x} + \mu^k(x - x^k)\}$$

$$0 \in \{\partial_{x^{k+1}} f(x^{k+1}, z^{k+1}) + \mu^k(x^{k+1} - x^k)\}$$

$$\exists \varepsilon \in \partial_{x^{k+1}} f(x^{k+1}, z^{k+1}) \diagup 0 = \varepsilon + \mu^k(x^{k+1} - x^k)$$

$$\varepsilon = \mu^k(x^k - x^{k+1}) \in \partial_{x^{k+1}} f(x^{k+1}, z^{k+1})$$

Para terminar esta prueba se tiene que:

$$0 = h(x,z) + \beta^k \langle \frac{1}{z^k} - \frac{z^k}{z} \frac{1}{z^k}, e \rangle$$

$$0 = h_i(x,z) + \beta^k(\frac{1}{z_i^k} - \frac{1}{z_i})$$

$$\frac{1}{z_i} - \frac{1}{z_i^k} = \frac{h_i}{\beta^k}$$

$$\frac{1}{z_i^{k+1}} - \frac{1}{z_i^k} = \frac{h_i(x^{k+1}, z^{k+1})}{\beta^k}$$

$\forall i = 1, ..., m$

$$\frac{1}{z_i^{k+1}} = \frac{h_i(x^{k+1}, z^{k+1})}{\beta^k} + \frac{1}{z_i^k}$$

$\forall i = 1, ..., m$

$$z_i^{k+1} = \frac{1}{\frac{h_i(x^{k+1}, z^{k+1})}{\beta^k} + \frac{1}{z_i^k}}$$

(B)
Por inducción matemática se obtiene que:

$$z^0 > 0$$

supongamos que $z^k > 0$ *estudiar (B)*

$$z^{k+1} > 0$$

ya que h *satisface (P4) se concluye que* $z_i^k > 0\ \forall i = 1, ..., m$
Cada función h_i *no es necesariamente separable, es decir, no es necesario* $h_i(x, z) = \xi(x) + \eta(z)$
En este trabajo se puede establecer la regla de parada igual que en Bonnel, es decir, si

$$(x^{k+1}, z^{k+1}) = (x^k, z^k)$$

Entonces x^k *es una solución débil de pareto para el problema de optimización multiobjetivo sin restricciones.*

3.2. Convergencia

Teorema 3.1. *(Ver([17])) Sea* $F : \mathbb{R}^n \to \mathbb{R}^m$ *función convexa y* $f : \mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}$ *una función que verifica las propiedades de (P1) a (P4). Supongamos que* $\Omega^0 = \{x \in \mathbb{R}^n / F(x) \leq F(x^0)\}$ *es acotada. Si* $\{\mu^k\}_{k\in\mathbb{N}}$ *y* $\{\beta^k\}_{k\in\mathbb{N}}$ *son suseciones de números reales positivos, con* $\{\mu^k\}_{k\in\mathbb{N}}$ *acotada, entonces la sucesión* $\{(x^k, z^k)\}_{k\in\mathbb{N}}$ *generada por el método de escalarización de punto proximal logaritmo cuadrático es acotada y cada punto adherente* $\{x^k\}_{k\in\mathbb{N}}$ *es un punto débil de pareto para el problema de optimización multiobjetivo sin restriciones.*

Prueba 3.2. *Considere el caso infinito, (es decir, el criterio de parada nunca se cumple). Note que* F *es una aplicación continua (ya que es convexa), esto implica la cerradura de* Ω^0*, más aún por hipótesis* Ω^0 *es acotado en* $\mathbb{R}^n$ *por tanto es compacto en* $\mathbb{R}^n$*. Se tiene que* $\Omega^0 \supseteq \Omega^1 \supseteq ... \supseteq \Omega^k \supseteq \Omega^{k+1} \supseteq ...$ *(ya que la sucesión hace que el valor*

de F decresca),más aún $\Omega = \{x \in \mathbb{R}^n / F(x) \leq F(x^k), \forall k \in \mathbb{N}\}$ es no vacia ($x^0 \in \Omega$)y como $\{x^k\} \subset \Omega^k$ se tiene que $\{x^k\}_{k\in\mathbb{N}}$ es acotada.
Veamos que $\lim_{k\to\infty} \| x^k - x^{k+1} \| = 0$.

Como $\mu^k(x^k - x^{k+1}) \in \partial_{x^{k+1}} f(x^{k+1}, z^{k+1})$ y f es convexa.

$$f(x, z^{k+1}) \geq f(x^{k+1}, z^{k+1}) + (x^k - x^{k+1})' \mu^k (x^k - x^{k+1})$$

$$f(x^{k+1}, z^{k+1}) + \mu^k \langle x^k - x^{k+1}, x^k - x^{k+1} \rangle$$

se puede escribir como:
$\mu^k \langle x^k - x^{k+1}, x^k - x^{k+1} \rangle \geq f(x^{k+1}, z^{k+1}) - f(x, z^{k+1})$ *para cada* $x \in \mathbb{R}^n$.
Por otra parte,

$$\| x^k - x \|^2 = \| x^k - x^{k+1} \|^2 + \| x^{k+1} - x \|^2 + 2\langle x^k - x^{k+1}, x^k - x^{k+1} \rangle$$

$$\geq \| x^k - x^{k+1} \|^2 + \| x^{k+1} - x \|^2 + \frac{2}{\mu_k}(f(x^{k+1}, z^{k+1}) - f(x, z^{k+1}))$$

para cada $x \in \mathbb{R}^n$
Como $F(x^k) \geq F(\overline{x})$ para cada $k \in \mathbb{N}$, con $\overline{x} \in \Omega$ y f satisface (P3) se tiene que:

$$\| x^k - \overline{x} \|^2 \geq \| x^k - x^{k+1} \|^2 + \| x^{k+1} - \overline{x} \|^2 + \frac{2}{\mu_k}(f(x^{k+1}, z^{k+1}) - f(\overline{x}, z^{k+1}))$$

$$\geq \| x^k - x^{k+1} \|^2 + \| x^{k+1} - \overline{x} \|^2 \geq \| \overline{x} - x^{k+1} \|^2$$

Así $\{\| x^{k+1} - \overline{x} \|^2\}_{k\in\mathbb{N}}$ es no- creciente y acotada por lo que es convergente. Esto implica $0 = \lim_{k\to\infty} \| x^k - \overline{x} \|^2 - \| x^{k+1} - \overline{x} \|^2 \geq \lim_{k\to\infty} \| x^k - x^{k+1} \|^2 \geq 0$.
Por otro lado, $\{z_i^k\}_{k\in\mathbb{N}}$ es acotada por debajo por cero y monotona no- creciente por lo que es convergente.
Ahora probaremos que todo punto de acumulacion de $\{x^k\}_{k\in\mathbb{N}}$ pertenece a el conjunto de soluciones debil de pareto.
Por la construcción $f(x^{k+1}, z^{k+1}) \leq f(x^k, z^k)$ para todo $k \in \mathbb{N}$. Como $\{x^k\}_{k\in\mathbb{N}}$ es acotada tiene una subsucesion convergente $\{x^{kj}\}_{j\in\mathbb{N}}$ tal que $x^{kj} \to x^*$, asociada a $\{x^{kj}\}_{j\in\mathbb{N}}$ tenemos $\{z^{kj}\}_{j\in\mathbb{N}}$ y como $\{z^k\}_{k\in\mathbb{N}}$ es convergente $z^{kj} \to z^*$, luego

$$f(x^*, z^*) = \lim_{j\to+\infty} f(x^{kj}, z^{kj}) = \inf_{k\in\mathbb{N}} \{f(x^k, z^k)\}$$

Notar que

$$f(x, z^{kj+1}) \geq f(x^{kj+1}, z^{kj+1}) + \mu^{kj}\langle x^{kj} - x^{kj+1}, x - x^{kj+1}\rangle$$

$$\geq f(x^*, z^*) + \mu^{kj}\langle x^{kj} - x^{kj+1}, x - x^{kj+1}\rangle$$

para todo $x \in \mathbb{R}^n$. *Tomando limite se tiene que*

$$f(x, z^*) \geq \lim_{j\to+\infty} f(x^{kj+1}, z^{kj+1}) \geq f(x^*, z^*)$$

así $x^* \in$ *argmin*$\{f(x, z^*)/x \in \mathbb{R}^n\}$. *En particular si reemplazamos* x *por* x^* *en la desigualdad anterior tenemos que* $\lim_{j\to+\infty} f(x^{kj+1}, z^{kj+1}) = f(x^*, z^*)$.
Para finalizar, $z^* \in \mathbb{R}^m_+$ *por P3* $f(x, z^*)$ *es una representacion escalar de* F *y como toda representación escalar es una representacion escalar débil por la definición (2.6) se tiene que* $x^* \in$ *argmin*$\{F(x)/x \in \mathbb{R}^n\}$

3.3. Criterio de parada

Propisicion 3.1. *(Ver([17])) Sea* $\{(x^k, z^k)\}_{k\in\mathbb{N}}$ *la sucesion generada por el método (LQPS). Si*

$$(x^{k+1}, z^{k+1}) = (x^k, z^k)$$

para cualquier entero k entonces x^k *es una solucion debil de pareto para el problema sin restricciones de optimización multiobjetivo.*

Prueba 3.3. *Supongamos que la regla de parada se verifica en la iteración k-ésima. Procedamos por reducción al absurdo*
supongamos que x^k *no es una solución débil de pareto. Entonces* $\bar{x} \in \mathbb{R}^n$ *tal que* $F(\bar{x}) \ll F(x^k)$ *por (P3) se tiene que* $f(\bar{x}, z^k) < f(x^k, z^k)$
Esto implica que existe $\alpha > 0$ *tal que* $f(\overline{x}, z^k) = f(x^k, z^k) - \alpha$
Definiendo $x\lambda = \lambda x^k + (1-\lambda)\bar{x}$, $\lambda \in (0, 1)$ *se obtiene:*

$$(x\lambda, z^k) = (\lambda x^k + (1-\lambda)\overline{x}, z^k)$$

$$= (\lambda x^k, z^k) + ((1-\lambda)\overline{x}, z^k)$$

$$= \lambda(x^k, z^k) + (1-\lambda)(\overline{x}, z^k)$$

$$g(\frac{z^k}{z^k}) = \frac{z^k}{z^k} - log\frac{z^k}{z^k} - 1 = 0$$

Puesto que (x^k, z^k) *resuelve el problema (LQPS) se tiene que:*

$$\varphi^k(x^k, z^k) \leq \varphi^k(x\lambda, z^k)$$

$$f(x^k, z^k) + \frac{\mu^k}{2}\|(x^k - x^k)\|^2 \leq f(x\lambda, z^k) + \frac{\mu^k}{2}\|(x\lambda - x^k)\|^2$$

$$f(x^k, z^k) = f(x\lambda, z^k) + \frac{\mu^k}{2}\|(x\lambda - x^k)\|^2$$

Estudiemos $\|(x\lambda - x^k)\|^2$

$$\|(x\lambda - x^k)\|^2 = \|(\lambda x^k + (1-\lambda)\overline{x}) - x^k\|^2$$

$$= \|(\lambda - 1)x^k + (1-\lambda)\overline{x}\|^2$$

$$= (1-\lambda)^2\|\overline{x} - x^k\|^2$$

Luego

$$f(x^k, z^k) \leq f(x\lambda, z^k) + (1-\lambda)^2\frac{\mu^k}{2}\|\overline{x} - x^k\|^2$$

Por otro lado, como f *es una función convexa implica que*

$$f(x^k, z^k) \leq \lambda f(x^k, z^k) + (1-\lambda)f(\overline{x}, z^k)$$

$$= \lambda f(x^k, z^k) + (1-\lambda)(f(x^k, z^k) - \alpha)$$

$$= f(x^k, z^k) - (1-\lambda)\alpha$$

Las últimas desigualdades implican que:

$$f(x^k, z^k) \leq f(x\lambda, z^k) + (1-\lambda)^2\frac{\mu^k}{2}\|\overline{x} - x^k\|^2$$

$$\leq f(x^k, z^k) - (1-\lambda)\alpha + (1-\lambda)^2\frac{\mu^k}{2}\|\overline{x} - x^k\|^2$$

Por lo tanto,

$$\alpha \leq (1-\lambda)\frac{\mu^k}{2}\|\overline{x} - x^k\|^2$$

Para cada $\lambda \in (0,1)$ *resulta que*

$$\alpha \leq \lim_{\lambda \longrightarrow \overline{1}} (1-\lambda)\frac{\mu^k}{2}\|\overline{x} - x^k\|^2 = 0$$

lo cual es una contradicción.

Conclusiones

En este trabajo se presenta una nueva familia de métodos de escalarización de programación multiobjetivo. Otros algoritmos, como los de [9] y [3], sólo trabajan con la hipótesis de la existencia de alguna de $Z \in \mathbb{R}^m_+ \backslash \{0\}$ tal que $\langle F(x), Z \rangle$ es una representación escalar estricta de F. Aquí, se muestra que la familia de la representación escalar estricta $f_z = f(x,z)$ converge a una representación escalar estricta $f_{z^*} = f(x, z^*)$ con $Z^* \in \mathbb{R}^m_+$.

Obtenemos el resultado de convergencia del método para una solución débil de Pareto del problema de optimización multiobjetivo no restringida. Pero, no hacemos ninguna referencia sobre el algoritmo que se debe utilizar para calcular (x^{k+1}, z^{k+1}) en el caso sin restricciones $(x^{k+1}, z^{k+1}, y^{k+1})$. Dado que φ^k es estrictamente convexa para todos $k \in \mathbb{N}$, si F es un función C^2, podemos aplicar el método de Newtons, que ha proporcionado una buena opción para esos problemas.

Oliveira [17] presenta un ejemplo de función que satisface las cuatro propiedades (P1), (P2), (P3) y (P4), pero este no es el único ejemplo. Se presenta un ejemplo con otras funciones con las mismas características, y estamos estudiando una forma parcial por el método de considerar aspectos inexactos en la determinación de las iteraciones x^{k+1} y z^{k+1} (es decir, no determinada las iteraciones exactas x^{k+1} y z^{k+1}, pero $\epsilon^k - \textit{iteraciones}$ en cada iteración del método, por cualquier secuencia dada de precisión $\epsilon^k > 0$, $k = 0, 1, 2, ...$) para hacer una implementación.

REFERENCIAS BIBLIOGRÁFICAS

[1] Auslender, A., Teboulle, M., Ben-Tiba, S. A *logarithmic-quadratic proximal method for variational inequalities.* Comput. Optim. Appl. 12 (1-3), 31-40 (1999)

[2] A. Auslender, M, Teboulle and S. Ben-Tiba. *Interior proximal and multiplier methods based on second order homogeneous kernels.* Mathematics of Operations Research 24, No. 3, pp. 645-669, (1999)

[3] Bonnel, H., Iusem, A.N., Svaiter, B.F. *Proximal methods in vector optimization.* SIAM J. Optim. 15 (4), 953-970 (2005)

[4] Chinchuluun, A., Pardalos, P.M.*A survery of recent developments in multiobjetive optimization.* Ann Oper. Res 154(1), 29-50(2007).

[5] Chinculuun, A., Migdalas, A., Pardalos, P.M., Pitsoulis, L.(eds.)*Pareto optimality, game theory and eqilibria.* springer, New York (2008)

[6] J. Eckstein. *Nonlinear proximal point algoritms using Bregman functions, with aplications to convex programming.* Mathematics of Operations REsearchs,18, pp. 202-226, 1993.

[7] M. Ergoth. *Multicriteria Optimization.* vol.2, Springer, New York,(2005).

[8] Fliege, J., Svaiter, B.F *Steespet descent methods for multicriteria optimization.* Math. Methods Oper. Res. 51(3), 479-494 (2000)

[9] Gopfert, A., Riahi, H., Tammer, C., Zalinescu, C. *Variational methods in partially ordered spaces.* Springer, New York (2003)

[10] A. Iusem. *Metodos do ponto proximal em Otimizacao.* 20 Coloquio Brasileiro de Matemática, IMPA,R.J., Brazil, (1995.)

[11] Luc,T.D *Theory of de vector optmization.*Lecture Notes In Economics and Mathematical Systems, 319, Springer, Berlin (1989)

[12] Luenberger, D., Yinyuye *Linear and Nonlinear Programming.* 451, Springer, (2008)

[13] Miettinen, K.M. Nonlinear multiobjective optimization. Kluver, Boston (1999)

[14] B. Martinet *Perturbation des méthodes dóptimization.* Application, R.A.I.R.O, Analyse numérique/ Numerical Analysis, (1978)

[15] R. Polyak and M. Teboulle. *Nonlinear rescaling and proximal-like methods in convex programming.* 76,pp. 667-739, (1997)

[16] R. T. Rockafeller. A dual approach to solving nonlinear programming problems by unconstrained optmization. Mathematical Programming, No. 5, pp. 354-373, 1973.

[17] R. Gregorio y P. Oliveira. *A logarithmic-quadratic proximal point scalarization method multiobjective.* J Glob Optim 49, 281-291 (2011).