



UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE MATEMÁTICA, ASTRONOMÍA, FÍSICA Y COMPUTACIÓN

TRABAJO ESPECIAL
LICENCIATURA EN MATEMÁTICA

MÉTODOS DE OPTIMIZACIÓN PARA EL
PROBLEMA DE LOCALIZACIÓN
DE FERMAT-WEBER

LAURA MONTES

DIRECTOR: DR. ELVIO A. PILOTTA

16 DE MARZO DE 2018



Métodos de optimización para el problema de localización de Fermat-Weber por Laura Montes se distribuye bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Resumen

El problema de Fermat-Weber es uno de los problemas de localización más conocidos. Muchos autores han estudiado este problema y diversas formulaciones, así como metodologías de resolución. El algoritmo clásico utilizado con este fin es el algoritmo de Weiszfeld, el cual se remonta a mediados del siglo pasado. A lo largo de este trabajo, estudiamos el problema de Fermat-Weber, y presentamos dos variantes del mismo: el problema en norma p y el problema con restricciones de tipo caja. Analizamos en profundidad el método de Weiszfeld y las modificaciones existentes para tratar las variantes arriba mencionadas. Además, describimos un método basado en la aplicación del método de Newton para resolver el problema de Fermat-Weber, el cual fue propuesto recientemente, y lo utilizamos para construir un nuevo algoritmo para resolver el problema con restricciones de tipo caja. Realizamos experimentos numéricos para analizar y comparar los métodos mencionados y concluimos que el método de Weiszfeld es, en efecto, un método eficiente para todas las variantes analizadas, y que el método de Newton es efectivo y muy estable, en el sentido de que tiene rendimientos similares para todos los tamaños de problemas. También concluimos que las variaciones del método de Newton propuestas para resolver el problema con restricciones de tipo caja no son eficaces, y resta investigar otras maneras de resolverlo.

Palabras clave:

- Problema de Fermat-Weber.
- Algoritmo de Weiszfeld.
- Método de Newton.
- Problema con restricciones de tipo caja.
- Optimización.
- Problema de localización.

Códigos de clasificación:

- 90C30 Nonlinear programming
- 90B85 Continuous location
- 90C25 Convex programming
- 49M15 Newton-type methods

Abstract

The Fermat-Weber problem is one of the best known location problems. Many authors have studied this problem, different formulations and solution methods. The Weiszfeld algorithm is the classical method for finding the solutions to this problem. Throughout this work, we study the Fermat-Weber problem and present two variations: the problem with l_p distances and the box-constrained problem. We analyze in depth Weiszfeld's method and the existing modifications for solving the problem's variants. We also describe a recent method for solving the Fermat-Weber problem based on Newton's method, and we use it to develop a new algorithm for solving the box-constrained problem. We perform numerical experiments and compare these methods, and we conclude that Weiszfeld's method solves the mentioned variations of Weber's problem efficiently, and that Newton's method is effective and very stable, in the sense of having similar performances in different problem sizes. We also conclude that the proposed variations of Newton's method for solving the box-constrained problem do not work, and further investigation of this matter should be done.

Key words:

- Fermat-Weber's problem.
- Weiszfeld's algorithm.
- Newton's method.
- Box-constrained problem.
- Optimization.
- Location problems.

Classification codes:

- 90C30 Nonlinear programming
- 90B85 Continuous location
- 90C25 Convex programming
- 49M15 Newton-type methods

Agradecimientos

A mi familia, principalmente a mis papás Patricia y Miguel, por asegurarse que nunca me falte nada y apoyarme a lo largo de toda la carrera.

A mi director Elvio, por acompañarme durante todo el último año con asuntos del trabajo, de la carrera y de la vida.

A mi novio Gustavo, por tranquilizarme en mis momentos de estrés y hacerme feliz todos los días.

A Sofía, por estar ahí siempre, como amiga y consejera.

A Santi, Alan, Rami y Juli, por compartir mates y horas de trabajo juntos todo el verano.

A mis maestros de Pakua, por enseñarme a controlar mis nervios y ayudarme a conocerme y superarme cada día más.

A todos mis amigos, compañeros y profesores, de la facultad y de afuera, que han estado en algún momento de mi vida, ayudándome a llegar hasta acá.

Índice general

1. Introducción	1
2. El problema de localización de Fermat-Weber	3
2.1. Un breve recorrido histórico	3
2.2. Propiedades del problema	5
3. El algoritmo de Weiszfeld	13
3.1. Historia del algoritmo	13
3.2. Resultados de convergencia	20
3.3. Variaciones para asegurar convergencia	30
4. Variantes y generalizaciones del algoritmo de Weiszfeld	35
4.1. El problema de Weber para norma p	35
4.2. Aceleración del algoritmo	44
4.3. El problema con restricciones de cotas en las variables	47
5. El método de Newton para el problema de Fermat-Weber	55
5.1. Desarrollo y convergencia del método	55
5.2. Newton aplicado al problema con restricciones de cotas en las variables	59
6. Resultados numéricos	63
6.1. Problema irrestricto	63
6.1.1. Comparación de métodos irrestrictos en norma 2	63
6.1.2. Comparación de métodos irrestrictos en norma p	66
6.2. Problema con restricciones de cajas	71
6.2.1. El problema en \mathbb{R}^2	73
6.2.2. El problema en \mathbb{R}^3	75
7. Conclusiones	79
A. Tablas de resultados para el problema irrestricto en norma 2	81
A.1. Resultados para $n = 2$	81
A.2. Resultados para $n = 3$	84
A.3. Resultados para $n = 5$	86
A.4. Resultados para $n = 10$	89
B. Tablas de resultados para el problema irrestricto en norma p	93
B.1. Resultados para $p = 1.1$	93
B.2. Problema en norma $p = 1.5$	96
B.3. Problema en norma $p = 2.5$	100

Capítulo 1

Introducción

En este trabajo se estudió el problema localización de Fermat-Weber, el cual consiste en encontrar un punto en el espacio que minimice la suma de distancias ponderadas a ciertos puntos dados, y algunas de sus variantes. Además, se analizó el primer método iterativo propuesto para resolver este problema, el método de Weiszfeld, y se lo comparó con otros más recientes.

En el Capítulo 2 presentamos un breve recorrido por la historia del problema y los primeros métodos geométricos para resolverlo. Posteriormente analizamos la función a minimizar, y probamos su continuidad y convexidad y la existencia de minimizadores.

En el Capítulo 3 estudiamos en profundidad el método de Weiszfeld. Comenzamos mencionando la historia del algoritmo y de la determinación de la cantidad de casos en los cuales el mismo se detiene sin hallar la solución. Luego realizamos una demostración completa de la convergencia del algoritmo cuando no se detiene y un análisis de su velocidad de convergencia. Por último, presentamos diversas estrategias para evitar que el algoritmo se detenga fuera de la solución.

En el Capítulo 4 mencionamos algunas variantes del algoritmo de Weiszfeld, para resolver el problema de Fermat-Weber original y algunas variaciones del mismo. Primero definimos una generalización del problema de Fermat-Weber utilizando las distancias l_p en lugar de la distancia euclídea, y describimos una variante del método de Weiszfeld para resolver este problema. Posteriormente estudiamos una nueva variante, partiendo de esta generalización, que permite acelerar el algoritmo, y tratamos de extender a \mathbb{R}^n los resultados existentes para \mathbb{R}^2 . Luego analizamos el problema de Fermat-Weber con restricciones en las variables, y presentamos el algoritmo de Weiszfeld proyectado.

En el Capítulo 5 presentamos un resultado reciente acerca de la convergencia del método de Newton para resolver el problema de Fermat-Weber. Luego, desarrollamos un método para resolver el problema con restricciones en las variables que utilice el método de Newton, con la intención de conseguir una convergencia más rápida que la del método de Weiszfeld proyectado.

En el Capítulo 6 realizamos experimentos numéricos para estudiar los métodos presentados en los capítulos anteriores. Para el problema irrestricto, se compararon los métodos de Weiszfeld, Weiszfeld acelerado y Newton, y para el problema con restricciones en las variables se analizó el funcionamiento de los métodos propuestos en la Sección 5.2 comparándolos con el método de Weiszfeld proyectado.

Finalmente, en el Capítulo 7 presentamos nuestras conclusiones sobre la efectividad de los distintos métodos estudiados, sus ventajas y sus aspectos a mejorar.

Capítulo 2

El problema de localización de Fermat-Weber

2.1. Un breve recorrido histórico

La teoría de localización estudia el problema de ubicar un punto en la mejor posición posible, cumpliendo algún requisito particular. Uno de los problemas de localización más conocidos es el problema de Fermat-Weber, en el cual el requisito que se busca satisfacer es minimizar la suma de las distancias ponderadas del nuevo punto hacia ciertos puntos dados.

El problema de Fermat-Weber se remonta al siglo XVII. Su origen se le atribuye al matemático Pierre de Fermat, quien habría propuesto el desafío de encontrar un punto en el plano que minimice la suma de las distancias a tres puntos dados. Para este problema simple, se propusieron múltiples soluciones geométricas. Las más conocidas son las del físico Evangelista Torricelli y el matemático Thomas Simpson, pero también realizaron aportes otros importantes investigadores de la época, como Cavalieri, Lebesgue, Steiner, entre otros.

El método constructivo de Torricelli fue propuesto para el caso en el cual el triángulo formado por los puntos dados tenga todos sus ángulos menores a 120° , y es el siguiente: a partir cada lado del triángulo inicial, se construyen tres triángulos equiláteros hacia afuera y se los circunscribe en tres circunferencias. El punto donde las tres circunferencias se intersecan es llamado el *punto de Torricelli*, y es la solución del problema de Fermat (ver Figura 2.1).

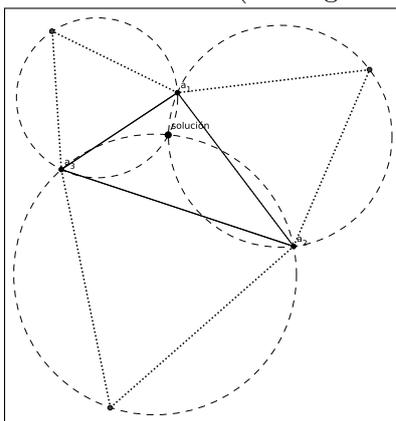


Figura 2.1: Método constructivo de Evangelista Torricelli.

El método de Simpson es muy similar: se comienza construyendo los tres triángulos equiláteros, al igual que en el método de Torricelli, y se trazan tres rectas, cada una conectando un vértice del triángulo inicial, con el vértice exterior del triángulo opuesto (ver Figura 2.2). A estas rectas se las conoce como *rectas de Simpson*, y el punto donde se intersecan es el punto de Fermat.

Todos estos desarrollos, fueron pensados para el problema de sumar las distancias en el plano a tres puntos dados, por lo que no se dieron grandes avances teóricos. Recién a principios del siglo XX,

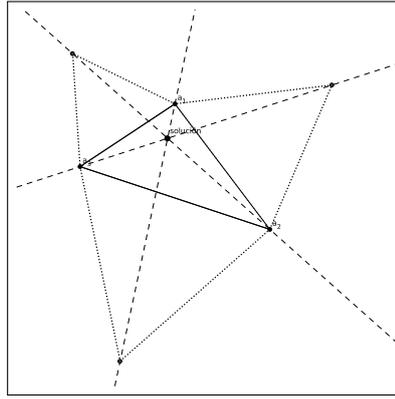


Figura 2.2: Método constructivo de Thomas Simpson.

el problema fue retomado para un estudio en mayor profundidad, al encontrarse aplicaciones reales. En 1909, el economista Alfred Weber [WF29] planteó el problema de Fermat desde un punto de vista económico, considerando a cada uno de los tres puntos dados como fuentes de materia prima o puntos de venta, y al punto buscado como una fábrica que se desea ubicar. El problema consistía en instalar una nueva fábrica de manera de minimizar los costos de transporte de la materia prima hacia la fábrica y del producto terminado hacia un punto de venta. Al incorporar la noción de *costos* se incorporan pesos, considerando que el costo no dependía sólo de la distancia entre las instalaciones sino también de la dificultad de transportar cada producto, la cual puede depender de la cantidad, la forma, la fragilidad, entre otros factores. La formalización de este problema entonces, era hallar un punto que minimice la suma de distancias ponderadas a tres puntos dados.

Si bien el planteo inicial es en el plano \mathbb{R}^2 , puede extenderse a \mathbb{R}^n . Además, puede generalizarse a m puntos iniciales, $m \in \mathbb{N}$, los cuales serán denominados *nodos* de aquí en más. El problema entonces consiste en encontrar el punto $x^* \in \mathbb{R}^n$ que minimice la suma de las distancias ponderadas a m nodos $a_1, \dots, a_m \in \mathbb{R}^n$, donde cada nodo a_i tiene asignado un peso positivo w_i , $i \in \{1, \dots, m\}$.

Este problema fue estudiado en 1937 por el matemático húngaro Endre Vaszonymi Weiszfeld [Wei37], dando origen a un método iterativo para hallar la solución. En su trabajo, demostró la existencia y unicidad de la solución, utilizando el gradiente de la función objetivo para generar un algoritmo de punto fijo. Este algoritmo sin embargo no se hizo conocido en el momento, y fue redescubierto décadas más tarde por distintos investigadores. El trabajo más conocido es el de Harold W. Kuhn y Robert E. Kuenne [KK62], quienes presentaron su algoritmo en 1962.

Tanto las demostraciones de Weiszfeld como las de Kuhn y Kuenne presentaban errores, al no considerar los problemas causados por la no diferenciabilidad de la función objetivo en cada uno de los nodos. Estos problemas fueron tratados extensivamente a lo largo del siglo pasado en trabajos de Kuhn [Kuh73], Chandrasekaran y Tamir [CT89], Brimberg [Bri95] y [Bri03], Cánovas, Cañavate y Marín [CCM02] y otros autores, los cuales llevaron a modificaciones del algoritmo para salvar este problema.

En situaciones de la vida cotidiana, este problema aparece con frecuencia. En el ámbito de la economía, su utilidad fue presentada por Weber en [WF29]. En su tesis de maestría, Kwarteng [Kwa11] presenta el problema de ubicar una biblioteca en el municipio Sunyani, de manera que esté cerca de las comunidades más apartadas. Otras de las aplicaciones, utilizan un modelo similar de localización, aunque más complejo, en el cual se busca ubicar más de un establecimiento. Esta formulación fue aplicada por ejemplo para estudiar la ubicación óptima de centros de salud en distintas regiones de Argentina: Bosque Sendra y Ramírez [RBS01] compararon las ubicaciones de los hospitales públicos existentes en la provincia de Chaco con las ubicaciones óptimas teóricas, y Buzai [Buz13] analizó la distribución de los centros primarios de atención sanitaria en la ciudad de Luján. Estas son algunas de las aplicaciones del problema de Fermat-Weber, entre muchas otras. Guerrero García [Gue07] presenta un poco de historia del problema y estudia otra variación del mismo, que es la ubicación de centros no deseados. Una aplicación para este problema por ejemplo es la ubicación de una fábrica altamente contaminante, por lo cual se busca que esté lo más alejada

posible de los centros urbanos.

A lo largo de este trabajo, estudiaremos el problema de Fermat-Weber y dos variantes del mismo, el problema en norma p y el problema con restricciones en las variables, y analizaremos distintas estrategias para resolverlos.

2.2. Propiedades del problema

La formulación matemática del problema de Fermat-Weber generalizado es:

$$\text{Minimizar } f(x) = \sum_{i=1}^m w_i \|x - a_i\|_2 \quad \text{con } a_1, \dots, a_m \in \mathbb{R}^n \text{ y } w_1, \dots, w_m > 0. \quad (2.2.1)$$

A lo largo de este capítulo presentaremos algunas características del problema que nos permitirán determinar si existen soluciones y si éstas son únicas. Primero presentaremos algunas notaciones que usaremos de aquí en más:

- $\|\cdot\|$ denotará la norma 2 en \mathbb{R}^n : $\|x\| = \|x\|_2 = \sqrt{\sum_{j=1}^n x_j^2}$, $x = (x_1, \dots, x_n)$.
- \mathcal{A} denotará el conjunto de nodos, $\mathcal{A} = \{a_1, \dots, a_m\}$.
- f denotará la función objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$f(x) = \sum_{i=1}^m w_i \|x - a_i\|. \quad (2.2.2)$$

Ahora sí, probaremos que el problema (2.2.1) tiene al menos una solución. Para ello, precisaremos conocer un poco más las características de la función f .

Lema 2.2.1. *La función f es continua y convexa.*

Demostración. Para cada $i \in \{1, \dots, m\}$, la función $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $f_i(x) := \|x - a_i\|$ es continua y convexa, y $w_i > 0$. Entonces $w_i f_i$ es una función continua y convexa, y $f = \sum_{i=1}^m w_i f_i$ resulta también continua y convexa. \square

Utilizando este lema y el comportamiento de la función cuando $\|x\|$ tiende a infinito, podemos demostrar el siguiente teorema:

Teorema 2.2.2. *La función f tiene al menos un minimizador global.*

Demostración. Cada f_i , $i = 1, \dots, m$, satisface que

$$\lim_{\|x\| \rightarrow +\infty} f_i(x) = +\infty.$$

Entonces

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty.$$

Luego $S_M = \{x \in \mathbb{R}^n : f(x) \leq M\}$ es acotado para todo $M \in \mathbb{R}$. Sea M tal que $S_M \neq \emptyset$. Como $0 < f(x) \leq M$ para todo $x \in S_M$ y f es continua, entonces existe $x^* \in S_M$ tal que $f(x^*) \leq f(x)$ para todo $x \in S_M$. Además, $f(x^*) \leq M < f(x)$ para todo $x \in S_M^c$, es decir que $f(x^*) \leq f(x)$ para todo $x \in \mathbb{R}^n = S_M \cup (S_M)^c$, y por lo tanto x^* es minimizador global de f . \square

El Teorema 2.2.2 nos garantiza la existencia de soluciones, pero no garantiza la unicidad. Para estudiar este aspecto, será necesario clasificar el problema en dos casos posibles: aquellos en que los nodos son colineales y el caso en que no lo son. Para el primer caso, es decir, si todos los nodos se encuentran en la misma recta, el problema no necesariamente tiene una única solución. Sin embargo, es sencillo de tratar pues es equivalente a un problema de minimización en una única dimensión, y sus soluciones pueden hallarse de manera explícita. Para el caso donde los nodos a_1, \dots, a_m no son colineales, se puede garantizar la unicidad de la solución y, si bien no se conocen métodos explícitos para determinarla, existen métodos iterativos para aproximarla.

A continuación veremos cómo resolver el primer caso: a_1, \dots, a_m pertenecen a una misma recta.

En primer lugar, veamos que el problema es equivalente a resolver un problema unidimensional. Si los nodos son colineales, existen $v, b \in \mathbb{R}^n$, con $\|v\| = 1$, y t_1, \dots, t_m tales que $a_i = t_i v + b$ para todo $i = 1, \dots, m$. Sin pérdida de generalidad, podemos asumir $t_1 < \dots < t_m$. Denotemos por \mathcal{R} a la recta $\{tv + b : t \in \mathbb{R}\}$.

Lema 2.2.3. *Sea x^* una solución al problema (2.2.1). Si $a_i \in \mathcal{R}$ para todo $i = 1, \dots, m$, entonces $x^* \in \mathcal{R}$.*

Demostración. Sea $P : \mathbb{R}^n \rightarrow \mathcal{R}$ la proyección a la recta \mathcal{R} . Como $a_i \in \mathcal{R}$ tenemos que $P(a_i) = a_i$ para todo $i = 1, \dots, m$. Sabemos que

$$\|x - y\| \geq \|P(x) - P(y)\| \quad \text{para todo } x, y \in \mathbb{R}^n.$$

En particular, tomando $y = a_i$ con $i \in \{1, \dots, m\}$, obtenemos

$$\|x - a_i\| \geq \|P(x) - P(a_i)\| = \|P(x) - a_i\|$$

y la igualdad vale sólo si $x \in \mathcal{R}$. Luego,

$$f(x) = \sum_{i=1}^m w_i \|x - a_i\| \geq \sum_{i=1}^m w_i \|P(x) - a_i\| = f(P(x)),$$

y como las igualdades valen sólo para $x \in \mathcal{R}$, resulta

$$f(x) > f(P(x)) \text{ si } x \notin \mathcal{R}, \text{ y} \tag{2.2.3}$$

$$f(x) = f(P(x)) \text{ si } x \in \mathcal{R}. \tag{2.2.4}$$

Si x^* es el minimizador de f , entonces $f(x^*) \leq f(y)$ para todo $y \in \mathbb{R}^n$. En particular, $f(x^*) \leq f(P(x^*))$. Luego, por (2.2.3) y (2.2.4), debe ser $f(x^*) = f(P(x^*))$ y $x^* \in \mathcal{R}$. \square

Como $x^* \in \mathcal{R}$, existe $t^* \in \mathbb{R}$ tal que $x^* = t^*v + b$ y por lo tanto

$$\min_{x \in \mathbb{R}^n} f(x) = \min_{x \in \mathcal{R}} f(x) = \min_{t \in \mathbb{R}} f(tv + b). \tag{2.2.5}$$

Usando que $a_i = t_i v + b$, obtenemos que $\|tv + b - a_i\| = \|(t - t_i)v\| = |t - t_i| \|v\| = |t - t_i|$ y por lo tanto

$$f(tv + b) = \sum_{i=1}^m w_i \|tv + b - a_i\| = \sum_{i=1}^m w_i |t - t_i|. \tag{2.2.6}$$

De (2.2.5) y (2.2.6), obtenemos que resolver el problema (2.2.1) para el caso donde los nodos son colineales es equivalente a resolver

$$\text{Minimizar } g(t) = \sum_{i=1}^m w_i |t - t_i|. \tag{2.2.7}$$

y se satisface que si t^* es el minimizador de (2.2.7), entonces $x^* = t^*v + b$.

Lo único que resta entonces es encontrar la forma de determinar t^* . Para hacerlo, seguiremos el razonamiento presentado por Love, Morris y Wesolowsky en el capítulo 2 de [LMW88]. El desarrollo planteado en el libro está dirigido a resolver el problema de Fermat-Weber en norma 1, el cual veremos más adelante, en la Sección 4.1. Sin embargo, puede adaptarse al problema (2.2.7) pues la estructura de ambos es la misma.

Primero notemos que la función $g : \mathbb{R} \rightarrow \mathbb{R}$ es continua y lineal a trozos. En efecto,

$$g(t) = \begin{cases} -\sum_{i=1}^m w_i(t - t_i) & \text{si } t \leq t_1 \\ \sum_{i=1}^j w_i(t - t_i) - \sum_{i=j+1}^m w_i(t - t_i) & \text{si } t_j \leq t \leq t_{j+1}, j = 1, \dots, m-1 \\ \sum_{i=1}^m w_i(t - t_i) & \text{si } t \geq t_m \end{cases} \quad (2.2.8)$$

Como es suma de las funciones convexas $w_i|t - t_i|$, la función g también resulta convexa. Como además $\lim_{|t| \rightarrow +\infty} g(t) = +\infty$, esto significa que tiene al menos un minimizador.

Por ser lineal a trozos, es diferenciable en todo punto salvo en los puntos de cambio, es decir, en t_1, \dots, t_m . Su derivada $g' : \mathbb{R} - \{t_i\}_{i=1}^m \rightarrow \mathbb{R}$ está dada por

$$g'(t) = \begin{cases} -\sum_{i=1}^m w_i & \text{si } t < t_1 \\ \sum_{i=1}^j w_i - \sum_{i=j+1}^m w_i & \text{si } t_j < t < t_{j+1}, j = 1, \dots, m-1 \\ \sum_{i=1}^m w_i & \text{si } t > t_m \end{cases} \quad (2.2.9)$$

Como los pesos w_i son positivos y

$$\sum_{i=1}^j w_i - \sum_{i=j+1}^m w_i = \left(\sum_{i=1}^{j-1} w_i - \sum_{i=j}^m w_i \right) + 2w_j,$$

entonces la función g' es no decreciente. Además, $g'(t) < 0$ para $t < t_1$ y $g'(t) > 0$ para $t > t_m$. Luego existe $s \in \{1, \dots, m\}$ tal que, para cualquier $t \neq t_i$ para todo i , se satisface que $g'(t) < 0$ si $t < t_s$ y $g'(t) \geq 0$ si $t > t_s$. Con esto en mente, podemos probar el siguiente resultado:

Teorema 2.2.4. *Sea $s \in \{1, \dots, m\}$ tal que*

$$\sum_{i=1}^{s-1} w_i - \sum_{i=s}^m w_i < 0 \quad (2.2.10)$$

y

$$\sum_{i=1}^s w_i - \sum_{i=s+1}^m w_i \geq 0. \quad (2.2.11)$$

Si la ecuación (2.2.11) se satisface como una desigualdad estricta, entonces la solución al problema (2.2.7) es $t^ = t_s$. Si se satisface como igualdad, entonces $t^* \in [t_s, t_{s+1}]$.*

Demostración. Estudiemos el comportamiento de la función g a través de su derivada. Por la elección de s , tenemos que $g'(t) < 0$ en el conjunto $\{t < t_s : t \neq t_i \text{ para todo } i\}$ y por lo tanto g es decreciente en ese conjunto. Por ser continua, resulta decreciente en toda la semirrecta $(-\infty, t_s]$.

Si la ecuación (2.2.11) satisface una desigualdad estricta, es decir, si

$$\sum_{i=1}^s w_i - \sum_{i=s+1}^m w_i > 0,$$

entonces $g'(t) > 0$ en el conjunto $\{t > t_s : t \neq t_i \text{ para todo } i\}$ y por lo tanto la función g es creciente en la semirrecta $[t_s, +\infty)$. Luego, como g decrece para valores menores que t_s y crece para valores mayores, debe ser el único minimizador, $t^* = t_s$.

Si en cambio se satisface la igualdad, es decir,

$$\sum_{i=1}^s w_i - \sum_{i=s+1}^m w_i = 0,$$

entonces $g'(t) = 0$ en el intervalo (t_s, t_{s+1}) y $g'(t) > 0$ en $\{t > t_{s+1} : t \neq t_i \text{ para todo } i\}$. Luego, g decrece en $(-\infty, t_s]$, es constante en $[t_s, t_{s+1}]$ y crece en $[t_{s+1}, +\infty)$, y por lo tanto todos los puntos en $[t_s, t_{s+1}]$ son minimizadores de g . \square

Esto concluye la estrategia para hallar los minimizadores del problema de Fermat-Weber para el caso de nodos colineales, y a su vez queda demostrado que puede haber una solución o una cantidad no numerable de soluciones.

Para el caso con nodos no colineales, su estudio es más complicado y no se conocen métodos exactos para resolverlo, salvo para los casos en el plano con tres o cuatro nodos, y algunos otros casos particulares sencillos. Existen otro tipo de métodos para encontrar soluciones, los cuales desarrollaremos más adelante. Primero, veremos el marco teórico necesario para hallar las soluciones.

De aquí en más asumiremos siempre que los nodos $a_1, \dots, a_m \in \mathbb{R}^n$ son puntos distintos y no colineales. Con esta nueva hipótesis, podemos probar la unicidad de la solución al problema (2.2.1). Para ello necesitaremos el siguiente lema:

Lema 2.2.5. *La función f es estrictamente convexa.*

Demostración. Ya hemos probado que la función f es convexa, es decir que

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \text{ para todo } x, y \in \mathbb{R}^n \text{ y } \alpha \in [0, 1].$$

Supongamos que existen $x \neq y$ y $0 < \alpha < 1$ tales que

$$f(\alpha x + (1 - \alpha)y) = \alpha f(x) + (1 - \alpha)f(y). \quad (2.2.12)$$

Como $f(\alpha x + (1 - \alpha)y) = \sum_{i=1}^m w_i \|(\alpha x + (1 - \alpha)y) - a_i\|$ y

$$\|\alpha x + (1 - \alpha)y - a_i\| = \|\alpha(x - a_i) + (1 - \alpha)(y - a_i)\| \leq \alpha \|x - a_i\| + (1 - \alpha) \|y - a_i\|$$

para todo $i \in \{1, \dots, m\}$, donde la igualdad vale si y sólo si $x - a_i = c_i(y - a_i)$ con $c_i > 0$, entonces (2.2.12) vale si y sólo si existen c_1, \dots, c_m tales que $x - a_i = c_i(y - a_i)$. Como $x \neq y$, debe cumplirse que $c_i \neq 1$ para todo $i = 1, \dots, m$. Esto significa que para todo $i \in \{1, \dots, m\}$,

$$a_i = \frac{1}{1 - c_i}(x - c_i y) = \frac{c_i}{1 - c_i}(x - y) + x,$$

y los nodos resultan colineales. Como estamos asumiendo que no lo son, entonces

$$f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y) \text{ para todo } x \neq y \text{ y } 0 < \alpha < 1$$

y f resulta estrictamente convexa. \square

Utilizando este resultado y el Teorema 2.2.2, es sencillo demostrar la unicidad del minimizador.

Teorema 2.2.6. *La función f tiene un único minimizador, que además es minimizador global.*

Demostración. Por el Teorema 2.2.2, f tiene un minimizador global. Como f es estrictamente convexa, entonces es el único minimizador. \square

Este teorema asegura la existencia de la solución pero no provee información respecto a cómo calcularla o caracterizarla. Sabemos que las funciones $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ son infinitamente diferenciables en todo punto salvo a_i , pues la función $\|\cdot\|$ no es diferenciable en el origen. Luego, la función $f = \sum_{i=1}^m w_i f_i$ resulta infinitamente diferenciable en todo punto salvo en \mathcal{A} por lo que no se satisfacen las hipótesis para la condición necesaria de optimalidad de primer orden. Sin embargo, como la no diferenciable ocurre sólo en los nodos, si la solución x^* no es un nodo, existe un entorno de la solución donde la función sí es diferenciable. Para ese caso, la condición necesaria de optimalidad de primer orden nos dice que $\nabla f(x^*) = 0$. Para el caso en que la solución coincida con un nodo, podremos caracterizar a la solución de manera similar. El siguiente teorema fue probado de tres maneras distintas por Weiszfeld en 1937 [Wei37], y nuevamente por Kuhn y Kuenne en 1962 [KK62].

Teorema 2.2.7. *Si $a_1, \dots, a_m \in \mathbb{R}^n$ son puntos distintos no colineales, y $w_1, \dots, w_m > 0$, entonces existe una única solución al problema (2.2.1). Además, dado $x^* \in \mathbb{R}^n$, se satisface:*

a. *Si $x^* \notin \mathcal{A}$, entonces x^* minimiza f si y sólo si*

$$\nabla f(x^*) = \sum_{i=1}^m w_i \frac{x^* - a_i}{\|x^* - a_i\|} = 0. \quad (2.2.13)$$

b. *Si $x^* = a_k$ para algún $k \in \{1, \dots, m\}$, entonces x^* minimiza f si y sólo si*

$$\left\| \sum_{i=1, i \neq k}^m w_i \frac{x^* - a_i}{\|x^* - a_i\|} \right\| \leq w_k. \quad (2.2.14)$$

Demostración. La demostración de la existencia y unicidad de la solución ya fue presentada (Teorema 2.2.2).

En el caso a., en el que x^* no es un nodo, sabemos que la función f es convexa y diferenciable en un entorno de x^* y por lo tanto $\nabla f(x^*) = 0$ es condición necesaria y suficiente de optimalidad de primer orden.

Para $x^* = a_k$ con $k \in \{1, \dots, m\}$, si bien no existe el gradiente de la función f , podemos analizar las derivadas direccionales de f en a_k . Dado $v \in \mathbb{R}^n$ con $\|v\| = 1$, sea $f_v : \mathbb{R}^+ \rightarrow \mathbb{R}$ dada por $f_v(t) = f(a_k + tv)$. La derivada direccional de f a partir de a_k en la dirección v está definida como

$$D_v f(a_k) = \lim_{t \rightarrow 0^+} \frac{f_v(t) - f_v(0)}{t}. \quad (2.2.15)$$

Como f es diferenciable fuera de los nodos, entonces f_v es diferenciable en $(0, s_0)$ para algún s_0 lo suficientemente pequeño, por lo que la ecuación (2.2.15) es equivalente a:

$$D_v f(a_k) = \lim_{s \rightarrow 0^+} \frac{df_v}{dt}(s).$$

Primero estudiemos la derivada de f_v :

$$\frac{df_v}{dt}(s) = \frac{d}{dt} \Big|_s \left(\sum_{i=1}^m w_i \|a_k + tv - a_i\| \right).$$

Si $i = k$, entonces $w_i \|a_k + tv - a_i\| = w_k \|tv\| = w_k |t| \|v\| = w_k t$. Por lo tanto,

$$\frac{df_v}{dt}(s) = \frac{d}{dt} \Big|_s \left(w_k t + \sum_{\substack{i=1 \\ i \neq k}}^m w_i \|a_k - a_i + tv\| \right) = w_k + \sum_{\substack{i=1 \\ i \neq k}}^m w_i \left\langle \frac{a_k - a_i + sv}{\|a_k - a_i + sv\|}, v \right\rangle.$$

Tomando límite en s , obtenemos

$$\begin{aligned} D_v f(a_k) &= \lim_{s \rightarrow 0^+} \left[w_k + \sum_{\substack{i=1 \\ i \neq k}}^m w_i \left\langle \frac{a_k - a_i + sv}{\|a_k - a_i + sv\|}, v \right\rangle \right] \\ &= w_k + \sum_{\substack{i=1 \\ i \neq k}}^m w_i \left\langle \frac{a_k - a_i}{\|a_k - a_i\|}, v \right\rangle \\ &= w_k + \underbrace{\left\langle \sum_{\substack{i=1 \\ i \neq k}}^m w_i \frac{a_k - a_i}{\|a_k - a_i\|}, v \right\rangle}_{R_k}. \end{aligned}$$

Esto nos asegura que existen las derivadas direccionales de f en todas las direcciones, y están dadas por

$$D_v f(a_k) = w_k + \langle R_k, v \rangle, \text{ donde } R_k = \sum_{\substack{i=1 \\ i \neq k}}^m w_i \frac{a_k - a_i}{\|a_k - a_i\|}. \quad (2.2.16)$$

Gracias a esto podemos afirmar que la función f tiene un mínimo en a_k si y sólo si no hay direcciones de descenso de f a partir de a_k , es decir, si y sólo si $D_v f(a_k) \geq 0$ para todo $v \in \mathbb{R}^n$ tal que $\|v\| = 1$. Equivalentemente, a_k es minimizador si y sólo si

$$\min_{\substack{v \in \mathbb{R}^n \\ \|v\|=1}} D_v f(a_k) \geq 0.$$

Debido a que $w_k > 0$ es constante, el mínimo sobre v de $D_v f(a_k)$ se alcanza cuando se minimiza $\langle R_k, v \rangle$, esto es, cuando $v = -\frac{R_k}{\|R_k\|}$. Por lo tanto a_k es minimizador de f si y sólo si

$$0 \leq D_{-\frac{R_k}{\|R_k\|}} f(a_k) = w_k + \left\langle R_k, -\frac{R_k}{\|R_k\|} \right\rangle = w_k - \|R_k\| \Leftrightarrow \|R_k\| = \left\| \sum_{\substack{i=1 \\ i \neq k}}^m w_i \frac{a_k - a_i}{\|a_k - a_i\|} \right\| \leq w_k.$$

□

Corolario 2.2.8. *El minimizador x^* de f está en la cápsula convexa de los nodos:*

$$x^* \in \mathcal{C}(\mathcal{A}) = \left\{ \sum_{i=1}^m \alpha_i a_i : 0 \leq \alpha_i \leq 1 \wedge \sum_{i=1}^m \alpha_i = 1 \right\}.$$

Demostración. Si $x^* = a_k$ la afirmación vale pues $a_k = \sum_{i=1}^m \alpha_i a_i$ con $\alpha_k = 1$ y $\alpha_i = 0$ si $i \neq k$.

Si x^* no es un nodo, por la ecuación (2.2.13) tenemos que

$$\sum_{i=1}^m w_i \frac{x^* - a_i}{\|x^* - a_i\|} = 0.$$

Manipulando esta ecuación obtenemos:

$$\begin{aligned} \sum_{i=1}^m w_i \frac{x^*}{\|x^* - a_i\|} &= \sum_{i=1}^m w_i \frac{a_i}{\|x^* - a_i\|} \\ x^* \sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} &= \sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} a_i \end{aligned}$$

Entonces

$$x^* = \frac{\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} a_i}{\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|}}. \quad (2.2.17)$$

Luego, obtenemos que $x^* = \sum_{i=1}^m \alpha_i a_i$ tomando para cada $i = 1, \dots, m$

$$\alpha_i = \frac{\frac{w_i}{\|x^* - a_i\|}}{\sum_{j=1}^m \frac{w_j}{\|x^* - a_j\|}}.$$

Como $\alpha_i > 0$ para todo i , y

$$\sum_{i=1}^m \alpha_i = \sum_{i=1}^m \frac{\frac{w_i}{\|x^* - a_i\|}}{\sum_{j=1}^m \frac{w_j}{\|x^* - a_j\|}} = 1,$$

entonces $x^* \in \mathcal{C}(\mathcal{A})$. □

La ecuación (2.2.17) es un resultado importante, pues da la idea para el algoritmo de Weiszfeld. En éste, se aplica el método de punto fijo a una transformación $T : \mathbb{R}^n - \mathcal{A} \rightarrow \mathbb{R}^n$ tal que $T(x^*) = x^*$, definida como:

$$T(y) = \frac{\sum_{i=1}^m \frac{w_i}{\|y - a_i\|} a_i}{\sum_{i=1}^m \frac{w_i}{\|y - a_i\|}}.$$

Capítulo 3

El algoritmo de Weiszfeld

3.1. Historia del algoritmo

Como mencionamos en la Sección 2.1, el algoritmo de Weiszfeld, que surgió en 1937 y fue redescubierto en múltiples ocasiones, es un método para hallar la solución al problema de Fermat-Weber generalizado (2.2.1).

En el año 1937, un matemático húngaro de tan sólo 16 años, Endre Vaszanyi Weiszfeld [Wei37], escribió un artículo dando tres demostraciones distintas al Teorema 2.2.7 para el caso sin pesos, es decir, $w_1 = \dots = w_m = 1$, las cuales pueden extenderse fácilmente al caso generalizado. La primera de ellas, consistía en construir una sucesión de puntos que convergía a la solución, la cual definió del siguiente modo:

Afirmación. Sean m puntos distintos en el espacio, a_1, a_2, \dots, a_m no colineales, y sea $x^{(0)}$ un punto arbitrario fuera de \mathcal{A} .

Para cada $k \in \mathbb{N} \cup \{0\}$, asignemos a los puntos a_1, a_2, \dots, a_m las respectivas masas

$$\frac{1}{\|x^{(k)} - a_1\|}, \frac{1}{\|x^{(k)} - a_2\|}, \dots, \frac{1}{\|x^{(k)} - a_m\|},$$

y definamos $x^{(k+1)}$ como el centro de gravedad de estas masas.

La sucesión obtenida $x^{(0)}, x^{(1)}, x^{(2)}, \dots$ converge al punto que minimiza la suma de las distancias a los puntos a_1, a_2, \dots, a_m , independientemente del punto $x^{(0)}$.

Esta sucesión es la que dio origen al algoritmo de Weiszfeld, el cual estudiaremos en profundidad más adelante. Un resultado importante probado en este artículo fue la monotonicidad de la sucesión de valores funcionales, es decir, que $f(x^{(k+1)}) \leq f(x^{(k)})$ para todo $k \in \mathbb{N}$.

El inconveniente con su demostración es que la sucesión no estaba bien definida. Weiszfeld aseguró que convergía independientemente del punto inicial, asumiendo que $x^{(k)} \notin \mathcal{A}$ para todo $k \in \mathbb{N}$. Sin embargo, podría ocurrir que $x^{(k)} \in \mathcal{A}$, llevando a la necesidad de dividir por cero para calcular $x^{(k+1)}$.

La sucesión fue redescubierta en varias ocasiones, pero el trabajo que más impacto tuvo fue el de Kuhn y Kuenne [KK62], en 1962, el cual considera el problema generalizado al caso con pesos para el problema en \mathbb{R}^2 . En él, definen una transformación $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ por

$$T(x) = \begin{cases} \frac{\sum_{i=1}^m \frac{w_i a_i}{\|x - a_i\|}}{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|}} & \text{si } x \notin \mathcal{A} \\ x & \text{si } x \in \mathcal{A} \end{cases} \quad (3.1.1)$$

y, partiendo de algún $x^{(0)} \in \mathbb{R}^2$ construyen la sucesión $x^{(k)} = T(x^{(k-1)})$, $k \in \mathbb{N}$. Esta transformación, para el caso con $w_i = 1$ para todo i , genera la sucesión propuesta por Weiszfeld. Los autores

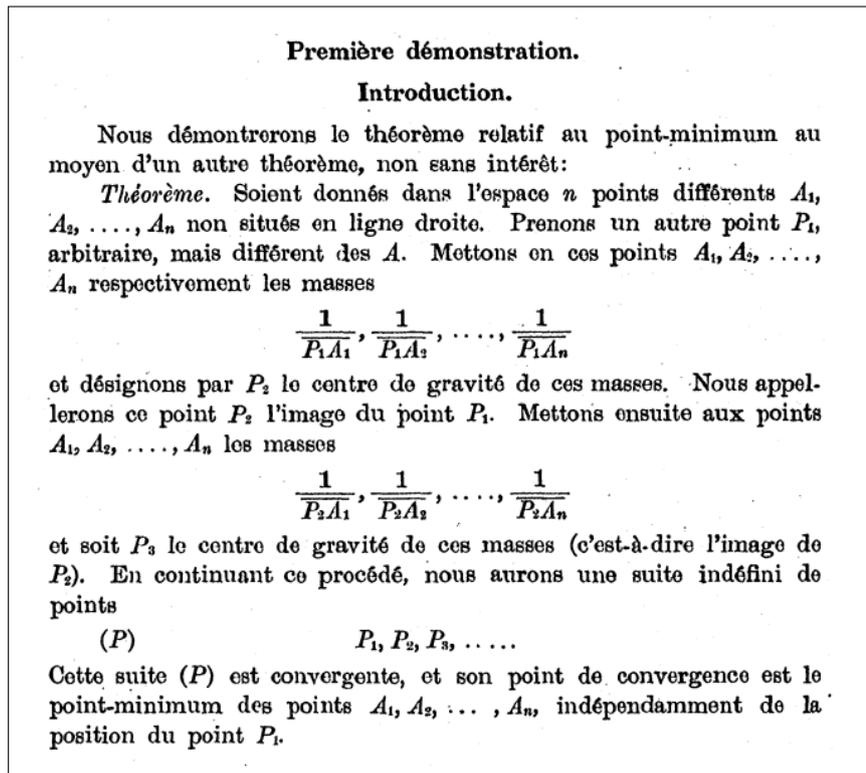


Figura 3.1: Fragmento del artículo “Sur le point pour lequel la somme des distances de n points donnés est minimum”

desconocían el trabajo de Weiszfeld antes de presentar su artículo, por lo cual al publicarlo anexaron el reconocimiento a Weiszfeld como el primer autor del algoritmo, pero aclarando el error que éste había cometido.

En 1973, Kuhn [Kuh73] presentó la generalización del algoritmo a \mathbb{R}^n , con $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida como en (3.1.1). En este nuevo trabajo, probó la convergencia de la sucesión para el caso $x^{(k)} \notin \mathcal{A}$ para todo $k \in \mathbb{N}$, explicó el error de Weiszfeld y demostró que efectivamente puede ocurrir que una de las iteraciones coincida con un nodo, mediante el siguiente ejemplo:

Ejemplo 3.1.1. Consideremos los puntos $a_1 = (-2, 0)$, $a_2 = (-1, 0)$, $a_3 = (1, 0)$, $a_4 = (2, 0)$, $a_5 = (0, 1)$ y $a_6 = (0, -1)$, y supongamos $w_i = 1$ para todo i . Es fácil probar que $\nabla f(0, 0) = 0$ y por lo tanto $x^* = (0, 0)$.

Recordemos que el operador T está definido por

$$T(x) = \frac{1}{\sum_{i=1}^6 \frac{1}{\|x - a_i\|}} \sum_{i=1}^6 \frac{a_i}{\|x - a_i\|}.$$

Si $x = (\alpha, 0) \notin \mathcal{A}$, entonces

$$\sum_{i=1}^6 \frac{1}{\|x - a_i\|} = \frac{1}{|\alpha + 2|} + \frac{1}{|\alpha + 1|} + \frac{1}{|\alpha - 1|} + \frac{1}{|\alpha - 2|} + \frac{2}{\sqrt{\alpha^2 + 1}}.$$

Luego

$$\begin{aligned} \sum_{i=1}^6 \frac{a_i}{\|x - a_i\|} &= \frac{(-2, 0)}{|\alpha + 2|} + \frac{(-1, 0)}{|\alpha + 1|} + \frac{(1, 0)}{|\alpha - 1|} + \frac{(2, 0)}{|\alpha - 2|} + \frac{(0, 1)}{\sqrt{\alpha^2 + 1}} + \frac{(0, -1)}{\sqrt{\alpha^2 + 1}} \\ &= \left(\frac{-2}{|\alpha + 2|} + \frac{-1}{|\alpha + 1|} + \frac{1}{|\alpha - 1|} + \frac{2}{|\alpha - 2|}, 0 \right) \end{aligned}$$

y

$$T((\alpha, 0)) = \left(\frac{\frac{-2}{|\alpha+2|} + \frac{-1}{|\alpha+1|} + \frac{1}{|\alpha-1|} + \frac{2}{|\alpha-2|}}{\frac{1}{|\alpha+2|} + \frac{1}{|\alpha+1|} + \frac{1}{|\alpha-1|} + \frac{1}{|\alpha-2|} + \frac{2}{\sqrt{\alpha^2+1}}}, 0 \right) =: (g(\alpha), 0).$$

Puede probarse que existe α_0 , $1 < \alpha_0 < 2$. tal que $g(\alpha_0) = 1$. Luego, si $x^{(0)} = (\alpha_0, 0)$ obtenemos que

$$x^{(1)} = T(x^{(0)}) = (g(\alpha_0), 0) = (1, 0) = a_3 \in \mathcal{A}$$

siendo que $x^{(0)} \notin \mathcal{A}$ y $a_3 \neq x^*$.

A pesar de haber hallado un ejemplo desfavorable, Kuhn [Kuh73] aseguró que se trataba de una situación poco probable, y demostró el siguiente teorema:

Teorema 3.1.2. *Salvo para una cantidad numerable de puntos iniciales $x^{(0)}$, la sucesión dada por $x^{(k)} = (T(x^{(0)}))^k$ converge a la solución x^* .*

Demostración. Si ningún $x^{(k)}$ coincide con un nodo, entonces la sucesión converge (como se probará en el Teorema 3.2.6).

Para encontrar a_i a través de T , debemos resolver ecuaciones algebraicas, por lo cual la ecuación $T(x) = a_i$ tiene finitas soluciones. Entonces para $k \in \mathbb{N}$, el conjunto

$$\{x^{(0)} : T^k(x^{(0)}) = a_i \text{ para algún } i = 1, \dots, m\}$$

es finito y el conjunto de puntos iniciales para los cuales la sucesión no converge,

$$\{x^{(0)} : T^k(x^{(0)}) = a_i \text{ para algún } i = 1, \dots, m \text{ y } k \in \mathbb{N}\},$$

resulta numerable. □

Pero este teorema es falso, pues también Kuhn cometió un error en su demostración. Éste fue puesto en evidencia por Chandrasekaran y Tamir en 1989 [CT89], quienes demostraron dando dos ejemplos que puede haber un conjunto continuo de soluciones a la ecuación $T(x) = a_i$, a pesar de ser ecuaciones algebraicas. En el primer ejemplo, la solución al problema es un nodo, y hay una recta cuya imagen es ese nodo.

Ejemplo 3.1.3. Consideremos el problema sin pesos en \mathbb{R}^3 dado por $a_1 = (1, 0, 0)$, $a_2 = (0, 1, 0)$, $a_3 = (-1, -1, 0)$, $a_4 = (0, 0, 0)$. El nodo a_4 es solución, lo cual puede verificarse mediante la ecuación (2.2.14).

$$R_4 = \sum_{i=1}^3 \frac{a_4 - a_i}{\|a_4 - a_i\|} = (-1, 0, 0) + (0, -1, 0) + \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right) = \left(\frac{1}{\sqrt{2}} - 1, \frac{1}{\sqrt{2}} - 1, 0 \right)$$

$$\|R_4\| = \left(\frac{1}{\sqrt{2}} - 1 \right)^2 + \left(\frac{1}{\sqrt{2}} - 1 \right)^2 = 2 \left(\frac{1 - \sqrt{2}}{\sqrt{2}} \right)^2 = 2 \frac{(1 - \sqrt{2})^2}{2} = (\sqrt{2} - 1)^2 < 1$$

Chandrasekaran y Tamir aseguraron que si $x = (-\frac{1}{6}, -\frac{1}{6}, \alpha)$ con $\alpha \in \mathbb{R}$, entonces $T(x) = a_4$.

En efecto, notemos que

$$\|x - a_1\| = \sqrt{\left(-\frac{1}{6} - 1\right)^2 + \left(-\frac{1}{6}\right)^2 + \alpha^2} = \sqrt{\frac{49}{36} + \frac{1}{36} + \alpha^2} = \sqrt{\frac{50}{36} + \alpha^2} =: \beta,$$

$$\|x - a_2\| = \sqrt{\left(-\frac{1}{6}\right)^2 + \left(-\frac{1}{6} - 1\right)^2 + \alpha^2} = \sqrt{\frac{1}{36} + \frac{49}{36} + \alpha^2} = \sqrt{\frac{50}{36} + \alpha^2} = \beta,$$

$$\|x - a_3\| = \sqrt{\left(-\frac{1}{6} + 1\right)^2 + \left(-\frac{1}{6} + 1\right)^2 + \alpha^2} = \sqrt{\frac{25}{36} + \frac{25}{36} + \alpha^2} = \sqrt{\frac{50}{36} + \alpha^2} = \beta,$$

$$\text{y } \|x - a_4\| = \sqrt{\left(-\frac{1}{6}\right)^2 + \left(-\frac{1}{6}\right)^2 + \alpha^2} = \sqrt{\frac{1}{36} + \frac{1}{36} + \alpha^2} = \sqrt{\frac{2}{36} + \alpha^2}$$

Entonces

$$\sum_{i=1}^4 \frac{a_i}{\|x - a_i\|} = \frac{(1, 0, 0)}{\beta} + \frac{(0, 1, 0)}{\beta} + \frac{(-1, -1, 0)}{\beta} + \frac{(0, 0, 0)}{\|x - a_4\|} = (0, 0, 0)$$

y por lo tanto

$$T(x) = \frac{1}{\sum_{i=1}^4 \frac{1}{\|x - a_i\|}} \sum_{i=1}^4 \frac{a_i}{\|x - a_i\|} = \frac{1}{\sum_{i=1}^4 \frac{1}{\|x - a_i\|}} (0, 0, 0) = (0, 0, 0).$$

Luego $T(x) = a_4 = x^*$ para todo x en la recta $\mathcal{R} = \left\{-\frac{1}{6}, -\frac{1}{6}, 0\right\} + \alpha(0, 0, 1) : \alpha \in \mathbb{R}\}$.

En el segundo ejemplo, se muestra que puede ocurrir que la transformación lleve a un nodo para todo punto en un conjunto continuo sin que el nodo sea solución.

Ejemplo 3.1.4. Consideremos el problema en \mathbb{R}^3 dado por los nodos $a_1 = (1, 0, 0)$, $a_2 = (-1, 0, 0)$, $a_3 = (0, 0, 0)$, $a_4 = (0, 2, 0)$ y $a_5 = (0, -2, 0)$, y los pesos $w_1 = w_2 = w_3 = w_5 = 1$ y $w_4 = 3$. El nodo a_3 no es solución, pues

$$R_3 = -\frac{a_1}{\|a_1\|} - \frac{a_2}{\|a_2\|} - 3\frac{a_4}{\|a_4\|} - \frac{a_5}{\|a_5\|} = -(1, 0, 0) + (1, 0, 0) - 3(0, 1, 0) + (0, 1, 0) = (0, -2, 0)$$

y por lo tanto $\|R_3\| = 2 > 1 = w_3$. Veremos que $T(x) = a_3$ tiene un conjunto continuo de soluciones. Como $a_3 = (0, 0, 0)$, entonces $T(x) = a_3$ si y sólo si

$$\sum_{i=1}^5 \frac{w_i a_i}{\|x - a_i\|} = (0, 0, 0).$$

Supongamos que $x = (0, \alpha, \beta)$. Entonces

$$\begin{aligned} \sum_{i=1}^5 \frac{w_i a_i}{\|x - a_i\|} &= \frac{(1, 0, 0)}{\|(-1, \alpha, \beta)\|} + \frac{(-1, 0, 0)}{\|(1, \alpha, \beta)\|} + \frac{(0, 0, 0)}{\|(0, \alpha, \beta)\|} + \frac{3(0, 2, 0)}{\|(0, \alpha - 2, \beta)\|} + \frac{(0, -2, 0)}{\|(0, \alpha + 2, \beta)\|} \\ &= \frac{3(0, 2, 0)}{\|(0, \alpha - 2, \beta)\|} + \frac{(0, -2, 0)}{\|(0, \alpha + 2, \beta)\|} \\ &= \left(0, \frac{6}{\|(0, \alpha - 2, \beta)\|} - \frac{2}{\|(0, \alpha + 2, \beta)\|}, 0\right) \end{aligned}$$

Para que $T(x) = 0$, debe ocurrir

$$0 = \frac{6}{\|(0, \alpha - 2, \beta)\|} - \frac{2}{\|(0, \alpha + 2, \beta)\|} = \frac{6}{\sqrt{(\alpha - 2)^2 + \beta^2}} - \frac{2}{\sqrt{(\alpha + 2)^2 + \beta^2}}.$$

Equivalentemente,

$$\begin{aligned}
6\sqrt{(\alpha+2)^2+\beta^2} &= 2\sqrt{(\alpha-2)^2+\beta^2} \\
6^2((\alpha+2)^2+\beta^2) &= 2^2((\alpha-2)^2+\beta^2) \\
36\alpha^2+144\alpha+144+36\beta^2 &= 4\alpha^2-16\alpha+16+4\beta^2 \\
32\alpha^2+160\alpha+32\beta^2 &= -128 \\
\alpha^2+5\alpha+\left(\frac{5}{2}\right)^2-\left(\frac{5}{2}\right)^2+\beta^2 &= -4 \\
\left(\alpha+\frac{5}{2}\right)^2+\beta^2 &= -4+\frac{25}{4} \\
\left(\alpha+\frac{5}{2}\right)^2+\beta^2 &= \frac{9}{4}
\end{aligned}$$

Por lo tanto, $T(x) = a_3$ para todo x en la circunferencia $\mathcal{C} = \{(0, \alpha, \beta) : (\alpha + \frac{5}{2})^2 + \beta^2 = \frac{9}{4}\}$.

Basados en estos ejemplos, Chandrasekaran y Tamir [CT89] plantearon la siguiente conjetura:

Conjetura. *Si la cápsula convexa de \mathcal{A} , $\mathcal{C}(\mathcal{A})$, tiene dimensión completa n , entonces la ecuación $T(x) = a_i$ tiene un número finito de soluciones para $i = 1, \dots, m$.*

Dicho de otra manera, para cada $i = 1, \dots, m$ hay una cantidad finita de soluciones en el menor espacio afín que contiene a $\mathcal{C}(\mathcal{A})$.

Si esta conjetura fuera cierta, se podría asegurar que si se comienza el algoritmo de Weiszfeld en un punto de la cápsula convexa de \mathcal{A} , la convergencia está garantizada salvo para una cantidad numerable de puntos iniciales.

En 1995, Brimberg [Bri95] formuló y demostró el siguiente teorema:

Teorema 3.1.5. *El conjunto de puntos iniciales $\{x^{(0)}\}$ que detendrá la secuencia generada por el algoritmo de Weiszfeld en un nodo a_i después de una cantidad finita de iteraciones es numerable si y sólo si $\mathcal{C}(\mathcal{A})$ tiene dimensión completa n .*

Sin embargo, en 2002, Cánovas, Cañavate y Marín [CCM02] descubrieron errores en su demostración. Por un lado, encontraron dos contraejemplos al “sólo si”, es decir, dos ejemplos en los que $\mathcal{C}(\mathcal{A})$ no tiene dimensión completa y sin embargo hay a lo sumo una cantidad numerable de puntos iniciales que detienen el algoritmo en un nodo.

Ejemplo 3.1.6. Sean $a_1 = (1, 0, 0)$, $a_2 = (0, 1, 0)$ y $a_3 = (0, 0, 0)$. Sin importar los pesos w_1, w_2 y w_3 , $T(x) = (T_1(x), T_2(x), 0)$ con

$$T_s(x) = \frac{\frac{w_s}{\|x-a_s\|}}{\sum_{i=1}^3 \frac{w_i}{\|x-a_i\|}} > 0 \quad s = 1, 2 \text{ para todo } x \notin \mathcal{A}.$$

Luego $T(x) = a_i$ si y sólo si $x = a_i$ y por lo tanto los puntos iniciales que finalizan el algoritmo en un nodo luego de un número finito de iteraciones son sólo los nodos. Sin embargo, $\mathcal{C}(\mathcal{A})$ está contenida en el plano $x_3 = 0$, por lo que tiene dimensión 2.

En este primer ejemplo, cada nodo era un vértice de $\mathcal{C}(\mathcal{A})$. En el siguiente ejemplo, uno de los nodos está en el interior de $\mathcal{C}(\mathcal{A})$.

Ejemplo 3.1.7. Sean $a_1 = (1, 0, 0)$, $a_2 = (0, 1, 0)$, $a_3 = (-1, -1, 0)$ y $a_4 = (0, 0, 0)$ con pesos $w_1 = w_2 = w_4 = 1$ y $w_3 = 4$.

$$T(x) = \frac{1}{\sum_{i=1}^3 \frac{w_i}{\|x - a_i\|}} \left(\frac{1}{\|x - a_1\|} - \frac{4}{\|x - a_3\|}, \frac{1}{\|x - a_1\|} - \frac{4}{\|x - a_3\|}, 0 \right).$$

Luego la ecuación $T(x) = a_4 = (0, 0, 0)$ es equivalente a

$$\begin{cases} \|x - a_3\| = 4 \|x - a_1\| \\ \|x - a_3\| = 4 \|x - a_2\| \end{cases}$$

La primer igualdad equivale a $(x_1 + 1)^2 + (x_2 + 1)^2 + x_3^2 = 16((x_1 - 1)^2 + x_2^2 + x_3^2)$, cuya solución es la esfera de centro $(\frac{17}{15}, \frac{1}{15}, 0)$ y radio $\frac{4\sqrt{5}}{15}$. Análogamente, la segunda igualdad tiene como solución a la esfera de centro $(\frac{1}{15}, \frac{17}{15}, 0)$ y radio $\frac{4\sqrt{5}}{15}$. Como los puntos $x \notin \mathcal{A}$ que satisfagan $T(x) = a_4$ deben satisfacer ambas igualdades, entonces deben estar en la intersección de ambas esferas, la cual es vacía. Por lo tanto, la ecuación $T(x) = a_4$ tiene como única solución al punto $x = a_4$. Al igual que en el ejemplo anterior, $\mathcal{C}(\mathcal{A})$ está contenido en el plano $x_3 = 0$, por lo que su dimensión no es completa.

Esto demuestra que el “sólo si” del Teorema 3.1.5 es falso. Por otro lado, los autores también encontraron un error en la demostración de la primer parte del teorema, es decir, de la conjetura. Una parte esencial de la demostración de Brimberg[Bri95] era lo siguiente:

Afirmación. Si una función $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de clase C^1 verifica que su Jacobiana $T'(x)$ es invertible para todo $x \in \mathbb{R}^n$ salvo un conjunto de medida cero, entonces $\{x : T(x) = b\}$ es a lo sumo numerable para todo $b \in \mathbb{R}^n$.

El siguiente ejemplo, dado por Cánovas, Cañavate y Marín [CCM02], muestra que esta afirmación no necesariamente es verdadera.

Ejemplo 3.1.8. Sea $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por $g(x) = (x_1^2, x_1 e^{x_2}, \dots, x_1 e^{x_n})$. Esta función es de clase C^1 y su matriz Jacobiana es

$$J(x) = \begin{pmatrix} 2x_1 & e^{x_2} & e^{x_3} & \dots & e^{x_n} \\ 0 & x_1 e^{x_2} & 0 & \dots & 0 \\ 0 & 0 & x_1 e^{x_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & x_1 e^{x_n} \end{pmatrix}$$

El determinante de $J(x)$ es $\det(J(x)) = 2x_1^n e^{x_2 + x_3 + \dots + x_n}$, el cual se anula sólo en el hiperplano $\mathcal{H} = \{x \in \mathbb{R}^n : x_1 = 0\}$. Luego es invertible en todo punto salvo en \mathcal{H} , que tiene medida cero. Sin embargo, el conjunto $\{x : g(x) = (0, \dots, 0)\}$ es el hiperplano \mathcal{H} , el cual no es numerable.

Este ejemplo muestra que la demostración presentada por Brimberg es errónea. Sin embargo en 2003, Brimberg [Bri03] logró probar finalmente la conjetura de Chandrasekaran y Tamir. Este artículo presenta tres resultados importantes.

El primer resultado que demostró no depende de la dimensión de $\mathcal{C}(\mathcal{A})$ sino de la ubicación de los nodos en la cápsula convexa, y es muy útil pues presenta una situación en la que no hay puntos iniciales $x^{(0)} \in \mathbb{R}^n - \mathcal{A}$ que lleven a que el algoritmo alcance un nodo en alguna iteración.

Teorema 3.1.9. *Si todos los nodos están en la frontera de $\mathcal{C}(\mathcal{A})$, entonces el conjunto*

$$\{x \in \mathbb{R}^n - \mathcal{A} : T(x) \in \mathcal{A}\}$$

es vacío.

Demostración. La transformación T dada por (3.1.1) puede reescribirse, para los puntos $x \in \mathbb{R}^n - \mathcal{A}$, como

$$T(x) = \sum_{i=1}^m \frac{\frac{w_i}{\|x-a_i\|}}{\sum_{s=1}^m \frac{w_s}{\|x-a_s\|}} a_i.$$

Como $0 < \frac{w_i}{\|x-a_i\|}$ para todo $i = 1, \dots, m$, entonces $0 < \frac{\frac{w_i}{\|x-a_i\|}}{\sum_{s=1}^m \frac{w_s}{\|x-a_s\|}} < 1$ para todo i .

Como los nodos están en la frontera de $\mathcal{C}(\mathcal{A})$ y $T(x) = \sum_{i=1}^m \alpha_i(x) a_i$ con $0 < \alpha_i(x) < 1$, esto implica que $T(x)$ está en el interior de $\mathcal{C}(\mathcal{A})$ y por lo tanto $T(x) \neq a_i$ para todo $i = 1, \dots, m$. \square

Esto a su vez nos asegura que para que el sistema $T(x) = a_i$ con $i \in \{1, \dots, m\}$ fijo tenga solución en $\mathbb{R}^n - \mathcal{A}$, debe satisfacerse que a_i esté en el interior de $\mathcal{C}(\mathcal{A})$, es decir

$$\{x \in \mathbb{R}^n - \mathcal{A} : T(x) \in \mathcal{A}\} = \{x \in \mathbb{R}^n - \mathcal{A} : T(x) = a_i \text{ y } a_i \in \mathcal{A} \text{ está en el interior de } \mathcal{C}(\mathcal{A})\}.$$

El segundo resultado del artículo, el cual el autor presentó primero con ejemplos y luego reforzó analizando la jacobiana del operador T , es que si $\mathcal{C}(\mathcal{A})$ no tiene dimensión completa, la cantidad de puntos iniciales “malos” puede ser nula, finita, numerable o incluso no numerable. Esto prueba, al igual que los ejemplos (3.1.6) y (3.1.7), que la recíproca del Teorema 3.1.5 es falsa.

El tercer y último resultado que Brimberg demostró en [Bri03] fue la conjetura de Chandrasekaran y Tamir. Para ello, se refirió primero a un lema que había probado inicialmente en [Bri95], sobre la jacobiana de T .

Lema 3.1.10. *Si la cápsula convexa de \mathcal{A} tiene dimensión completa n , entonces la matriz jacobiana de T , $J(x)$, es invertible para todo $x \in \mathbb{R}^n$, salvo un subconjunto de medida cero.*

Con este resultado, Brimberg prueba la conjetura de Chandrasekaran y Tamir, la cual reformuló del siguiente modo:

Teorema 3.1.11. *Si la cápsula convexa de \mathcal{A} tiene dimensión completa, entonces el conjunto de puntos iniciales que causan que el algoritmo de Weiszfeld alcance un nodo luego de una cantidad finita de iteraciones es a lo sumo numerable.*

El resultado de Chandrasekaran y Tamir nos asegura que eligiendo $x^{(0)} \in \mathcal{C}(\mathcal{A})$ en lugar de cualquier punto en $\mathbb{R}^n - \mathcal{A}$, podemos garantizar que a lo sumo en numerables ocasiones esto lleva a detener el algoritmo en un nodo, es decir que la probabilidad de elegir un punto problemático es nula. Sin embargo, muchos autores han trabajado sobre variaciones del algoritmo para mejorar aún más este resultado. Estas variaciones garantizan que el algoritmo nunca se detendrá sin llegar a un nodo. Algunas de ellas no dejan lugar a elegir el punto inicial sino que lo construyen de manera de garantizar la convergencia. Otras redefinen el operador en los nodos para que siga avanzando el método. La principal idea en ambas versiones consiste en encontrar una dirección de descenso a partir de los nodos, y definir

$$T(a_k) = \begin{cases} a_k & \text{si } a_k \text{ es solución} \\ S_k \text{ tal que } f(S_k) < f(a_k) & \text{caso contrario.} \end{cases}$$

Ostreich [OJ78] desarrolló un algoritmo de este tipo en 1978, tomando $T(a_k) = a_k + td$ donde $d = -\frac{R_k}{\|R_k\|}$ es la dirección de máximo descenso, como se probó en la demostración del ítem *b.* del Teorema 2.2.7, y $t > 0$ es alguna longitud de paso que puede variar. Una idea similar la plantean Vardi y Zhang en 2001 [VZ01], con longitud de paso fija. Ambos son mencionados en el trabajo de Beck y Sabach de 2015 [BS15], donde además demuestran que el paso elegido por Vardi y Zhang efectivamente es un buen paso. Profundizaremos sobre estas variantes en la Sección 3.3.

3.2. Resultados de convergencia

A continuación plantearemos y demostraremos los teoremas de convergencia del algoritmo de Weiszfeld. Primero recordemos la formulación del problema (2.2.1) y del algoritmo.

El problema de Fermat-Weber generalizado:

Dados $\mathcal{A} = \{a_1, \dots, a_m\} \subset \mathbb{R}^n$ y $w_1, \dots, w_m > 0$, el problema de Fermat-Weber generalizado es

$$\text{Minimizar } f(x) = \sum_{i=1}^m w_i \|x - a_i\|. \quad (3.2.1)$$

Como probamos en el Teorema 2.2.7, si el minimizador x^* de f no es un nodo, se satisface que

$$\nabla f(x^*) = \sum_{i=1}^m w_i \frac{x^* - a_i}{\|x^* - a_i\|} = 0.$$

Si manipulamos esta igualdad como en la demostración del Corolario 2.2.8, obtenemos la igualdad (2.2.17):

$$x^* = \frac{\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} a_i}{\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|}}.$$

Con esta igualdad en mente, Weiszfeld construyó la función $T : \mathbb{R}^n - \mathcal{A} \rightarrow \mathbb{R}^n$, dada por

$$T(x) = \frac{\sum_{i=1}^m \frac{w_i a_i}{\|x - a_i\|}}{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|}}, \quad (3.2.2)$$

la cual satisface que $x^* = T(x^*)$ por la igualdad (2.2.7). Esto dio lugar a un método de punto fijo para resolver el problema (3.2.1):

Algoritmo 1: Algoritmo de Weiszfeld

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$

Resultado: $x^* \in \mathbb{R}^n$ que minimiza $f(x)$

Elegir $x^{(0)} \neq a_1, \dots, a_m$;

para $j = 0, 1, 2, \dots$ **hacer**

$x^{(j+1)} \leftarrow T(x^{(j)})$;

si se satisface algún criterio de parada entonces

$x^* \leftarrow x^{(j+1)}$

El criterio de parada podría ser que la distancia entre iteraciones sea menor que un cierto $\varepsilon > 0$ o que la norma del gradiente en la iteración sea menor a cierto $\varepsilon > 0$, entre otros. En el Capítulo 6 presentamos los criterios que utilizamos en nuestro análisis.

A continuación daremos una demostración de convergencia basada en las de Kuhn [Kuh73] y de Beck y Sabach en [BS15] (con la notación que venimos usando). Para poder realizarla, precisaremos conocer varias propiedades del operador T .

Lo primero que resalta Kuhn es que el método de Weiszfeld es también un método de gradiente,

pues

$$\begin{aligned} T(x) &= \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} \sum_{i=1}^m \frac{w_i a_i}{\|x - a_i\|} \\ T(x) &= x + \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} \sum_{i=1}^m \frac{w_i (a_i - x)}{\|x - a_i\|} \\ T(x) &= x - \frac{1}{L(x)} \nabla f(x), \end{aligned} \quad (3.2.3)$$

donde la función $L : \mathbb{R}^n - \mathcal{A} \rightarrow \mathbb{R}^+$ está dada por

$$L(x) := \sum_{i=1}^m \frac{w_i}{\|x - a_i\|}. \quad (3.2.4)$$

Este resultado será utilizado en algunas demostraciones, así como también en algunos métodos de aceleración del algoritmo. Además utilizaremos la extensión del operador T a todo el espacio \mathbb{R}^n , definiéndolo como en (3.1.1), es decir, fijando

$$T(a_i) = a_i \quad \text{para } i = 1, \dots, m. \quad (3.2.5)$$

Lema 3.2.1. *La extensión de T dada en (3.2.5) hace que la función $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ sea continua.*

Demostración. Es claro que T es continua en $\mathbb{R}^n - \mathcal{A}$ pues para cada $i = 1, \dots, m$, las funciones $\|x - a_i\|$ son continuas y se anulan sólo en $x = a_i \in \mathcal{A}$.

Veamos que $\lim_{x \rightarrow a_j} T(x) = a_j$:

Como

$$\lim_{x \rightarrow a_j} \frac{w_i}{\|x - a_i\|} = \frac{w_i}{\|a_j - a_i\|} > 0 \quad \text{si } i \neq j$$

y

$$\lim_{x \rightarrow a_j} \frac{w_j}{\|x - a_j\|} = +\infty,$$

entonces

$$\lim_{x \rightarrow a_j} L(x) = \lim_{x \rightarrow a_j} \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} = +\infty.$$

Además, $\nabla f(x)$ es acotado pues

$$\|\nabla f(x)\| = \left\| \sum_{i=1}^m w_i \frac{x - a_i}{\|x - a_i\|} \right\| \leq \sum_{i=1}^m w_i \left\| \frac{x - a_i}{\|x - a_i\|} \right\| \leq \sum_{i=1}^m w_i.$$

$$\text{Luego } \lim_{x \rightarrow a_j} \frac{1}{L(x)} \nabla f(x) = (0, \dots, 0) \in \mathbb{R}^n \quad \text{y} \quad \lim_{x \rightarrow a_j} T(x) = \lim_{x \rightarrow a_j} \left(x - \frac{1}{L(x)} \nabla f(x) \right) = a_j. \quad \square$$

Lema 3.2.2. *Para todo $x \in \mathbb{R}^n$ se cumple que $T(x) \in \mathcal{C}(\mathcal{A})$, es decir, $T : \mathbb{R}^n \rightarrow \mathcal{C}(\mathcal{A})$.*

Demostración. Si $x \in \mathcal{A}$, entonces $T(x) = x$ y por lo tanto $T(x) \in \mathcal{C}(\mathcal{A})$.

Consideremos ahora $x \notin \mathcal{A}$. Tenemos que

$$T(x) = \sum_{i=1}^m \left(\frac{\frac{w_i}{\|x - a_i\|}}{\sum_{j=1}^m \frac{w_j}{\|x - a_j\|}} \right) a_i = \sum_{i=1}^m \alpha_i(x) a_i, \quad \text{con} \quad \alpha_i(x) = \frac{\frac{w_i}{\|x - a_i\|}}{\sum_{j=1}^m \frac{w_j}{\|x - a_j\|}}, \quad i = 1, \dots, m.$$

Los coeficientes $\alpha_i(x)$ satisfacen que $0 < \alpha_i(x) < 1$ para todo i y $\sum_{i=1}^m \alpha_i(x) = 1$. Luego, $T(x)$ está en la cápsula convexa de \mathcal{A} , $\mathcal{C}(\mathcal{A})$. \square

Estudiemos un poco el comportamiento del operador T con respecto a la función f . En primer lugar, como trabajaremos con un método de punto fijo, veamos la relación entre los puntos fijos de T y la función f .

Lema 3.2.3. $T(x) = x$ si y sólo si se da una de las siguientes situaciones:

1) $x \in \mathcal{A}$.

2) x es el minimizador de f .

Demostración. Si $x \in \mathcal{A}$, entonces $T(x) = x$ por definición.

Si $x \notin \mathcal{A}$, por el inciso a. del Teorema 2.2.7, x es minimizador de f si y sólo si $\nabla f(x) = 0$, si y sólo si

$$T(x) = x - \frac{1}{L(x)} \cdot 0 = x.$$

□

A continuación veremos cómo se comporta f en los demás puntos. El siguiente resultado es muy importante, pues prueba que el algoritmo genera una sucesión tal que el valor funcional en cada iteración vaya decreciendo.

Teorema 3.2.4. (Monotonicidad de f respecto al operador T)

Si $T(x) \neq x$, entonces $f(T(x)) < f(x)$.

Demostración. Notemos que como $T(x) \neq x$, entonces $x \notin \mathcal{A}$ y T está definida como en (3.2.2).

$T(x)$ es el único minimizador de la función $h_x : \mathbb{R}^n - \mathcal{A} \rightarrow \mathbb{R}$ dada por

$$h_x(y) = \sum_{i=1}^m w_i \frac{\|y - a_i\|^2}{\|x - a_i\|}$$

pues h_x es estrictamente convexa por ser suma de funciones cuadráticas definidas positivas, y su gradiente

$$\nabla h_x(y) = \sum_{i=1}^m w_i \frac{2(y - a_i)}{\|x - a_i\|}$$

se anula en $T(x)$.

Si evaluamos la función h_x en el punto x obtenemos

$$h_x(x) = \sum_{i=1}^m w_i \frac{\|x - a_i\|^2}{\|x - a_i\|} = \sum_{i=1}^m w_i \|x - a_i\| = f(x) \quad (3.2.6)$$

y evaluando en $T(x)$ obtenemos

$$\begin{aligned} h_x(T(x)) &= \sum_{i=1}^m w_i \frac{\|T(x) - a_i\|^2}{\|x - a_i\|} \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} ((\|T(x) - a_i\| - \|x - a_i\|) + \|x - a_i\|)^2 \\ &= \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (\|T(x) - a_i\| - \|x - a_i\|)^2 \right) + 2(f(T(x)) - f(x)) + f(x). \end{aligned} \quad (3.2.7)$$

Usando que $\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (\|T(x) - a_i\| - \|x - a_i\|)^2 \geq 0$ en (3.2.7) obtenemos que

$$h_x(T(x)) \geq 2f(T(x)) - f(x). \quad (3.2.8)$$

Como $T(x)$ es el minimizador de h_x , entonces $h_x(T(x)) < h(x)$. Reemplazando con las ecuaciones (3.2.6) y (3.2.8), obtenemos

$$2f(T(x)) - f(x) \leq h_x(T(x)) < h_x(x) = f(x)$$

y por lo tanto $f(T(x)) < f(x)$. □

Consideremos la sucesión $\{x^{(k)}\}$ dada por $x^{(0)} \in \mathbb{R}^n$ y $x^{(k)} = T(x^{(k-1)})$ para todo $k \in \mathbb{N}$.

Definición 1. Diremos que la sucesión $\{x^{(k)}\}$ “se detiene” si existe $K \in \mathbb{N}$ tal que $x^{(k)} = x^{(K)}$ para todo $k \geq K$. Diremos que “se detiene en \tilde{x} ” si $x^{(K)} = \tilde{x}$.

Afirmación 1. Aplicando el Lema 3.2.3 y el Teorema 3.2.4 a la sucesión $\{x^{(k)}\}$, podemos deducir que se da uno de los siguientes casos:

1. La sucesión $\{x^{(k)}\}$ se detiene en un nodo.
2. $x^{(k)} \notin \mathcal{A}$ para todo $k \in \mathbb{N}$ y la sucesión $\{x^{(k)}\}$ se detiene en la solución al problema (3.2.1).
3. $x^{(k)} \notin \mathcal{A}$ para todo $k \in \mathbb{N}$ y la sucesión de valores funcionales $\{f(x^{(k)})\}$ es monótona decreciente.

En el Teorema 3.2.6 veremos que si se no se da el primer caso, entonces la serie converge a la solución. Para poder probarlo, precisaremos un último resultado que asegura el decrecimiento de la sucesión $\{\|x^{(k)} - x\|\}$ para ciertos valores de $x \in \mathbb{R}^n$.

Lema 3.2.5. (Monotonidad de Fejér) Supongamos que $x^{(k)} \notin \mathcal{A}$ para todo k . Si $x \in \mathbb{R}^n$ satisface que $f(x) \leq f(x^{(k)})$ para todo k , entonces

$$\|x^{(k+1)} - x\| \leq \|x^{(k)} - x\| \text{ para todo } k \in \mathbb{N}.$$

Demostración. Consideremos la función $h_x : \mathbb{R}^n - \mathcal{A} \rightarrow \mathbb{R}$ definida como en la demostración del Teorema 3.2.4. h_x es cuadrática, y su desarrollo de Taylor centrado en x es

$$h_x(y) = h_x(x) + \langle \nabla h_x(x), y - x \rangle + \frac{1}{2}(y - x)^T \nabla^2 h_x(x) \cdot (y - x).$$

Haciendo las cuentas obtenemos que

$$\begin{aligned} h_x(x) &= f(x) \text{ como probamos antes} \\ \nabla h_x(y) &= 2 \sum_{i=1}^m \frac{w_i(y - a_i)}{\|x - a_i\|} \Rightarrow \nabla h_x(x) = 2\nabla f(x) \\ \nabla^2 h_x(y) &= 2 \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} Id = 2L(x)Id. \end{aligned}$$

Reemplazando en el desarrollo de Taylor obtenemos

$$h_x(y) = f(x) + 2\langle \nabla f(x), y - x \rangle + L(x) \|y - x\|^2. \quad (3.2.9)$$

Si tomamos $x = x^{(k)}$ e $y = T(x) = x^{(k+1)}$ obtenemos que

$$h_{x^{(k)}}(x^{(k+1)}) = f(x^{(k)}) + 2\langle \nabla f(x^{(k)}), x^{(k+1)} - x^{(k)} \rangle + L(x^{(k)}) \|x^{(k+1)} - x^{(k)}\|^2,$$

y utilizando la ecuación (3.2.8) obtenemos que

$$2f(x^{(k+1)}) - f(x^{(k)}) \leq f(x^{(k)}) + 2\langle \nabla f(x^{(k)}), x^{(k+1)} - x^{(k)} \rangle + L(x^{(k)}) \|x^{(k+1)} - x^{(k)}\|^2.$$

Por lo tanto,

$$f(x^{(k+1)}) \leq f(x^{(k)}) + \langle \nabla f(x^{(k)}), x^{(k+1)} - x^{(k)} \rangle + \frac{L(x^{(k)})}{2} \|x^{(k+1)} - x^{(k)}\|^2. \quad (3.2.10)$$

Como f es convexa, entonces vale que $f(x) \geq f(x^{(k)}) + \langle \nabla f(x^{(k)}), x - x^{(k)} \rangle$ para todo $x \in \mathbb{R}^n$, o equivalentemente,

$$f(x^{(k)}) \leq f(x) + \langle \nabla f(x^{(k)}), x^{(k)} - x \rangle$$

Combinando esto con (3.2.10), obtenemos que

$$\begin{aligned} f(x^{(k+1)}) &\leq f(x^{(k)}) + \langle \nabla f(x^{(k)}), x^{(k+1)} - x^{(k)} \rangle + \frac{L(x^{(k)})}{2} \|x^{(k+1)} - x^{(k)}\|^2 \\ f(x^{(k+1)}) &\leq f(x) + \langle \nabla f(x^{(k)}), x^{(k)} - x \rangle + \langle \nabla f(x^{(k)}), x^{(k+1)} - x^{(k)} \rangle + \frac{L(x^{(k)})}{2} \|x^{(k+1)} - x^{(k)}\|^2 \\ f(x^{(k+1)}) &\leq f(x) + \langle \nabla f(x^{(k)}), x^{(k+1)} - x \rangle + \frac{L(x^{(k)})}{2} \|x^{(k+1)} - x^{(k)}\|^2. \end{aligned}$$

Por la ecuación (3.2.3), $\nabla f(x^{(k)}) = L(x^{(k)})(x^{(k)} - x^{(k+1)})$ y por lo tanto

$$\begin{aligned} f(x^{(k+1)}) &\leq f(x) + L(x^{(k)}) \langle x^{(k)} - x^{(k+1)}, x^{(k+1)} - x \rangle + \frac{L(x^{(k)})}{2} \|x^{(k+1)} - x^{(k)}\|^2 \\ f(x^{(k+1)}) - f(x) &\leq \frac{L(x^{(k)})}{2} \left(-2 \langle x^{(k+1)} - x^{(k)}, x^{(k+1)} - x \rangle + \|x^{(k+1)} - x^{(k)}\|^2 \right) \\ f(x^{(k+1)}) - f(x) &\leq \frac{L(x^{(k)})}{2} \left(\|x^{(k)} - x\|^2 - \|x^{(k+1)} - x\|^2 \right). \end{aligned} \quad (3.2.11)$$

Si x satisface que $f(x) \leq f(x^{(k)})$ para todo $k \in \mathbb{N}$, en particular se satisface que

$$f(x^{(k+1)}) - f(x) \geq 0.$$

Aplicando esto a (3.2.11), y usando que $L(x^{(k)}) > 0$, tenemos que

$$0 \leq \|x^{(k)} - x\|^2 - \|x^{(k+1)} - x\|^2$$

lo cual concluye la demostración. \square

Ahora sí tenemos todas las herramientas para demostrar el teorema de convergencia del método de Weiszfeld.

Teorema 3.2.6. *Sea $\{x^{(k)}\}$ la sucesión generada por el método de Weiszfeld, con $x^{(0)} \in \mathbb{R}^n - \mathcal{A}$. Si $x^{(k)} \notin \mathcal{A}$ para todo $k \in \mathbb{N}$, entonces la sucesión converge a la solución x^* del problema de Fermat-Weber.*

Demostración. Como $x^{(k)} \notin \mathcal{A}$, pueden ocurrir el segundo o el tercer caso de la Afirmación 1.

En el caso 2, la sucesión se detiene en la solución, es decir, existe $K \in \mathbb{N}$ tal que $x^{(k)} = x^*$ para todo $k \geq K$. Luego, $\{x^{(k)}\}$ converge a x^* .

Veamos el caso 3. Por el Lema 3.2.2, $x^{(k)} \in \mathcal{C}(\mathcal{A})$ para todo $k > 0$, por lo que $\{x^{(k)}\}$ es acotada. Por el teorema de Bolzano-Weierstrass, existe al menos una subsucesión convergente $\{x^{(k_j)}\}$.

Supongamos que $\{x^{(k_j)}\}$ es una subsucesión que converge a \tilde{x} . Como $\{f(x^{(k)})\}$ es monótona decreciente y acotada inferiormente por $f(x^*)$, entonces converge a algún $f^* \in \mathbb{R}$ que satisface $f^* \leq f(x^{(k)})$ para todo k . Por continuidad de f ,

$$f(\tilde{x}) = \lim_{j \rightarrow +\infty} f(x^{(k_j)}) = \lim_{j \rightarrow +\infty} f(x^{(k)}) = f^*$$

y por lo tanto $f(\tilde{x}) \leq f(x^{(k)})$ para todo k . Por el Lema 3.2.5, la sucesión $\{\|x^{(k)} - \tilde{x}\|\}$ es no creciente, y es acotada inferiormente por 0, por lo cual converge a una constante $\gamma \in \mathbb{R}$. Como $\{\|x^{(k_j)} - \tilde{x}\|\}$ es una subsucesión de $\{\|x^{(k)} - \tilde{x}\|\}$ obtenemos

$$\gamma = \lim_{k \rightarrow +\infty} \|x^{(k)} - \tilde{x}\| = \lim_{j \rightarrow +\infty} \|x^{(k_j)} - \tilde{x}\| = 0.$$

Luego $\{x^{(k)}\}$ converge a \tilde{x} . Solo resta ver que \tilde{x} es la solución del problema de Fermat-Weber.

Como $x^{(k+1)} = T(x^{(k)})$ y T es continua, entonces

$$\tilde{x} = \lim_{k \rightarrow +\infty} x^{(k+1)} = \lim_{k \rightarrow +\infty} T(x^{(k)}) = T\left(\lim_{k \rightarrow +\infty} x^{(k)}\right) = T(\tilde{x})$$

es decir, $\tilde{x} = T(\tilde{x})$.

Si $\tilde{x} \notin \mathcal{A}$, entonces $\tilde{x} = x^*$ por el Lema 3.2.3.

Si $\tilde{x} \in \mathcal{A}$, entonces $\tilde{x} = a_l$ para algún $l = 1, \dots, m$. Como para cada k $x^{(k+1)} = T(x^{(k)})$ es el minimizador de $h_{x^{(k)}}$, entonces $\nabla h_{x^{(k)}}(x^{(k+1)}) = 0$, lo cual implica que

$$\sum_{i=1}^m \frac{w_i(x^{(k+1)} - a_i)}{\|x^{(k)} - a_i\|} = 0.$$

Manipulando esta igualdad obtenemos

$$\begin{aligned} \sum_{\substack{i=1 \\ i \neq l}}^m \frac{w_i(x^{(k+1)} - a_i)}{\|x^{(k)} - a_i\|} &= -\frac{w_l(x^{(k+1)} - a_l)}{\|x^{(k)} - a_l\|} \text{ y por lo tanto} \\ \left\| \sum_{\substack{i=1 \\ i \neq l}}^m \frac{w_i(x^{(k+1)} - a_i)}{\|x^{(k)} - a_i\|} \right\| &= w_l \left\| \frac{x^{(k+1)} - a_l}{\|x^{(k)} - a_l\|} \right\|. \end{aligned} \quad (3.2.12)$$

Como $\{\|x^{(k)} - \tilde{x}\|\}$ es no creciente, entonces $\frac{\|x^{(k+1)} - a_l\|}{\|x^{(k)} - a_l\|} \leq 1$. Usando este resultado, la ecuación (3.2.12) implica que

$$\|R_l\| = \left\| \sum_{\substack{i=1 \\ i \neq l}}^m \frac{w_i(x^{(k+1)} - a_i)}{\|x^{(k)} - a_i\|} \right\| = w_l \left\| \frac{x^{(k+1)} - a_l}{\|x^{(k)} - a_l\|} \right\| \leq w_l$$

Por el Teorema 2.2.7, esto significa que $a_l = \tilde{x}$ es el minimizador de f , $\tilde{x} = x^*$. \square

Con esto hemos probado que la sucesión generada por el algoritmo de Weiszfeld, si no se detiene en un nodo, converge a la solución del problema de Fermat-Weber generalizado. Además de converger, es importante que el algoritmo tenga una tasa de convergencia que haga factible su aplicación para problemas reales. Katz [Kat74] probó que la tasa de convergencia local del algoritmo es lineal si la solución no es un nodo, y puede variar entre sublineal, lineal o cuadrática para el caso en que la solución es un nodo. La demostración de este resultado la veremos a continuación.

Teorema 3.2.7. *Supongamos que la solución x^* del problema de Fermat-Weber no es un nodo. Entonces existen λ_{\min} y λ_{\max} con $0 < \lambda_{\min} < \lambda_{\max} < 1$ tales que, si $x^{(0)}$ está lo suficientemente cerca de x^* ,*

$$\lambda_{\min} \|x^{(k)} - x^*\| + \mathcal{O}(\|x^{(k)} - x^*\|^2) \leq \|x^{(k+1)} - x^*\| \leq \lambda_{\max} \|x^{(k)} - x^*\| + \mathcal{O}(\|x^{(k)} - x^*\|^2) \quad (3.2.13)$$

para todo $k \in \mathbb{N}$.

Demostración. Estudiemos la transformación $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ alrededor de x^* . Como no es un nodo, existe un entorno de x^* tal que todas las funciones $f_i(x) = \|x - a_i\|$ son continuamente diferenciables y la transformación T está dada por

$$T(x) = \frac{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} a_i}{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|}}$$

y resulta continuamente diferenciable. Luego, podemos escribir su desarrollo de Taylor de orden dos alrededor de x^* ,

$$T(x) = T(x^*) + J(x^*)(x - x^*) + \mathcal{O}(\|x - x^*\|^2) \text{ para } x \text{ cercano a } x^*,$$

donde $J(x^*)$ es la jacobiana de T evaluada en x^* . Como x^* es solución, entonces $T(x^*) = x^*$. Luego, para todo x lo suficientemente cerca de x^* se tiene que

$$T(x) - x^* = J(x^*)(x - x^*) + \mathcal{O}(\|x - x^*\|^2). \quad (3.2.14)$$

Como la sucesión generada por el algoritmo converge a la solución, si $x^{(0)}$ está cerca de x^* , entonces todo $x^{(k)}$ lo está. Luego, podemos tomar $x = x^{(k)}$ en (3.2.14), y como $T(x^{(k)}) = x^{(k+1)}$ obtenemos que

$$x^{(k+1)} - x^* = J(x^*)(x^{(k)} - x^*) + \mathcal{O}(\|x^{(k)} - x^*\|^2) \text{ para todo } k. \quad (3.2.15)$$

Estudiemos la matriz jacobiana de $T = (T_1, \dots, T_n)$.

$$\begin{aligned} \frac{\partial T_s}{\partial x_j}(x) &= \frac{\partial}{\partial x_j} \Big|_x \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} a_{is} \right) \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} \\ &= - \left(\sum_{i=1}^m \frac{w_i a_{is} (x_j - a_{ij})}{\|x - a_i\|^3} \right) \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} + \\ &\quad \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} a_{is} \right) \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-2} \left(\sum_{i=1}^m \frac{w_i (x_j - a_{ij})}{\|x - a_i\|^3} \right) \\ &= \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} \left[\sum_{i=1}^m \frac{w_i (x_j - a_{ij})}{\|x - a_i\|^3} (-a_{is}) + T_s(x) \left(\sum_{i=1}^m \frac{w_i (x_j - a_{ij})}{\|x - a_i\|^3} \right) \right] \\ &= \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} \left(\sum_{i=1}^m \frac{w_i (x_j - a_{ij})}{\|x - a_i\|^3} (T_s(x) - a_{is}) \right) \\ &= \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|^3} [(x - a_i)(T(x) - a_i)^T]_{js} \right), \end{aligned}$$

donde dada una matriz B , la notación $[B]_{js}$ indica el elemento de su fila j y columna s .

Por lo tanto

$$J(x) = \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right)^{-1} \left(\sum_{i=1}^m \frac{w_i}{\|x - a_i\|^3} (x - a_i)(T(x) - a_i)^T \right). \quad (3.2.16)$$

Evaluando en x^* , y usando que $T(x^*) = x^*$, obtenemos

$$J(x^*) = \left(\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} \right)^{-1} \left(\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|^3} (x^* - a_i)(x^* - a_i)^T \right). \quad (3.2.17)$$

Como $J(x^*)$ es una matriz simétrica con coeficientes reales, entonces es diagonalizable, es decir que existe una base ortonormal de autovectores de $J(x^*)$, $\{v_1, \dots, v_n\}$, donde cada v_i tiene autovalor λ_i .

Puede probarse que si $v \in \mathbb{R}^n$, la matriz $B = vv^T$ es semidefinida positiva, es decir que $y^T B y \geq 0$ para todo $y \in \mathbb{R}^n$, y la igualdad vale si y sólo si y es ortogonal a v . En particular, $(x^* - a_i)(x^* - a_i)^T$ es semidefinida positiva para todo $i = 1, \dots, m$, y como

$$\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} > 0 \quad \text{y} \quad \sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|^3} > 0$$

obtenemos que $y^T J(x^*) y \geq 0$ para todo $y \in \mathbb{R}^n$, y la igualdad vale si y sólo si $y^T (x^* - a_i) = 0$ para todo i . Esto significa que $J(x^*)$ es semidefinida positiva, y por lo tanto todos sus autovalores son no negativos. Además,

$$\lambda_1 + \dots + \lambda_n = \text{tr}[J(x^*)] = \sum_{j=1}^n \left(\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} \right)^{-1} \left(\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|^3} (x_j^* - a_{ij})^2 \right) = 1 \quad (3.2.18)$$

por lo que $0 \leq \lambda_i \leq 1$ para todo i .

Como $\{v_1, \dots, v_n\}$ es base ortonormal de \mathbb{R}^n , para cada x existen $\alpha_j(x)$, $j = 1, \dots, n$, tales que

$$x - x^* = \sum_{j=1}^n \alpha_j(x) v_j \quad \text{y} \quad \|x - x^*\|^2 = \left\| \sum_{j=1}^n \alpha_j(x) v_j \right\|^2 = \sum_{j=1}^n |\alpha_j(x)|^2.$$

Además,

$$J(x^*)(x - x^*) = J(x^*) \left(\sum_{j=1}^n \alpha_j(x) v_j \right) = \sum_{j=1}^n \alpha_j(x) \lambda_j v_j$$

por lo que

$$\|J(x^*)(x - x^*)\|^2 = \sum_{j=1}^n |\alpha_j(x)|^2 \lambda_j^2. \quad (3.2.19)$$

Si y es ortogonal a $x^* - a_i$ para todo $i = 1, \dots, m$, entonces y es ortogonal al subespacio \mathcal{S} generado por los vectores $x^* - a_i$, $i = 1, \dots, n$. Luego, $y^T J(x^*) y > 0$ si y sólo si $y \in \mathcal{S}$. Como los nodos son no colineales, \mathcal{S} tiene dimensión $s \geq 2$ y por lo tanto existen $s \geq 2$ autovalores no nulos de $J(x^*)$, no necesariamente distintos. Más aún, \mathcal{S} es la unión de los autoespacios asociados a autovalores positivos.

Sean $\lambda_{j_1}, \dots, \lambda_{j_s}$ los autovalores no nulos y v_{j_1}, \dots, v_{j_s} sus autovectores asociados, y definamos $\lambda_{\min} = \min\{\lambda_{j_i} : 1 \leq i \leq s\}$ y $\lambda_{\max} = \max\{\lambda_j : 1 \leq j \leq n\}$. En particular, por (3.2.18), tenemos que $0 < \lambda_{\min} \leq \frac{1}{s} \leq \lambda_{\max} < 1$.

Para $x \in \mathcal{C}(\mathcal{A})$, por definición de $\mathcal{C}(\mathcal{A})$ y por $x^* \in \mathcal{C}(\mathcal{A})$, se cumple que $x - x^* \in \mathcal{S}$, y como \mathcal{S} está generado por v_{j_1}, \dots, v_{j_s} , podemos escribir

$$x - x^* = \sum_{i=1}^s \alpha_{j_i}(x) v_{j_i}$$

es decir que $\alpha_j(x) = 0$ si $j \neq j_1, \dots, j_s$. Usando (3.2.19) obtenemos que

$$\|J(x^*)(x - x^*)\|^2 = \sum_{i=1}^s \lambda_{j_i}^2 |\alpha_{j_i}(x)|^2.$$

En particular, vale tomando $x = x^{(k)}$, pues $x^{(k)} \in \mathcal{C}(\mathcal{A})$ para todo $k \in \mathbb{N}$. Utilizando que $\lambda_{\min} \leq \lambda_{j_i} \leq \lambda_{\max}$, obtenemos las siguientes desigualdades:

$$\begin{aligned} \left\| J(x^*)(x^{(k)} - x^*) \right\|^2 &= \sum_{i=1}^s \lambda_{j_i}^2 |\alpha_{j_i}|^2 \geq \sum_{i=1}^s \lambda_{\min}^2 |\alpha_{j_i}|^2 = \lambda_{\min}^2 \left\| x^{(k)} - x^* \right\|^2 \quad y \\ \left\| J(x^*)(x^{(k)} - x^*) \right\|^2 &= \sum_{i=1}^s \lambda_{j_i}^2 |\alpha_{j_i}|^2 \leq \sum_{i=1}^s \lambda_{\max}^2 |\alpha_{j_i}|^2 = \lambda_{\max}^2 \left\| x^{(k)} - x^* \right\|^2, \end{aligned}$$

por lo que

$$\lambda_{\min} \left\| x^{(k)} - x^* \right\| \leq \left\| J(x^*)(x^{(k)} - x^*) \right\| \leq \lambda_{\max} \left\| x^{(k)} - x^* \right\|. \quad (3.2.20)$$

Como para $x^{(0)}$ lo suficientemente cerca de x^* vale la igualdad (3.2.15), si tomamos norma a ambos lados de ella tenemos que

$$\left\| x^{(k+1)} - x^* \right\| = \left\| J(x^*)(x^{(k)} - x^*) \right\| + \mathcal{O}\left(\left\| x^{(k)} - x^* \right\|^2\right)$$

e introduciendo (3.2.20) obtenemos

$$\lambda_{\min} \left\| x^{(k)} - x^* \right\| + \mathcal{O}\left(\left\| x^{(k)} - x^* \right\|^2\right) \leq \left\| x^{(k+1)} - x^* \right\| \leq \lambda_{\max} \left\| x^{(k)} - x^* \right\| + \mathcal{O}\left(\left\| x^{(k)} - x^* \right\|^2\right).$$

□

Esto nos indica que la peor tasa de convergencia posible para el caso donde $x^* \notin \mathcal{A}$ es lineal, aunque también la mejor tasa posible lo es.

En el caso donde x^* es un nodo, veremos que

$$\lim_{x \rightarrow a_i} \frac{\|T(x) - a_i\|}{\|x - a_i\|} = \frac{\|R_i\|}{\|w_i\|} \quad \text{para todo } i = 1, \dots, m$$

y determinaremos la tasa de convergencia usando que si $x^* = a_s$, entonces $\|R_s\| \leq w_s$.

Teorema 3.2.8. *Supongamos que $x^* = a_s$, con $s \in \{1, \dots, m\}$. Para $x^{(0)}$ lo suficientemente cerca de x^* , se cumple que:*

- Si $0 < \|R_s\| < w_s$, la convergencia es lineal con factor de convergencia asintótica $\frac{\|R_s\|}{w_s}$.
- Si $\|R_s\| = 0$, la convergencia es cuadrática.
- Si $\|R_s\| = w_s$, la convergencia es sublineal.

Demostración. Analicemos la transformación T :

$$T(x) = \frac{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} a_i}{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|}} = \frac{\frac{w_s}{\|x - a_s\|} a_s + \sum_{i \neq s} \frac{w_i}{\|x - a_i\|} a_i}{\frac{w_s}{\|x - a_s\|} + \sum_{i \neq s} \frac{w_i}{\|x - a_i\|}} = \frac{a_s + \frac{\|x - a_s\|}{w_s} \sum_{i \neq s} \frac{w_i}{\|x - a_i\|} a_i}{1 + \frac{\|x - a_s\|}{w_s} \sum_{i \neq s} \frac{w_i}{\|x - a_i\|}},$$

por lo que

$$T(x) = \left(a_s + \frac{\|x - a_s\|}{w_s} G(x) \right) \left(1 + \frac{\|x - a_s\|}{w_s} g(x) \right)^{-1},$$

donde

$$G(x) = \sum_{i \neq s} \frac{w_i}{\|x - a_i\|} a_i \quad y \quad g(x) = \sum_{i \neq s} \frac{w_i}{\|x - a_i\|}. \quad (3.2.21)$$

Para x suficientemente cerca de a_s , podemos expresar

$$\left(1 + \frac{\|x - a_s\|}{w_s} g(x)\right)^{-1} = 1 - \frac{\|x - a_s\|}{w_s} g(x) + \left(\frac{\|x - a_s\|}{w_s}\right)^2 g^2(x) + \mathcal{O}(\|x - a_s\|^3). \quad (3.2.22)$$

Como G y g son diferenciables en un entorno de a_s , podemos calcular su desarrollo de Taylor de orden dos para x cerca de a_s , obteniendo

$$G(x) = G(a_s) - \sum_{i \neq s} c_i(x) a_i + \mathcal{O}(\|x - a_s\|^2) \quad y \quad (3.2.23)$$

$$g(x) = g(a_s) - \sum_{i \neq s} c_i(x) + \mathcal{O}(\|x - a_s\|^2), \quad (3.2.24)$$

con

$$c_i(x) = \frac{w_i}{\|a_s - a_i\|^3} (a_s - a_i)^T (x - a_s).$$

Además,

$$G(a_s) - a_s g(a_s) = \sum_{i \neq s} \frac{w_i}{\|a_s - a_i\|} a_i - \sum_{i \neq s} \frac{w_i}{\|a_s - a_i\|} a_s = \sum_{i \neq s} \frac{w_i}{\|a_s - a_i\|} (a_i - a_s) = -R_s. \quad (3.2.25)$$

Entonces

$$\begin{aligned} T(x) &= \left(a_s + \frac{\|x - a_s\|}{w_s} G(a_s) - \frac{\|x - a_s\|}{w_s} \sum_{i \neq s} c_i(x) a_i + \mathcal{O}(\|x - a_s\|^3) \right) \\ &\quad \left(1 - \frac{\|x - a_s\|}{w_s} g(a_s) + \frac{\|x - a_s\|}{w_s} \sum_{i \neq s} c_i(x) + \frac{\|x - a_s\|^2}{w_s^2} g^2(a_s) + \mathcal{O}(\|x - a_s\|^3) \right) \\ &= a_s + \frac{\|x - a_s\|}{w_s} \left(G(a_s) - a_s g(a_s) - \sum_{i \neq s} c_i(x) a_i + \sum_{i \neq s} c_i(x) a_s \right) \\ &\quad - \frac{\|x - a_s\|^2}{w_s^2} (G(a_s) g(a_s) - a_s g^2(a_s)) + \mathcal{O}(\|x - a_s\|^3) \\ &= a_s - \frac{\|x - a_s\|}{w_s} R_s + \frac{\|x - a_s\|}{w_s} \sum_{i \neq s} c_i(x) (a_s - a_i) + \frac{\|x - a_s\|^2}{w_s^2} g(a_s) R_s + \mathcal{O}(\|x - a_s\|^3). \end{aligned} \quad (3.2.26)$$

Sustrayendo a_s en ambos miembros obtenemos

$$T(x) - a_s = -\frac{\|x - a_s\|}{w_s} R_s + \frac{\|x - a_s\|}{w_s} \sum_{i \neq s} \frac{w_i}{\|a_s - a_i\|^3} (a_s - a_i) (a_s - a_i)^T (x - a_s) + \mathcal{O}(\|x - a_s\|^2). \quad (3.2.27)$$

Tomando norma a ambos lados y dividiendo por $\|x - a_s\|$ obtenemos que

$$\frac{\|T(x) - a_s\|}{\|x - a_s\|} = \frac{\|R_s\|}{w_s} + \mathcal{O}(\|x - a_s\|)$$

y por lo tanto

$$\lim_{x \rightarrow a_s} \frac{\|T(x) - a_s\|}{\|x - a_s\|} = \frac{\|R_s\|}{w_s}. \quad (3.2.28)$$

Supongamos ahora que $x^* = a_s$. Como $x^{(k+1)} = T(x^{(k)})$ para todo $k \in \mathbb{N}$ y la sucesión $\{x^{(k)}\}$ converge a la solución, la ecuación (3.2.28) implica que

$$\lim_{k \rightarrow +\infty} \frac{\|x^{(k+1)} - a_s\|}{\|x^{(k)} - a_s\|} = \frac{\|R_s\|}{w_s}. \quad (3.2.29)$$

Sabemos que a_s es solución si y sólo si $\|R_s\| \leq w_s$.

Si $0 < \|R_s\| < w_s$, entonces la velocidad de convergencia resulta lineal con factor de convergencia asintótica $\frac{\|R_s\|}{w_s}$.

Si $\|R_s\| = 0$, estudiando en mayor profundidad la ecuación (3.2.27), podemos probar que la velocidad de convergencia es cuadrática, pues

$$x^{(k+1)} - a_s = \frac{\|x^{(k)} - a_s\|}{w_s} \sum_{i \neq s} \frac{w_i}{\|a_s - a_i\|^3} (a_s - a_i)(a_s - a_i)^T (x^{(k)} - a_s) + \mathcal{O}(\|x^{(k)} - a_s\|^3)$$

por lo que

$$\|x^{(k+1)} - a_s\| = \mathcal{O}(\|x^{(k)} - a_s\|^2). \quad (3.2.30)$$

Por último, si $\|R_s\| = w_s$, entonces

$$\lim_{k \rightarrow +\infty} \frac{\|x^{(k+1)} - a_s\|}{\|x^{(k)} - a_s\|} = 1$$

y por lo tanto la convergencia es sublineal. \square

3.3. Variaciones para asegurar convergencia

Vimos en la Sección 3.1 que no cualquier punto inicial $x^{(0)}$ para el algoritmo de Weiszfeld genera una sucesión que converja a la solución, sino que hay puntos iniciales “malos” que generan una sucesión que se detiene en un nodo. Más aún, vimos que si la cápsula convexa de $\mathcal{A} = \{a_1, \dots, a_m\}$ tiene dimensión completa, podría haber numerables puntos iniciales malos, y si no es completa podría incluso ser un conjunto continuo de puntos iniciales malos. En vista de esto, si bien es poco probable elegir dichos puntos, es recomendable encontrar métodos que permitan evitar estas situaciones.

Como mencionamos anteriormente, Ostrech [OJ78] y Vardi y Zhang [VZ01] presentan variaciones al algoritmo de Weiszfeld tales que el algoritmo se detiene sólo si alcanzó la solución x^* del problema de Fermat-Weber, redefiniendo la transformación T en los nodos.

Ostrech utiliza el hecho de que el método de Weiszfeld es un método de gradiente

$$T(x) = x - \frac{1}{L(x)} \nabla f(x)$$

y crea una nueva transformación que permite variar la longitud del paso. Como las funciones L y ∇f no están definidas en los nodos, construye una nueva función que extiende al gradiente, $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$, y extiende L a todo \mathbb{R}^n . Estas extensiones están dadas por:

$$g(x) = \begin{cases} \nabla f(x) & \text{si } x \notin \mathcal{A} \\ 0 & \text{si } x = a_k \text{ y } \|R_k\| \leq w_k \\ R_k \left(1 - \frac{w_k}{\|R_k\|}\right) & \text{si } x = a_k \text{ y } \|R_k\| > w_k \end{cases} \quad (3.3.1)$$

y

$$L(x) = \begin{cases} \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} & \text{si } x \notin \mathcal{A} \\ \sum_{i \neq k} \frac{w_i}{\|x - a_i\|} & \text{si } x = a_k, k = 1, \dots, m. \end{cases} \quad (3.3.2)$$

Con esta definición, por el Teorema 2.2.7, $x^* \in \mathbb{R}^n$ minimiza f si y sólo si $g(x) = 0$. Usando ambas extensiones, Ostrech construye una nueva transformación definida por

$$T_c(x) = x - c \frac{1}{L(x)} g(x) \text{ para todo } x \in \mathbb{R}^n, \quad (3.3.3)$$

con $c \in \mathbb{R}$. A partir de ella genera un método iterativo dado por $x^{(k+1)} = T_{c_k}(x^{(k)})$ y demuestra que éste converge a la solución x^* si $x^{(0)} \in \mathcal{C}(\mathcal{A})$ y $1 \leq c_k \leq 2$ para todo $k \geq 0$.

Notemos que si $c_k = 1$ para todo k y nunca se llega a un nodo, el algoritmo de Ostrech coincide con el de Weiszfeld. Este caso particular es justamente el algoritmo de Vardi y Zhang [VZ01], el cual puede escribirse del siguiente modo:

$$x^{(s+1)} := T_{VZ}(x^{(k)}) = \begin{cases} T(x^{(k)}) & \text{si } x^{(k)} \notin \mathcal{A} \\ a_k & \text{si } x = a_k \text{ y } \|R_k\| \leq w_k \\ a_k - \left(\frac{\|R_k\| - w_k}{L(a_k)} \right) \frac{R_k}{\|R_k\|} & \text{si } x = a_k \text{ y } \|R_k\| > w_k \end{cases} \quad (3.3.4)$$

donde T es la transformación de Weiszfeld.

A continuación probaremos que en efecto, el método de Vardi y Zhang converge a la solución del problema de Fermat-Weber para cualquier punto inicial. Para ello, veremos primero un teorema de [BS15] que demuestra que si a_k no es solución del problema de Fermat, entonces $f(S_k) < f(a_k)$, donde

$$S_k = T_{VZ}(a_k) = a_k - \left(\frac{\|R_k\| - w_k}{L(a_k)} \right) \frac{R_k}{\|R_k\|}.$$

Lema 3.3.1. *Si a_k no es el minimizador de f , entonces*

$$f(a_k) - f(S_k) \geq \frac{(\|R_k\| - w_k)^2}{2L(a_k)} > 0.$$

Demostración.

$$S_k = a_k + t_k d_k \quad \text{con} \quad t_k = \frac{\|R_k\| - w_k}{L(a_k)} \quad \text{y} \quad d_k = -\frac{R_k}{\|R_k\|}.$$

Por un lado, notemos que

$$\begin{aligned} -2 \frac{1}{t_k} \|S_k - a_k\|^2 &= -2 \langle S_k - a_k, \frac{S_k - a_k}{t_k} \rangle \\ -\frac{1}{t_k} \|S_k - a_k\|^2 &= \frac{1}{t_k} \|S_k - a_k\|^2 - 2 \langle S_k - a_k, d_k \rangle \\ -\frac{1}{t_k} \|S_k - a_k\|^2 &= \frac{1}{t_k} \|S_k - a_k\|^2 - 2 \langle S_k - a_k, -\frac{R_k}{\|R_k\|} \rangle \\ -\frac{\|S_k - a_k\|^2}{t_k} &= \frac{1}{t_k} \|S_k - a_k\|^2 + \frac{2}{\|R_k\|} \langle S_k - a_k, R_k \rangle. \end{aligned} \quad (3.3.5)$$

Por otro lado,

$$\begin{aligned} \frac{1}{t_k} \|S_k - a_k\|^2 &= \frac{L(a_k)}{\|R_k\| - w_k} \|S_k - a_k\|^2 \\ &= \frac{1}{\|R_k\| - w_k} \sum_{\substack{i=1 \\ i \neq k}}^m \frac{w_i}{\|a_k - a_i\|} \|S_k - a_k\|^2 \\ &= \frac{1}{\|R_k\| - w_k} \sum_{\substack{i=1 \\ i \neq k}}^m \frac{w_i}{\|a_k - a_i\|} \left(\|S_k - a_i\|^2 + 2 \langle S_k - a_i, a_i - a_k \rangle + \|a_i - a_k\|^2 \right). \end{aligned}$$

Como $\langle S_k - a_i, a_i - a_k \rangle = \langle S_k - a_k, a_i - a_k \rangle + \langle a_k - a_i, a_i - a_k \rangle = \langle S_k - a_k, a_i - a_k \rangle - \|a_i - a_k\|^2$

entonces obtenemos que

$$\begin{aligned} \frac{1}{t_k} \|S_k - a_k\|^2 &= \frac{1}{\|R_k\| - w_k} \sum_{\substack{i=1 \\ i \neq k}}^m w_i \left(\frac{\|S_k - a_i\|^2}{\|a_k - a_i\|} + 2 \frac{\langle S_k - a_k, a_i - a_k \rangle}{\|a_k - a_i\|} - \|a_k - a_i\| \right) \\ \frac{1}{t_k} \|S_k - a_k\|^2 &= \frac{1}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \frac{\|S_k - a_i\|^2}{\|a_k - a_i\|} + 2 \langle S_k - a_k, \sum_{\substack{i=1 \\ i \neq k}}^m w_i \frac{a_i - a_k}{\|a_k - a_i\|} \rangle - f(a_k) \right) \\ \frac{1}{t_k} \|S_k - a_k\|^2 &= \frac{1}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \frac{\|S_k - a_i\|^2}{\|a_k - a_i\|} - 2 \langle S_k - a_k, R_k \rangle - f(a_k) \right). \end{aligned}$$

Como $0 \leq (\|S_k - a_i\| - \|a_k - a_i\|)^2 = \|S_k - a_i\|^2 - 2\|S_k - a_i\|\|a_k - a_i\| + \|a_k - a_i\|^2$, entonces

$$\frac{\|S_k - a_i\|^2}{\|a_k - a_i\|} \geq 2\|S_k - a_i\| - \|a_k - a_i\|.$$

Luego

$$\begin{aligned} \frac{1}{t_k} \|S_k - a_k\|^2 &\geq \frac{1}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i (2\|S_k - a_i\| - \|a_k - a_i\|) - 2 \langle S_k - a_k, R_k \rangle - f(a_k) \right) \\ \frac{1}{t_k} \|S_k - a_k\|^2 &\geq \frac{2}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \|S_k - a_i\| - \langle S_k - a_k, R_k \rangle - f(a_k) \right). \end{aligned} \quad (3.3.6)$$

Aplicando la desigualdad (3.3.6) en la ecuación (3.3.5), obtenemos que

$$\begin{aligned} -\frac{\|S_k - a_k\|^2}{t_k} &\geq \frac{2}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \|S_k - a_i\| - \langle S_k - a_k, R_k \rangle - f(a_k) \right) + \frac{2}{\|R_k\|} \langle S_k - a_k, R_k \rangle \\ -\frac{\|S_k - a_k\|^2}{t_k} &\geq \frac{2}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \|S_k - a_i\| - f(a_k) \right) - \left(\frac{2}{\|R_k\| - w_k} - \frac{2}{\|R_k\|} \right) \langle S_k - a_k, R_k \rangle \\ -\frac{\|S_k - a_k\|^2}{t_k} &\geq \frac{2}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \|S_k - a_i\| - f(a_k) \right) - \frac{2w_k}{\|R_k\| (\|R_k\| - w_k)} \langle S_k - a_k, R_k \rangle \\ -\frac{\|S_k - a_k\|^2}{t_k} &\geq \frac{2}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \|S_k - a_i\| - f(a_k) \right) + \frac{2w_k}{\|R_k\| - w_k} \langle S_k - a_k, d_k \rangle \\ -\frac{1}{t_k} \|S_k - a_k\|^2 &\geq \frac{2}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \|S_k - a_i\| - f(a_k) + w_k \langle S_k - a_k, d_k \rangle \right). \end{aligned}$$

Como $S_k - a_k = t_k d_k$ y $\|d_k\| = 1$, entonces $\|S_k - a_k\| = t_k$ y

$$\langle S_k - a_k, d_k \rangle = \langle S_k - a_k, \frac{S_k - a_k}{\|S_k - a_k\|} \rangle = \|S_k - a_k\|.$$

Por lo tanto,

$$\begin{aligned} -\frac{\|S_k - a_k\|^2}{t_k} &\geq \frac{2}{\|R_k\| - w_k} \left(\sum_{\substack{i=1 \\ i \neq k}}^m w_i \|S_k - a_i\| - f(a_k) + w_k \|S_k - a_k\| \right) \\ -\frac{t_k^2}{t_k} &\geq \frac{2}{\|R_k\| - w_k} (f(S_k) - f(a_k)). \end{aligned}$$

Como a_k no es minimizador, por el Teorema 2.2.7 $\|R_k\| - w_k > 0$. Entonces

$$\begin{aligned} -t_k \frac{\|R_k\| - w_k}{2} &\geq f(S_k) - f(a_k) \\ \left(\frac{\|R_k\| - w_k}{L(a_k)} \right) \frac{\|R_k\| - w_k}{2} &\leq -(f(S_k) - f(a_k)) \\ \frac{(\|R_k\| - w_k)^2}{2L(a_k)} &\leq f(a_k) - f(S_k). \end{aligned}$$

□

Este lema demuestra que si $x^{(k)}$ está generada por el método de Vardi y Zhang, la sucesión $\{f(x^{(k)})\}$ es no creciente. Más aún, $f(x^{(k+1)}) = f(x^{(k)})$ si y sólo si $x^{(k)}$ es el minimizador de f , y en ese caso la sucesión se detiene. Gracias a esto, podemos demostrar el siguiente teorema.

Teorema 3.3.2. *Para cualquier $x^{(0)} \in \mathbb{R}^n$, la sucesión $\{x^{(k)}\}$ generada por (3.3.4) converge al minimizador x^* de f .*

Demostración. Por el Teorema 3.2.6, sabemos que si la sucesión $\{x^{(k)}\}$ satisface alguna de las siguientes situaciones, entonces converge a la solución del problema de Fermat-Weber:

1. $x^{(k)} \notin \mathcal{A}$ para todo k .

En este caso, el método de Vardi y Zhang coincide con el método de Weiszfeld, y por lo tanto converge a x^* .

2. Existe algún $K \in \mathbb{N}$ tal que $x^{(K)} = a_l$ y a_l es el minimizador de f .

En este caso, $T(a_l) = a_l$ y por lo tanto $x^{(k)} = a_l$ para todo $k \geq K$, por lo que el método converge a la solución $x^* = a_l$.

Supongamos que algún $x^{(K)} = a_l$ y que a_l no es solución. Por el Lema 3.3.1, $x^{(K+1)} \neq a_l$ y $f(x^{(k)}) < f(a_l)$ para todo $k > K$. Luego $x^{(k)} \neq a_l$ para todo $k > K$.

Cada vez que una iteración coincida con un nodo que no sea minimizador, las siguientes iteraciones no volverán a ese nodo, y como \mathcal{A} es finito, para algún $M \in \mathbb{N}$ se cumplirá que $x^{(k)} \notin (\mathcal{A} - \{x^*\})$ para todo $k \geq M$.

La sucesión $\{y^{(s)}\}$ dada por el método de Vardi y Zhang con punto inicial $y^{(0)} = x^{(M)}$ satisface que $y^{(s)} = x^{(s+M)}$ para todo $s \geq 0$. Luego, $y^{(s)} \notin (\mathcal{A} - \{x^*\})$ para todo s , por lo cual cumple alguna de las dos situaciones mencionadas arriba que garantizan su convergencia a x^* .

$$\lim_{k \rightarrow +\infty} x^{(k)} = \lim_{\substack{k \rightarrow +\infty \\ k \geq M}} x^{(k)} = \lim_{s \rightarrow +\infty} x^{(s+M)} = \lim_{s \rightarrow +\infty} y^{(s)} = x^*.$$

Luego, la sucesión $\{x^{(k)}\}$ converge a x^* , cualquiera sea $x^{(0)}$. □

Este método permite ejecutar el algoritmo desde cualquier punto inicial sin que se detenga fuera de la solución. Otra idea distinta para evitar esto es realizar una buena elección del punto inicial. En [BS15], Beck y Sabach proponen un método sencillo, al cual denominan método SP (en inglés,

“*Simple Procedure*”), que consiste en tomar el punto inicial $x^{(0)}$ de manera que $f(x^{(0)}) < f(a_i)$ para todo $i = 1, \dots, m$, si ningún nodo es solución. Si alguno de los nodos fuera solución, el método daría como resultado dicho nodo. Este método utiliza $T_{VZ}(a_i) = S_i$ para hallar un punto con menor valor funcional.

Algoritmo 2: Método SP

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$

Resultado: $x^{(0)} \in \mathbb{R}^n$

Determinar $i = \operatorname{argmin}\{f(a_j) : 1 \leq j \leq m\}$;

si $\|R_i\| \leq w_i$ **entonces**

$x^{(0)} \leftarrow a_i$;

 “El nodo a_i es solución”;

 Parar;

en otro caso

$x^{(0)} \leftarrow S_i$;

Si algún nodo es solución, el método SP lo encuentra y no es necesario aplicar el método de Weiszfeld o de Vardi y Zhang. Si ningún nodo es solución, el método SP provee un punto inicial $x^{(0)}$ partiendo del nodo a_i con menor valor funcional. Como $f(a_i) \leq f(a_j)$ para todo $j = 1, \dots, m$ y $f(x^{(0)}) = f(S_i) < f(a_i)$ por el Lema 3.3.1, el punto $x^{(0)}$ satisface que $f(x^{(0)}) < f(a_j)$ para todo $j = 1, \dots, m$. La sucesión generada por el método de Weiszfeld partiendo de $x^{(0)}$ satisface que $f(x^{(k)}) \leq f(x^{(0)}) < f(a_j)$ para todo $j = 1, \dots, m$ y todo $k \in \mathbb{N}$ y por lo tanto $x^{(k)} \notin \mathcal{A}$ para todo k . Luego, por el Teorema 3.2.6, la sucesión converge.

En conclusión, elegir el punto inicial a partir del método SP es una buena estrategia para garantizar la convergencia del método de Weiszfeld.

Capítulo 4

Variantes y generalizaciones del algoritmo de Weiszfeld

4.1. El problema de Weber para norma p

El problema de Fermat-Weber fue formulado originalmente para la distancia euclídea, es decir, utilizando la norma $\|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$. Sin embargo, en algunas aplicaciones podría ser más acertado utilizar otras medidas de distancia. Una opción es la norma rectangular, o norma 1, dada por $\|x\|_1 = |x_1| + \dots + |x_n|$. Debido a que la principal aplicación del problema de Fermat-Weber es la de ubicar instalaciones minimizando distancias hacia otros puntos de oferta y demanda, si estas instalaciones se encontraran dentro de una ciudad, pareciera ser acertado utilizar la norma rectangular. Sin embargo, algunos autores han estudiado distintas medidas para modelar el transporte en distintas rutas (por ejemplo, [LM72] y [LM79]), y concluyeron que tomar $\|x\|_p = (|x_1|^p + \dots + |x_n|^p)^{\frac{1}{p}}$ con $1 < p < 2$ en ocasiones aproxima mejor las distancias en distintas regiones geográficas.

Empirically Determined p Values for 17 Regions			
Geographical Region	p Value	Geographical Region	p Value
Australia	1.7545	Brussels	1.7802
British Columbia	1.7080	London City	1.7901
Canada	1.4584	London North	1.6171
France	1.7417	Los Angeles	1.5684
Great Britain	1.8826	New York City	1.6975
New York State	1.4950	Paris	1.6649
Pennsylvania	1.6958	Sydney	1.3940
United States	1.7427	Tokyo	1.8901
		Toronto	1.1261

Figura 4.1: Valores de p determinados empíricamente para ciertas regiones (extraído de [BL93])

Teniendo en cuenta esto, una generalización al problema de Fermat-Weber es cambiar la norma euclídea por la norma p . Con este cambio, el problema en norma p definido en \mathbb{R}^n es el siguiente:

El problema de Fermat-Weber en norma p :

$$\text{Minimizar } f^{(p)}(x) = \sum_{i=1}^m w_i \|x - a_i\|_p = \sum_{i=1}^m w_i \left(\sum_{j=1}^n |x_j - a_{ij}|^p \right)^{\frac{1}{p}} \quad (4.1.1)$$

donde x_j y a_i denotan la coordenada j de x y a_i respectivamente, es decir, $x = (x_1, \dots, x_n)$ y $a_i = (a_{i1}, \dots, a_{in})$ para $i = 1, \dots, m$, y $p \geq 1$ para garantizar que $\|\cdot\|_p$ sea norma. Los nodos a_1, \dots, a_m pueden o no ser colineales.

En la Sección 2.2 demostramos la continuidad y convexidad de la función f , y la existencia de minimizadores. Los argumentos presentados para probar el Lema 2.2.1 y el teorema 2.2.2 utilizan propiedades de la norma euclídea que son comunes para toda norma, es decir que valen también para la norma $p > 1$. Estos resultados demuestran el siguiente teorema:

Teorema 4.1.1. *Dado $p > 1$, la función $f^{(p)}$ definida como en (4.1.1) es continua y convexa. Además, tiene al menos un minimizador local, y todo mínimo local es un mínimo global.*

Para resolver el problema en norma rectangular, no es necesario ejecutar un método iterativo, sino que es posible hallar una solución explícitamente. El problema (4.1.1) para el caso $p = 1$ puede escribirse como

$$f^{(1)}(x) = \sum_{i=1}^m w_i \left(\sum_{j=1}^n |x_j - a_{ij}| \right) = \sum_{j=1}^n \left(\sum_{i=1}^m w_i |x_j - a_{ij}| \right) = \sum_{j=1}^n f_j^{(1)}(x_j) \quad (4.1.2)$$

donde $f_j^{(1)}(x_j) = \sum_{i=1}^m w_i |x_j - a_{ij}| \geq 0$ depende sólo de la j -ésima coordenada de x y de a_1, \dots, a_m . Luego, resolverlo es equivalente a resolver

$$\text{Minimizar } f_j^{(1)}(t) = \sum_{i=1}^m w_i |t - a_{ij}| \text{ con } t \in \mathbb{R} \text{ para } j = 1, \dots, n. \quad (4.1.3)$$

La solución $x^* = (x_1^*, \dots, x_n^*)$ de (4.1.2) satisface que $x_j^* = t_j^*$ para $j = 1, \dots, n$, donde t_j^* es el minimizador de $f_j^{(1)}$.

Al comienzo de la Sección 2.2, presentamos la fórmula cerrada para resolver el problema de Fermat-Weber en norma euclídea con nodos colineales, sabiendo que éste era equivalente a resolver un problema unidimensional de la forma:

$$\text{Minimizar } g(t) = \sum_{i=1}^m w_i |t - t_i|.$$

Para hallar la solución utilizamos el Teorema 2.2.4, que Love, Morris y Wesolowsky [LMW88] formularon para resolver el problema (4.1.3). Para este problema, el teorema se aplica para minimizar cada $f_j^{(1)}$ pero reescribiéndolas de la siguiente manera: para cada j , reordenamos la j -ésima coordenada de los nodos, a_{1j}, \dots, a_{mj} , de menor a mayor y sin repetición, obteniendo $a_{(1)j} < \dots < a_{(r_j)j}$ con $r_j \leq m$, y llamamos $w_1^j, \dots, w_{r_j}^j$ a sus respectivos pesos, sumando los pesos de las coordenadas coincidentes. Con esta notación, la función $f_j^{(1)}$ resulta $f_j^{(1)}(t) = \sum_{i=1}^{r_j} w_i^j |t - a_{(i)j}|$.

Ejemplo 4.1.2. Si $m = 3$ y $n = 2$, con $a_1 = (1, 1)$, $a_2 = (2, 1)$, $a_3 = (1, 2)$, $w_1 = w_2 = w_3 = 1$, entonces obtendríamos $r_1 = 2$, $a_{(1)1} = 1$, $a_{(2)1} = 2$, $w_1^1 = 2$, $w_2^1 = 1$ y $f_1^{(1)}(t) = 2|t - 1| + 1|t - 2|$.

El Teorema 2.2.4 reformulado para resolver el problema (4.1.3) es entonces:

Teorema 4.1.3. *Sea $j \in \{1, \dots, n\}$ y sea $s_j \in \{1, \dots, r_j\}$ tal que que*

$$\sum_{i=1}^{s_j-1} w_i^j - \sum_{i=s_j}^{r_j} w_i^j < 0 \quad (4.1.4)$$

y

$$\sum_{i=1}^{s_j} w_i^j - \sum_{i=s_j+1}^{r_j} w_i^j \geq 0. \quad (4.1.5)$$

Si la ecuación (4.1.5) se satisface como una desigualdad estricta, entonces la solución al problema

$$\text{Minimizar } f_j^{(1)}(t) = \sum_{i=1}^m w_i |t - a_{ij}| = \sum_{i=1}^{r_j} w_i^j |t - a_{(i)j}|$$

es $t_j^* = a_{(s_j)j}$. Si se satisface como igualdad, entonces $t_j^* \in [a_{(s_j)j}, a_{(s_j+1)j}]$.

Con este teorema, podemos encontrar explícitamente cada coordenada de las soluciones del problema (4.1.2), las cuales puede no ser únicas. Si $p > 1$ no se conoce una fórmula explícita para hallar las soluciones, pero podemos generalizar el algoritmo de Weiszfeld, incluso si los nodos son colineales. Antes de eso, podemos generalizar las propiedades de la solución dadas en el Teorema 2.2.7 para el problema en norma $p > 1$. El primer inciso del teorema presentaba las condiciones de optimalidad para la solución si ésta no es un nodo gracias a la diferenciabilidad de f en ese punto, lo cual puede aplicarse también a $f^{(p)}$. En el segundo inciso se definía un vector $R_k \in \mathbb{R}^n$ tal que el nodo a_k es solución si y sólo si $\|R_k\| \leq w_k$. Juel y Love [JL81] desarrollaron una propiedad análoga para diversas normas. A continuación presentaremos este resultado, para el caso de la norma p .

Denotaremos con $\text{sgn}(t)$ al signo de $t \in \mathbb{R}$, es decir,

$$\text{sgn}(t) = \begin{cases} 1 & \text{si } t \geq 0 \\ -1 & \text{si } t < 0. \end{cases}$$

Teorema 4.1.4. Consideremos el problema

$$\text{Minimizar } f^{(p)}(x) = \sum_{i=1}^m w_i \|x - a_i\|_p.$$

Dado $x^* \in \mathbb{R}^n$, se satisface:

a. Si $x^* \notin \mathcal{A}$, entonces x^* minimiza $f^{(p)}$ si y sólo si $\nabla f^{(p)}(x^*) = 0$, es decir,

$$\frac{\partial f^{(p)}}{\partial x_j}(x^*) = \sum_{i=1}^m w_i \text{sgn}(x_j^* - a_{ij}) \frac{|x_j^* - a_{ij}|^{p-1}}{\|x^* - a_i\|_p^{p-1}} = 0 \text{ para todo } j = 1, \dots, n. \quad (4.1.6)$$

b. Si $x^* = a_k$ para algún $k \in \{1, \dots, m\}$, entonces x^* minimiza $f^{(p)}$ si y sólo si $\|R_k\|_{\frac{p}{p-1}} \leq w_k$ donde $\|R_k\|_{\frac{p}{p-1}}$ denota la norma $\frac{p}{p-1}$ y $R_k = (R_{k1}, \dots, R_{kn})$ está dado por

$$R_{kj} = \sum_{\substack{i=1 \\ i \neq k}}^m w_i \text{sgn}(a_{kj} - a_{ij}) \frac{|a_{kj} - a_{ij}|^{p-1}}{\|a_k - a_i\|_p^{p-1}}. \quad (4.1.7)$$

Demostración. Para el inciso a., en el que x^* no es un nodo, sabemos que la función $f^{(p)}$ es convexa y diferenciable en un entorno de x^* y por lo tanto $\nabla f(x^*) = 0$ es condición necesaria y suficiente de optimalidad de primer orden.

Para $x^* = a_k$ con $k \in \{1, \dots, m\}$, podemos analizar las derivadas direccionales de $f^{(p)}$ en a_k como realizamos para el caso en norma 2. Dado $v \in \mathbb{R}^n$, sea $f_v^{(p)} : \mathbb{R}^+ \rightarrow \mathbb{R}$ dada por

$$f_v^{(p)}(t) = f^{(p)}(a_k + tv).$$

La derivada direccional de $f^{(p)}$ a partir de a_k en la dirección v está definida como

$$D_v f^{(p)}(a_k) = \lim_{s \rightarrow 0^+} \frac{df_v^{(p)}}{dt}(s).$$

Notemos que $w_k \|a_k + tv - a_k\|_p = w_k \|tv\|_p = w_k |t| \|v\|_p = tw_k \|v\|_p$. Por lo tanto,

$$\begin{aligned} \frac{df_v}{dt}(s) &= \frac{d}{dt} \Big|_s \left(tw_k \|v\|_p + \sum_{\substack{i=1 \\ i \neq k}}^m w_i \|a_k - a_i + tv\|_p \right) \\ &= w_k \|v\|_p + \sum_{\substack{i=1 \\ i \neq k}}^m w_i \left[\sum_{j=1}^n \operatorname{sgn}(a_{kj} + sv_j - a_{ij}) \frac{|a_{kj} + sv_j - a_{ij}|^{p-1}}{\|a_k + sv - a_i\|_p^{p-1}} v_j \right]. \end{aligned}$$

Tomando límite en s , obtenemos

$$\begin{aligned} \lim_{s \rightarrow 0^+} \frac{df_v^{(p)}}{dt}(s) &= w_k \|v\|_p + \sum_{j=1}^n \left[\sum_{\substack{i=1 \\ i \neq k}}^m w_i \operatorname{sgn}(a_{kj} - a_{ij}) \frac{|a_{kj} - a_{ij}|^{p-1}}{\|a_k - a_i\|_p^{p-1}} \right] v_j \\ &= w_k \|v\|_p + \sum_{j=1}^n R_{kj} v_j. \end{aligned}$$

Es decir que existen las derivadas direccionales de f en todas las direcciones, y están dadas por

$$D_v f^{(p)}(a_k) = w_k \|v\|_p + v^T R_k. \quad (4.1.8)$$

a_k minimiza $f^{(p)}$ si y sólo si no hay direcciones de descenso de $f^{(p)}$ a partir de a_k , es decir, si y sólo si $D_v f^{(p)}(a_k) \geq 0$ para todo $v \in \mathbb{R}^n$. A continuación probaremos que esto ocurre si y sólo si $\|R_k\|_{\frac{p}{p-1}} \leq w_k$:

(\Rightarrow) Supongamos que $w_k \|v\|_p + v^T R_k \geq 0$ para todo $v \in \mathbb{R}^n$. Como la norma $\frac{p}{p-1}$ es la norma dual de la norma p , existe $y \in \mathbb{R}^n$, $y \neq 0$, tal que

$$\frac{R_k^T y}{\|y\|_p} = \max_{x \neq 0} \frac{R_k^T x}{\|x\|_p} = \|R_k\|_{\frac{p}{p-1}}.$$

Luego, tomando $v = -y$ y usando que $w_k \|v\|_p \geq -v^T R_k$, obtenemos que

$$\|R_k\|_{\frac{p}{p-1}} = \frac{R_k^T(-v)}{\|-v\|_p} = \frac{-v^T R_k}{\|v\|_p} \leq \frac{w_k \|v\|_p}{\|v\|_p} = w_k$$

es decir, $\|R_k\|_{\frac{p}{p-1}} \leq w_k$.

(\Leftarrow) Supongamos ahora que $\|R_k\|_{\frac{p}{p-1}} \leq w_k$. Aplicando la desigualdad de Hölder, tenemos que

$$-v^T R_k \leq |v^T R_k| \leq \|v\|_p \|R_k\|_{\frac{p}{p-1}} \leq \|v\|_p w_k \text{ para todo } v \in \mathbb{R}^n$$

y por lo tanto $0 \leq \|v\|_p w_k + v^T R_k$ para todo $v \in \mathbb{R}^n$. □

El Teorema 4.1.4 nos asegura que si la solución x^* se encuentra en un punto donde $f^{(p)}$ es diferenciable, entonces

$$0 = \nabla f^{(p)}(x^*) = \left(\frac{\partial f^{(p)}}{\partial x_1}(x^*), \dots, \frac{\partial f^{(p)}}{\partial x_m}(x^*) \right).$$

Las derivadas parciales de $f^{(p)}$ están dadas por

$$\frac{\partial f^{(p)}}{\partial x_j}(x) = \sum_{i=1}^m w_i \operatorname{sgn}(x_j - a_{ij}) \frac{|x_j - a_{ij}|^{p-1}}{\|x - a_i\|_p^{p-1}} \quad j = 1, \dots, n. \quad (4.1.9)$$

Como $t = \operatorname{sgn}(t)|t|$, entonces la ecuación (4.1.9) es equivalente a

$$\frac{\partial f^{(p)}}{\partial x_j}(x) = \sum_{i=1}^m w_i (x_j - a_{ij}) \frac{|x_j - a_{ij}|^{p-2}}{\|x - a_i\|_p^{p-1}} \quad j = 1, \dots, n. \quad (4.1.10)$$

Evaluando en x^* , igualando a cero y manipulando la ecuación $\frac{\partial f^{(p)}}{\partial x_j}(x^*)$, obtenemos que para cada j ,

$$\begin{aligned} \sum_{i=1}^m w_i (x_j^* - a_{ij}) \frac{|x_j^* - a_{ij}|^{p-2}}{\|x^* - a_i\|_p^{p-1}} &= 0 \\ \sum_{i=1}^m w_i x_j^* \frac{|x_j^* - a_{ij}|^{p-2}}{\|x^* - a_i\|_p^{p-1}} &= \sum_{i=1}^m w_i a_{ij} \frac{|x_j^* - a_{ij}|^{p-2}}{\|x^* - a_i\|_p^{p-1}} \\ x_j^* \left(\sum_{i=1}^m w_i \frac{|x_j^* - a_{ij}|^{p-2}}{\|x^* - a_i\|_p^{p-1}} \right) &= \sum_{i=1}^m w_i a_{ij} \frac{|x_j^* - a_{ij}|^{p-2}}{\|x^* - a_i\|_p^{p-1}}, \end{aligned}$$

con lo cual conseguimos la siguiente igualdad:

$$x_j^* = \left[\sum_{i=1}^m w_i \frac{|x_j^* - a_{ij}|^{p-2}}{\|x^* - a_i\|_p^{p-1}} \right]^{-1} \sum_{i=1}^m w_i a_{ij} \frac{|x_j^* - a_{ij}|^{p-2}}{\|x^* - a_i\|_p^{p-1}}. \quad (4.1.11)$$

Esta última ecuación sugiere, al igual que en lo trabajado para norma euclídea, la siguiente sucesión:

$$x_j^{(k+1)} = \left[\sum_{i=1}^m w_i \frac{|x_j^{(k)} - a_{ij}|^{p-2}}{\|x^{(k)} - a_i\|_p^{p-1}} \right]^{-1} \sum_{i=1}^m w_i \frac{|x_j^{(k)} - a_{ij}|^{p-2}}{\|x^{(k)} - a_i\|_p^{p-1}} a_{ij} =: \Phi_j(x^{(k)}). \quad (4.1.12)$$

Notemos que si $p = 2$, esta iteración coincide con la iteración de Weiszfeld. Al igual que la transformación de Weiszfeld, la función $\Phi = (\Phi_1, \dots, \Phi_n)$ no está definida en todo punto. Si $p \geq 2$ la función no está definida en los nodos a_1, \dots, a_m , y para $1 \leq p < 2$, la función es discontinua en nm hiperplanos, $H_{ij} = \{x \in \mathbb{R}^n : x_j = a_{ij}\}$. Brimberg y Love [BL93] presentaron una extensión continua T de la función Φ , para el caso $1 \leq p \leq 2$, y probaron la convergencia del método iterativo dado por $x^{(k+1)} = T(x^{(k)})$, el cual llamaremos *método de Weiszfeld para norma p*. Para el caso $p > 2$, no siempre puede construirse una extensión continua, y en general el método no converge (ver [BL92]), aunque Brimberg, Chen y Chen [BCC98] propusieron una variación del algoritmo que garantiza su convergencia. A continuación, veremos algunos de los resultados que forman parte de la prueba de convergencia propuesta por Brimberg y Love [BL93] para el caso con $1 < p < 2$.

De ahora en más, asumiremos que $1 < p < 2$, y utilizaremos la siguiente notación:

$$\begin{aligned} H_{ij} &= \{x \in \mathbb{R}^n : x_j = a_{ij}\} \quad \text{para } i = 1, \dots, m \text{ y } j = 1, \dots, n, \\ H_j &= \bigcup_{i=1}^m H_{ij} \quad \text{para } j = 1, \dots, n \\ \text{y } S &= \bigcup_{j=1}^n H_j. \end{aligned}$$

La función Φ dada en la ecuación (4.1.12) está bien definida en $\mathbb{R}^n - S$. Además, su coordenada j -ésima, con $j = 1, \dots, m$, está definida para todo $x \notin H_j$ y puede escribirse como

$$\Phi_j(x) = \frac{\sum_{i=1}^m y_{ij}(x) a_{ij}}{\sum_{i=1}^m y_{ij}(x)} \quad \text{con } y_{ij}(x) = w_i \frac{|x_j - a_{ij}|^{p-2}}{\|x - a_i\|_p^{p-1}}. \quad (4.1.13)$$

Podemos definir la extensión $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de la función Φ coordenada a coordenada, del siguiente modo:

$$T_j(x) = \begin{cases} a_{ij} & \text{si } x_j = a_{ij} \text{ para algún } i = 1, \dots, m \\ \Phi_j(x) & \text{caso contrario.} \end{cases} \quad (4.1.14)$$

Si $x \notin S$, entonces $T(x) = \Phi(x)$. Para demostrar que $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es continua, notemos antes que cada función Φ_j se asemeja a una iteración del método del gradiente. En efecto, si $x_j \notin H_j$, $\Phi_j(x)$ satisface que

$$\begin{aligned} \Phi_j(x) &= \frac{\sum_{i=1}^m y_{ij}(x) a_{ij}}{\sum_{i=1}^m y_{ij}(x)} \\ &= x_j - \frac{\sum_{i=1}^m y_{ij}(x) (x_j - a_{ij})}{\sum_{i=1}^m y_{ij}(x)} \\ &= x_j - \left[\sum_{i=1}^m y_{ij}(x) \right]^{-1} \frac{\partial f^{(p)}}{\partial x_j}(x). \end{aligned}$$

Ahora sí, podemos probar la continuidad de la transformación T .

Lema 4.1.5. *La función $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por (4.1.14) es continua.*

Demostración. Para facilitar las cuentas, probaremos que T es continua coordenada a coordenada. Sea T_j la j -ésima coordenada de T .

Es claro que T_j es continua en $\mathbb{R}^n - H_j$ pues para cada $i = 1, \dots, m$, las funciones $\|x - a_i\|_p$ y $|x_j - a_{ij}|$ son continuas y no se anulan en $\mathbb{R}^n - H_j$.

Veamos que si $y \in H_{ij}$, entonces $\lim_{x \rightarrow y} T_j(x) = a_{ij}$:

Como $p - 2 < 0$ y $\lim_{x \rightarrow y} x_j = y_j = a_{ij}$, tenemos que

$$\lim_{x \rightarrow y} y_{ij}(x) = \lim_{x \rightarrow y} w_i \frac{|x_j - a_{ij}|^{p-2}}{\|x - a_i\|_p^{p-1}} = +\infty.$$

Además, si $k \neq i$,

$$\lim_{x \rightarrow y} y_{kj}(x) = \lim_{x \rightarrow y} w_k \frac{|x_j - a_{kj}|^{p-2}}{\|x - a_k\|_p^{p-1}} \geq 0,$$

pudiendo ser $+\infty$. Luego,

$$\lim_{x \rightarrow y} \sum_{k=1}^m y_{kj}(x) = +\infty.$$

Además, $\frac{\partial f^{(p)}}{\partial x_j}(x)$ es acotado pues

$$\left| \frac{\partial f^{(p)}}{\partial x_j}(x) \right| = \left| \sum_{i=1}^m w_i \operatorname{sgn}(x_j - a_{ij}) \frac{|x_j - a_{ij}|^{p-1}}{\|x - a_i\|_p^{p-1}} \right| \leq \sum_{i=1}^m w_i \left| \frac{|x_j - a_{ij}|}{\|x - a_i\|_p} \right|^{p-1} \leq \sum_{i=1}^m w_i.$$

Luego

$$\lim_{x \rightarrow y} \frac{1}{\sum_{k=1}^m y_{kj}(x)} \frac{\partial f^{(p)}}{\partial x_j}(x) = 0 \in \mathbb{R}$$

y

$$\lim_{x \rightarrow y} T_j(x) = \lim_{x \rightarrow y} \left(x_j - \frac{1}{\sum_{k=1}^m y_{kj}(x)} \frac{\partial f^{(p)}}{\partial x_j}(x) \right) = a_{ij}.$$

□

Lema 4.1.6. *La imagen de la transformación T , $\{T(x) : x \in \mathbb{R}^n\}$, es acotada.*

Demostración. Vimos que para $x \notin H_j$, la j -ésima coordenada de Φ es

$$\Phi_j(x) = \sum_{i=1}^m \frac{y_{ij}(x)}{\sum_{k=1}^m y_{kj}(x)} a_{ij}.$$

Como $0 < y_{ij}$ para todo i , entonces $0 \leq \frac{y_{ij}(x)}{\sum_{k=1}^m y_{kj}(x)} \leq 1$. Además, $\sum_{i=1}^m \frac{y_{ij}(x)}{\sum_{k=1}^m y_{kj}(x)} = 1$. Luego, $\Phi_j(x)$ es combinación convexa de a_{1j}, \dots, a_{mj} y por lo tanto

$$\min_{1 \leq i \leq m} a_{ij} \leq \Phi_j(x) \leq \max_{1 \leq i \leq m} a_{ij}.$$

$$T_j(x) = \begin{cases} a_{ij} & \text{si } x_j \in H_{ij}, i = 1, \dots, m \\ \Phi_j(x) & \text{caso contrario.} \end{cases}$$

Entonces,

$$\min_{1 \leq i \leq m} a_{ij} \leq T_j(x) \leq \max_{1 \leq i \leq m} a_{ij}$$

y por lo tanto $T(x)$ está contenido en el hipercubo

$$\mathbf{C} = \left\{ y \in \mathbb{R}^n : \min_{1 \leq i \leq m} a_{ij} \leq y_j \leq \max_{1 \leq i \leq m} a_{ij} \text{ para todo } j = 1, \dots, n \right\}.$$

□

La siguiente propiedad es muy importante, pues es la que determina la relación entre los puntos fijos de T y la solución del problema (4.1.1).

Lema 4.1.7. *Sea x^* la solución al problema de Fermat-Weber en norma p , y sea $\{x^{(k)}\}$ la sucesión generada por el algoritmo de Weiszfeld para norma p partiendo de $x^{(0)} \in \mathbb{R}^n$. Entonces:*

Si $x^{(K)} = x^$ para algún $K \in \mathbb{N}$, entonces $x^{(k)} = x^*$ para todo $k \geq K$.*

Si $x^{(K+1)} \notin S$ y $x^{(K+1)} = x^{(K)}$ para algún $K \in \mathbb{N}$, entonces $x^{(K)} = x^$.*

Demostración. Para la primer afirmación, consideremos por separado los casos $x^* \in S$ y $x^* \notin S$.

Si $x^* \notin S$, entonces $\nabla f^{(p)}(x^*) = 0$ y por (4.1.11), $x^* = \Phi(x^*) = T(x^*)$. Luego, si $x^{(K)} = x^*$, obtenemos que $x^{(K+1)} = T(x^{(K)}) = T(x^*) = x^*$. Inductivamente puede probarse que $x^{(k)} = x^*$ para todo $k > K$.

Si $x^* \in S$, sea $\mathcal{J} = \{j : x^* \notin H_j\}$. Si $j \in \mathcal{J}$, entonces $\frac{\partial f^{(p)}}{\partial x_j}(x^*) = 0$, por lo que $\Phi_j(x^*) = x_j^*$. Si $j \notin \mathcal{J}$, entonces $x^* \in H_j = \bigcup_{1 \leq i \leq m} H_{ij}$, es decir que existe $i_j \in \{1, \dots, m\}$ tal que $x_j^* = a_{i_j j}$.

Si $x^{(K)} = x^*$, por (4.1.14) obtenemos que

$$x_j^{(K+1)} = T_j(x^{(K)}) = \begin{cases} \Phi_j(x^*) = x_j^* & \text{si } j \in \mathcal{J} \\ a_{i_j j} = x_j^* & \text{si } j \notin \mathcal{J}. \end{cases}$$

Luego, $x^{(K+1)} = x^*$ e inductivamente se prueba que $x^{(k)} = x^*$ para todo $k > K$.

Para la segunda afirmación del teorema, si $x^{(K)} \notin S$, entonces $x^{(K+1)} = T(x^{(K)}) = \Phi(x^{(K)})$. Luego $x^{(K)} = \Phi(x^{(K)})$ y por lo tanto $\nabla f^{(p)}(x^{(K)}) = 0$, por lo que $x^{(K)} = x^*$. □

La siguiente propiedad indispensable para la prueba de convergencia es la monotonicidad de $f^{(p)}$ respecto a la transformación T .

Lema 4.1.8. *Si $T(x) \neq x$, entonces $f^{(p)}(T(x)) < f^{(p)}(x)$.*

Demostración. A lo largo de esta demostración, denotaremos $y_{ij} = y_{ij}(x)$ para todo $i = 1, \dots, m$ y $j = 1, \dots, n$. Con esta notación, si $x \notin H_j$,

$$T_j(x) = \frac{\sum_{i=1}^m y_{ij} a_{ij}}{\sum_{i=1}^m y_{ij}}$$

Sea $\mathcal{J} = \{j : x \notin H_j\}$. Como $T(x) \neq x$, entonces $\mathcal{J} \neq \emptyset$.

Para $j \in \mathcal{J}$, definamos

$$g_j(t) = \sum_{i=1}^m y_{ij} (t - a_{ij})^2 \quad \text{para } t \in \mathbb{R}. \quad (4.1.15)$$

g_j es una función estrictamente convexa por ser cuadrática positiva, y tiene su único mínimo en $T_j(x) = \Phi_j(x)$, pues

$$\begin{aligned} g'_j(t) &= \sum_{i=1}^m 2y_{ij} (t - a_{ij}) \\ g'_j(\Phi_j(x)) &= \sum_{i=1}^m 2y_{ij} \left(\frac{\sum_{k=1}^m y_{kj} a_{kj}}{\sum_{k=1}^m y_{kj}} - a_{ij} \right) \\ &= \sum_{i=1}^m 2y_{ij} \frac{\sum_{k=1}^m y_{kj} a_{kj} - \sum_{k=1}^m y_{kj} a_{ij}}{\sum_{k=1}^m y_{kj}} \\ &= \frac{2}{\sum_{k=1}^m y_{kj}} \sum_{i=1}^m \sum_{k=1}^m y_{ij} y_{kj} (a_{kj} - a_{ij}) \end{aligned}$$

y como $y_{ij} y_{kj} (a_{kj} - a_{ij}) = -y_{kj} y_{ij} (a_{ij} - a_{kj})$ para todo $i, k = 1, \dots, m$, resulta $g'_j(\Phi_j(x)) = 0$. Luego, para todo $j \in \mathcal{J}$ se tiene que

$$g_j(T_j(x)) \leq g_j(x_j) = \sum_{i=1}^m \left(\frac{w_i |x_j - a_{ij}|^{p-2}}{\|x - a_i\|_p^{p-1}} \right) (x_j - a_{ij})^2 = \sum_{i=1}^m \frac{w_i |x_j - a_{ij}|^p}{\|x - a_i\|_p^{p-1}} \quad (4.1.16)$$

y la desigualdad es estricta para al menos un j , pues $T(x) \neq x$.

Si $j \notin \mathcal{J}$, entonces $T_j(x) = x_j$. Si definimos

$$h_j(t) = \sum_{i=1}^m \frac{w_i |t - a_{ij}|^p}{\|x - a_i\|_p^{p-1}} \quad \text{para } t \in \mathbb{R} \quad (4.1.17)$$

obtenemos que

$$h_j(T_j(x)) = h_j(x_j) \quad \text{para todo } j \notin \mathcal{J}. \quad (4.1.18)$$

Combinando (4.1.16) y (4.1.18), tenemos que

$$\begin{aligned} \sum_{j \in \mathcal{J}} g_j(T_j(x)) + \sum_{j \notin \mathcal{J}} h_j(T_j(x)) &< \sum_{j \in \mathcal{J}} g_j(x_j) + \sum_{j \notin \mathcal{J}} h_j(x_j) \\ &= \sum_{j=1}^n \sum_{i=1}^m \frac{w_i |x_j - a_{ij}|^p}{\|x - a_i\|_p^{p-1}} \\ &= \sum_{i=1}^m \frac{w_i \|x - a_i\|_p^p}{\|x - a_i\|_p^{p-1}} \\ &= \sum_{i=1}^m w_i \|x - a_i\|_p \end{aligned}$$

por lo cual

$$\sum_{j \in \mathcal{J}} g_j(T_j(x)) + \sum_{j \notin \mathcal{J}} h_j(T_j(x)) < f^{(p)}(x). \quad (4.1.19)$$

Por otro lado, tenemos que

$$\begin{aligned} \sum_{j \in \mathcal{J}} g_j(T_j(x)) + \sum_{j \notin \mathcal{J}} h_j(T_j(x)) &= \sum_{j \in \mathcal{J}} \sum_{i=1}^m \frac{w_i |x_j - a_{ij}|^{p-2} |T_j(x) - a_{ij}|^2}{\|x - a_i\|_p^{p-1}} + \sum_{j \notin \mathcal{J}} \sum_{i=1}^m \frac{w_i |x_j - a_{ij}|^p}{\|x - a_i\|_p^{p-1}} \\ &= \sum_{j=1}^n \sum_{i=1}^m \frac{w_i |x_j - a_{ij}|^{p-2} |T_j(x) - a_{ij}|^2}{\|x - a_i\|_p^{p-1}} \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|_p^{p-1}} (|x_j - a_{ij}|^p)^{1-\frac{2}{p}} (|T_j(x) - a_{ij}|^p)^{\frac{2}{p}}. \end{aligned} \quad (4.1.20)$$

Puede probarse que si $a > 0$, $b > 0$ y $0 < q < 1$ se satisface que

$$a^{1-\frac{1}{q}} b^{\frac{1}{q}} \geq \left(1 - \frac{1}{q}\right) a + \frac{1}{q} b. \quad (4.1.21)$$

Una demostración completa de esta afirmación puede verse por ejemplo en la Sección 1.14 del libro [BB71]. Aplicando la desigualdad (4.1.21) a la ecuación (4.1.20), tomando $q = \frac{p}{2} \in (\frac{1}{2}, 1)$ obtenemos que

$$\begin{aligned} \sum_{j \in \mathcal{J}} g_j(T_j(x)) + \sum_{j \notin \mathcal{J}} h_j(T_j(x)) &\geq \sum_{i=1}^m \frac{w_i}{\|x - a_i\|_p^{p-1}} \left[\left(1 - \frac{2}{p}\right) |x_j - a_{ij}|^p + \frac{2}{p} |T_j(x) - a_{ij}|^p \right] \\ &= \frac{p-2}{p} \sum_{i=1}^m w_i \|x - a_i\|_p + \frac{2}{p} \sum_{i=1}^m w_i \|x - a_i\|_p^{1-p} \|T(x) - a_i\|_p^p. \end{aligned}$$

Usando nuevamente (4.1.21) con $q = \frac{1}{p} \in (\frac{1}{2}, 1)$, resulta

$$\begin{aligned} \sum_{j \in \mathcal{J}} g_j(T_j(x)) + \sum_{j \notin \mathcal{J}} h_j(T_j(x)) &\geq \frac{p-2}{p} f^{(p)}(x) + \frac{2}{p} \sum_{i=1}^m w_i \left[(1-p) \|x - a_i\|_p + p \|T(x) - a_i\|_p \right] \\ &= \frac{p-2}{p} f^{(p)}(x) + \frac{2(1-p)}{p} f^{(p)}(x) + 2f^{(p)}(T(x)) \\ &= -f^{(p)}(x) + 2f^{(p)}(T(x)). \end{aligned} \quad (4.1.22)$$

De (4.1.19) y (4.1.22) obtenemos que

$$-f^{(p)}(x) + 2f^{(p)}(T(x)) \leq \sum_{j \in \mathcal{J}} g_j(T_j(x)) + \sum_{j \notin \mathcal{J}} h_j(T_j(x)) < f^{(p)}(x)$$

y por lo tanto

$$f^{(p)}(T(x)) < f^{(p)}(x). \quad (4.1.23)$$

□

Por la convexidad de $f^{(p)}$, el Lema 4.1.8 nos asegura que, si todas las iteraciones satisfacen $x^{(k)} \notin S$, entonces cada iteración está más cerca de la solución que la anterior. En efecto, la sucesión generada por el método de Weiszfeld modificado converge a la solución si ninguna iteración pertenece a S . Una demostración completa de la convergencia es presentada por Brimberg y Love [BL93].

4.2. Aceleración del algoritmo

Hemos visto que el algoritmo de Weiszfeld en norma 2 tiene tasa de convergencia cuadrática en algunos casos, pero a lo sumo lineal en la mayoría. Muchos autores han tratado de mejorar esta tasa de convergencia mediante distintos métodos. En este capítulo presentaremos algunas de las variaciones para acelerar el algoritmo de Weiszfeld.

Katz [Kat74] sugirió aplicar el método de Steffensen, el cual podría llevar a una convergencia cuadrática del algoritmo, pero cuya implementación es muy costosa. Otros autores propusieron aprovechar que el algoritmo es de la forma

$$x^{(k+1)} = x^{(k)} - \frac{1}{L(x)} \nabla f(x^{(k)})$$

y cambiar la longitud del paso para acelerar la convergencia, multiplicándolo por un factor $\lambda > 0$, es decir,

$$x^{(k+1)} = x^{(k)} - \frac{\lambda}{L(x)} \nabla f(x^{(k)}). \quad (4.2.1)$$

Ostreich [OJ78] demostró que para $n = 2$, el algoritmo generado por (4.2.1) converge para cualquier $\lambda \in [1, 2]$.

Drezner [Dre92] y [Dre95] estudia el problema en \mathbb{R}^n , y propone variar el valor de λ en cada iteración. Sin embargo, esto puede ser costoso y en muchos casos no reporta una diferencia significativa respecto a tomar un valor de λ fijo.

En [BCC98], Brimberg, Chen y Chen extendieron el método dado por (4.2.1) al problema de Fermat norma $p > 1$, tomando

$$x_j^{(k+1)} = x_j^{(k)} - \frac{\lambda}{S_j(x)} \frac{\partial f^{(p)}}{\partial x_k}(x^{(k)}), \quad j = 1, \dots, n \quad (4.2.2)$$

donde S_j se define de manera similar a $L(x)$, como veremos más adelante, y demostraron su convergencia para ciertos valores de λ , considerando siempre el problema en el plano \mathbb{R}^2 . A continuación generalizaremos sus resultados a \mathbb{R}^n , para el caso $1 < p \leq 2$.

Generalización del trabajo de Brimberg, Chen y Chen [BCC98] a \mathbb{R}^n :

Recordemos que si $x^{(k)}$ no es un nodo, la iteración $k + 1$ del algoritmo de Weiszfeld para norma p puede escribirse coordenada a coordenada como

$$x_j^{(k+1)} = \Phi_j(x^{(k)}) = x_j^{(k)} - \frac{1}{S_j(x^{(k)})} \frac{\partial f^{(p)}}{\partial x_j}(x^{(k)}) \quad (4.2.3)$$

donde la función S_j está dada por $S_j(x) = \sum_{i=1}^m y_{ij}(x)$, con

$$y_{ij}(x^{(k)}) = w_i \frac{|x_j^{(k)} - a_{ij}|^{p-2}}{\|x^{(k)} - a_i\|_p^{p-1}}.$$

El método propuesto por Brimberg, Chen y Chen para acelerar la convergencia es un método iterativo donde cada iteración está dada por $x^{(k+1)} = \psi(x^{(k)})$, con

$$\psi_j(x) = x_j - \frac{\lambda}{S_j(x)} \frac{\partial f^{(p)}}{\partial x_j}(x) \quad \text{para } j = 1, \dots, n, \quad (4.2.4)$$

donde λ es una constante positiva.

Para estudiar la convergencia del método, debemos estudiar la jacobiana de la función ψ .

$$\frac{\partial \psi_j}{\partial x_k}(x) = \delta_{kj} - \frac{\lambda}{S_j^2(x)} \left(\frac{\partial^2 f^{(p)}}{\partial x_k \partial x_j}(x) S_j(x) - \frac{\partial f^{(p)}}{\partial x_j}(x) \frac{\partial S_j}{\partial x_k}(x) \right) \quad (4.2.5)$$

donde δ_{kj} es la función *delta* de Kronecker. Si x^* es minimizador de $f^{(p)}$ y las derivadas parciales existen en ese punto, entonces $\frac{\partial f^{(p)}}{\partial x_j}(x^*) = 0$ para todo $j = 1, \dots, n$, y por lo tanto

$$\frac{\partial \psi_j}{\partial x_k}(x) = \delta_{kj} - \lambda \frac{\partial^2 f^{(p)}}{\partial x_k \partial x_j}(x^*) \frac{1}{S_j(x^*)}. \quad (4.2.6)$$

Si denotamos con S a la matriz diagonal con $S_{jj} = \frac{1}{S_j(x^*)}$, y con H a la matriz de derivadas segundas de $f^{(p)}$ en x^* , es decir

$$S = \begin{pmatrix} \frac{1}{S_1(x^*)} & 0 & \cdots & 0 \\ 0 & \frac{1}{S_2(x^*)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{S_n(x^*)} \end{pmatrix} \quad \text{y} \quad H = \begin{pmatrix} \frac{\partial^2 f^{(p)}}{\partial x_1 \partial x_1}(x^*) & \cdots & \frac{\partial^2 f^{(p)}}{\partial x_1 \partial x_n}(x^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f^{(p)}}{\partial x_n \partial x_1}(x^*) & \cdots & \frac{\partial^2 f^{(p)}}{\partial x_n \partial x_n}(x^*) \end{pmatrix}. \quad (4.2.7)$$

entonces la jacobiana de ψ en x^* es $\psi'(x^*) = Id - \lambda SH$.

Para el caso $1 < p \leq 2$, si los nodos no son colineales podemos asegurar que $f^{(p)}$ es estrictamente convexa y que H es definida positiva, como demuestran Brimberg y Love en [BL99]. De aquí en más, asumiremos que esto se cumple.

Como $S_j(x^*) > 0$ para todo j , la matriz S es definida positiva y por lo tanto SH también lo es. Sean $\alpha_1, \dots, \alpha_n$ los autovalores de SH . Los autovalores de $\psi'(x^*)$ son

$$\sigma_j = 1 - \lambda \alpha_j, \quad j = 1, \dots, n$$

y como $\alpha_j > 0$ para todo $j = 1, \dots, m$ y $\lambda > 0$, entonces $\sigma_j < 1$ para todo j . Usando esto podemos probar el siguiente teorema.

Teorema 4.2.1. *Si el operador ψ satisface que*

$$\text{tr}[\psi'(x^*)] \geq n - 2 \quad (4.2.8)$$

entonces el algoritmo $x^{(k+1)} = \psi(x^{(k)})$ converge localmente a x^* .

Demostración. Por (4.2.8) tenemos que $\sum_{j=1}^n \sigma_j = \text{tr}[\psi'(x^*)] \geq n - 2$. Además, habíamos probado que $\sigma_j < 1$ para todo $j = 1, \dots, n$, es decir, $-\sigma_j > -1$ para todo j . Luego,

$$\sigma_k \geq (n - 2) - \sum_{\substack{j=1 \\ j \neq k}}^n \sigma_j > (n - 2) - (n - 1) = -1 \quad \text{para todo } k = 1, \dots, n \quad (4.2.9)$$

por lo que $-1 < \sigma_k < 1$ para todo $k = 1, \dots, n$.

El radio espectral de $\psi'(x^*)$ es entonces $\rho = \max\{|\sigma_1|, \dots, |\sigma_n|\} < 1$, lo cual garantiza la convergencia local del algoritmo (ver por ejemplo [OR70]). Más aún, la convergencia es al menos lineal, pues ρ da una cota superior para el factor de convergencia asintótica. \square

Estudiemos entonces la traza de $\psi'(x^*)$. Puede probarse que

$$\begin{aligned} \frac{\partial^2 f^{(p)}}{\partial x_j^2}(x) &= \sum_{i=1}^m w_i (p-1) \frac{|x_j - a_{ij}|^{p-2}}{\|x - a_i\|_p^{p-1}} \left(1 - \frac{|x_j - a_{ij}|^p}{\|x - a_i\|_p^p} \right) \\ &= (p-1) \sum_{i=1}^m y_{ij}(x) \left(1 - \frac{|x_j - a_{ij}|^p}{\|x - a_i\|_p^p} \right) \\ &= (p-1) \left(S_j(x) - \sum_{i=1}^m y_{ij}(x) \frac{|x_j - a_{ij}|^p}{\|x - a_i\|_p^p} \right). \end{aligned}$$

Entonces

$$\begin{aligned}
\operatorname{tr}[\psi'(x^*)] &= \operatorname{tr}[Id] - \lambda \operatorname{tr}[SH] \\
&= n - \lambda \sum_{j=1}^n \frac{1}{S_j(x^*)} \frac{\partial^2 f^{(p)}}{\partial x_j^2}(x^*) \\
&= n - \lambda(p-1) \sum_{j=1}^n \frac{1}{S_j(x^*)} \left(S_j(x^*) - \sum_{i=1}^m y_{ij}(x^*) \frac{|x_j^* - a_{ij}|^p}{\|x^* - a_i\|_p^p} \right) \\
&= n - \lambda(p-1) \sum_{j=1}^n \left(1 - \frac{1}{S_j(x^*)} \sum_{i=1}^m y_{ij}(x^*) \frac{|x_j^* - a_{ij}|^p}{\|x^* - a_i\|_p^p} \right)
\end{aligned}$$

es decir que

$$\operatorname{tr}[\psi'(x^*)] = n - \lambda(p-1)(n - Q^*) \quad \text{con} \quad Q^* = \sum_{j=1}^n \frac{1}{S_j(x^*)} \sum_{i=1}^m y_{ij}(x^*) \frac{|x_j^* - a_{ij}|^p}{\|x^* - a_i\|_p^p}. \quad (4.2.10)$$

Ahora,

$$\frac{|x_j^* - a_{ij}|^p}{\|x^* - a_i\|_p^p} \leq 1 \quad \text{para todo } i = 1, \dots, m, j = 1, \dots, n$$

donde la igualdad se da si y sólo si $x_l^* = a_{il}$ para todo $l \neq j$. Entonces la desigualdad es estricta en al menos un par i, j y obtenemos que $0 < Q^* < n$ y por lo tanto $0 < n - Q^* < n$. Luego $n - 2 \leq \operatorname{tr}[\psi'(x^*)] = n - \lambda(p-1)(n - Q^*)$ si y sólo si

$$\lambda \leq \frac{2}{(p-1)(n - Q^*)}. \quad (4.2.11)$$

Por el Teorema 4.2.1, para todo λ que satisfaga (4.2.11) la sucesión dada por $x^{(k+1)} = \psi(x^{(k)})$ converge a x^* . Pero Q^* es una constante desconocida, pues depende de x^* , por lo que es necesario aproximar el valor de λ .

Sabemos que $n - Q^* < n$, por lo que tomar

$$\lambda = \frac{2}{n(p-1)} \quad (4.2.12)$$

garantiza que se satisface la ecuación (4.2.8) y que la sucesión converge.

Para $p = 2$, como $y_{ij}(x^*) = \frac{w_i}{\|x^* - a_i\|}$ para todo $j = 1, \dots, n$ y todo $i = 1, \dots, m$, podemos calcular Q^* de manera exacta:

$$\begin{aligned}
Q^* &= \sum_{j=1}^n \frac{1}{\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|}} \sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} \frac{(x_j^* - a_{ij})^2}{\|x^* - a_i\|^2} \\
&= \frac{1}{\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|}} \sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|^3} \sum_{j=1}^n (x_j^* - a_{ij})^2 \\
&= \frac{1}{\sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|}} \sum_{i=1}^m \frac{w_i}{\|x^* - a_i\|} \\
&= 1.
\end{aligned}$$

Entonces para $p = 2$ podemos asegurar la convergencia para todo $\lambda \leq \frac{2}{n-1}$. Además, podríamos asumir que para otros valores de p , $Q^* \approx 1$ y estimar la cota superior por

$$\lambda = \frac{2}{(p-1)(n-1)}. \quad (4.2.13)$$

Para el caso $n = 2$, que es el caso con más aplicaciones conocidas, las cotas para λ sugeridas arriba permiten realizar pasos más amplios en cada iteración. En efecto, por la ecuación (4.2.12) podemos tomar $\lambda = \frac{1}{p-1}$ sabiendo que la convergencia está garantizada, y como $1 < p \leq 2$, entonces $\lambda \geq 1$. Además, la aproximación dada por (4.2.13) sugiere utilizar $\lambda = \frac{2}{p-1}$, y si suponemos que la convergencia está garantizada también para este valor, podemos elegir cualquier $\lambda \in \left[\frac{1}{p-1}, \frac{2}{p-1} \right]$. En [BCC98], proponen por ejemplo utilizar

$$\lambda = \frac{p}{p-1} \quad (4.2.14)$$

y numéricamente obtienen que este paso es el óptimo, entre (4.2.12), (4.2.13) y (4.2.14).

Para el caso $n > 2$, el valor propuesto por (4.2.12) puede no ser mayor que 1. Para $n = 2$ por ejemplo, obtenemos $\lambda = \frac{2}{n}$, lo cual no es recomendable para acelerar la convergencia, pues se estarían haciendo pasos más cortos. Para el caso $p = 1 + \varepsilon$ con $0 < \varepsilon < 1$ en cambio, es posible obtener valores mayores de λ , y que aumentan a medida que p decrece. Si $\varepsilon < \frac{2}{n}$, podemos tomar valores de λ mayores que 1, pues $\frac{2}{(p-1)n} = \frac{2}{\varepsilon n} > 1$. La ecuación (4.2.13) sugiere que para $\varepsilon < \frac{2}{n-1}$ pueden encontrarse pasos más largos.

Si bien no realizamos la generalización para el caso en que $p > 2$, Brimberg, Chen y Chen [BCC98] afirman que, construyendo el paso λ de manera similar, puede asegurarse la convergencia, incluso aunque no esté garantizada para el algoritmo de Weiszfeld original. La demostración de la convergencia es muy similar a la demostración para el caso $1 < p \leq 2$, con la diferencia de que varían las condiciones necesarias para garantizar que la función $f^{(p)}$ sea estrictamente convexa. Los valores de λ sugeridos serían los dados por las ecuaciones (4.2.12), (4.2.13) y (4.2.14).

En la Sección 6.1 veremos algunos resultados numéricos variando n y p .

4.3. El problema con restricciones de cotas en las variables

Volvamos al problema de Fermat-Weber en norma 2. Éste fue pensado originalmente como un problema de minimización irrestricta. Sin embargo, para algunas aplicaciones reales tiene sentido restringir la ubicación del punto que se desea incorporar. Pensemos por ejemplo en la formulación de Weber, en la cual se busca ubicar un establecimiento de manera de minimizar los costos de transporte desde, o hacia, ciertos establecimientos dados, que pueden ser clientes o proveedores. Para este problema, podrían haber distintas razones que impidan ubicar el establecimiento en una cierta ubicación, por ejemplo:

- Si se busca ubicar una fábrica, podría haber regulaciones sanitarias que impidan su ubicación en una zona poblada, o por el contrario, decisiones políticas que exijan que se ubique en una cierta población para dar nuevas fuentes de trabajo.
- Si se busca ubicar un hospital, podría ser necesario ubicarlo en una zona tranquila, con bajos niveles de contaminación.
- Las características geográficas de la zona, como ríos o montañas, podrían ser barreras naturales para la ubicación del establecimiento.

Dependiendo de las causas de cada restricción, éstas se modelarán de distintas maneras. A lo largo de esta sección, presentaremos una modificación del algoritmo de Weiszfeld propuesta por Pilotta y Torres [PT11] para resolver el problema de Fermat-Weber con restricciones de tipo caja, es decir, para el caso en que cada variable está acotada. Este problema puede formalizarse del siguiente modo:

$$\text{Minimizar } f(x) = \sum_{i=1}^m w_i \|x - a_i\| \quad \text{sujeto a } \quad l \leq x \leq u \quad (4.3.1)$$

donde $l = (l_1, \dots, l_n)$ y $u = (u_1, \dots, u_n)$ satisfacen $-\infty \leq l_j < u_j \leq +\infty$ para todo $j = 1, \dots, n$, y donde $x \leq y$ denota que $x_j \leq y_j$ para cada coordenada $j = 1, \dots, n$. Esta formulación es útil para representar las situaciones en las que, por razones ambientales, políticas o económicas, hay una región ya definida para ubicar el nuevo establecimiento. De aquí en más, denotaremos con Ω al conjunto de ubicaciones permitidas, es decir, $\Omega = \{x \in \mathbb{R}^n : l \leq x \leq u\}$.

Para resolver este problema, Pilotta y Torres propusieron una transformación $\tilde{T} : \mathbb{R}^n - \mathcal{A} \rightarrow \Omega$, definida como $\tilde{T}(x) = P(T(x))$, donde T es la transformación de Weiszfeld y $P : \mathbb{R}^n \rightarrow \Omega$ es la proyección de \mathbb{R}^n a la caja Ω , la cual está dada por

$$[P(x)]_j = \begin{cases} l_j & \text{si } x_j < l_j \\ x_j & \text{si } l_j \leq x_j \leq u_j \\ u_j & \text{si } u_j < x_j \end{cases} \quad \text{para } j = 1, \dots, n$$

Como la transformación T es continua en $\mathbb{R}^n - \mathcal{A}$ y P es continua en \mathbb{R}^n , entonces la transformación \tilde{T} resulta continua. El algoritmo de Weiszfeld para el problema con restricciones de tipo caja es entonces:

Algoritmo 3: Algoritmo de Weiszfeld proyectado

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$; $-\infty \leq l_j < u_j \leq +\infty$, $j = 1, \dots, n$

Resultado: $x^* \in \mathbb{R}^n$ que minimiza $f(x)$

Elegir $x^{(0)} \neq a_1, \dots, a_m$;

para $k = 0, 1, 2, \dots$ **hacer**

$y^{(k+1)} \leftarrow T(x^{(k)})$;

$x^{(k+1)} \leftarrow P(y^{(k+1)})$;

si se satisface algún criterio de parada entonces

$\perp x^* \leftarrow x^{(k+1)}$

El criterio de parada puede variar según la precisión que se busque. En el Capítulo 6 presentamos los criterios que utilizamos en nuestro análisis.

A continuación presentaremos algunos resultados necesarios para probar la convergencia del algoritmo a la solución del problema (4.3.1). Al igual que para el caso irrestricto, la demostración se basa en hallar una función auxiliar que relacione $f(x)$ con $f(\tilde{T}(x))$ para probar la monotonicidad de f respecto a \tilde{T} y concluir que el algoritmo genera una sucesión convergente. Recordemos que en el caso irrestricto, se utilizó la función

$$h_x(y) = \sum_{i=1}^m w_i \frac{\|y - a_i\|^2}{\|x - a_i\|}$$

la cual satisface $2f(T(x)) - f(x) \leq h_x(T(x)) < h_x(x) = f(x)$. Para el caso con restricciones, encontrar esta función auxiliar no es sencillo, y deberemos antes presentar varias definiciones.

Dado $x \notin \mathcal{A}$, definimos

$$\alpha_j(x) = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (l_j - a_{ij}) \quad j = 1, \dots, n$$

$$\text{y } \beta_j(x) = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (a_{ij} - u_j) \quad j = 1, \dots, n$$

y construimos los siguientes conjuntos de índices:

$$L(x) = \{j \in \{1, \dots, n\} : T_j(x) < l_j\},$$

$$U(x) = \{j \in \{1, \dots, n\} : T_j(x) > u_j\},$$

$$I(x) = \{j \in \{1, \dots, n\} : T_j(x) \in [l_j, u_j]\}.$$

Con estas definiciones, para cada $x \notin \mathcal{A}$ y cada $j \in \{1, \dots, n\}$, se cumple que

- I. $j \in L(x)$ si y sólo si $T_j(x) < l_j$, si y sólo si $\alpha_j > 0$.
- II. $j \in U(x)$ si y sólo si $T_j(x) > u_j$, si y sólo si $\beta_j > 0$.
- III. $j \in I(x)$ si y sólo si $T_j(x) \in [l_j, u_j]$, si y sólo si $\alpha_j \leq 0$ y $\beta_j \leq 0$.

Además definiremos algunas funciones auxiliares, relacionadas fuertemente con la función proyección. Dado $x \in \mathbb{R}^n - \mathcal{A}$ fijo, definimos $E_x : \mathbb{R}^n \rightarrow \mathbb{R}^n$ y $g_x : \mathbb{R}^n - \mathcal{A} \rightarrow \mathbb{R}$ por

$$[E_x(y)]_j = \begin{cases} l_j & \text{si } j \in L(x) \\ y_j & \text{si } j \in I(x) \\ u_j & \text{si } j \in U(x) \end{cases} \quad \text{para } j = 1, \dots, n \quad (4.3.2)$$

$$y \quad g_x(y) = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \|E_x(y) - a_i\|^2. \quad (4.3.3)$$

Además, si $I(x) = \{j_1, \dots, j_r\} \neq \emptyset$, definimos $I_x : \mathbb{R}^r \rightarrow \mathbb{R}^n$ y $P_x : \mathbb{R}^n \rightarrow \mathbb{R}^r$ por

$$[I_x(z)]_j = \begin{cases} l_j & \text{si } j \in L(x) \\ z_s & \text{si } j = j_s \in I(x) \\ u_j & \text{si } j \in U(x) \end{cases} \quad \text{para } j = 1, \dots, n \quad (4.3.4)$$

$$y \quad [P_x(y)]_s = y_{j_s} \quad \text{para } s = 1, \dots, r. \quad (4.3.5)$$

Lema 4.3.1. Si $x \in \mathbb{R}^n - \mathcal{A}$ y $I(x) = \{j_1, \dots, j_r\} \neq \emptyset$, entonces $E_x \circ I_x = I_x$.

Demostración. Sean $z \in \mathbb{R}^r$ y $j \in \{1, \dots, n\}$. Si $j \in I(x)$, entonces $[E_x(I_x(z))]_j = [I_x(z)]_j$ por (4.3.2). Si $j \in L(x)$, $[E_x(I_x(z))]_j = l_j = [I_x(z)]_j$ por (4.3.2) y (4.3.4), y análogamente, si $j \in U(x)$, $[E_x(I_x(z))]_j = u_j = [I_x(z)]_j$. Luego, $E_x \circ I_x(z) = I_x(z)$ para todo $x \in \mathbb{R}^n$. \square

Ahora veremos la relación entre $f(x)$ y $g_x(x)$.

Lema 4.3.2. Si $x \in \mathbb{R}^n - \mathcal{A}$, entonces

$$g_x(x) = f(x) + 2 \sum_{j \in L(x)} (l_j - x_j) \alpha_j(x) - 2 \sum_{j \in U(x)} (u_j - x_j) \beta_j(x) + \gamma \quad (4.3.6)$$

con

$$\gamma = - \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left[\sum_{j \in L(x)} (l_j - x_j)^2 + \sum_{j \in U(x)} (u_j - x_j)^2 \right].$$

Demostración. Para cualquier $A, B, C \in \mathbb{R}$ se cumple que

$$(A - B)^2 = ((A - C) + (C - B))^2 = (A - C)^2 + (C - B)^2 + 2(A - C)(C - B)$$

por lo que

$$(A - C)^2 = (A - B)^2 - (C - B)^2 - 2(A - C)(C - B) = (A - B)^2 - (C - B)^2 + 2(C - A)(C - B).$$

Luego, para cada $i = 1, \dots, m$, si $j \in L(x)$, podemos tomar $A = a_{ij}$, $B = x_j$ y $C = l_j$ y obtenemos

$$(l_j - a_{ij})^2 = (x_j - a_{ij})^2 - (l_j - x_j)^2 + 2(l_j - x_j)(l_j - a_{ij}). \quad (4.3.7)$$

Análogamente, para $j \in U(x)$ podemos tomar $A = a_{ij}$, $B = x_j$ y $C = u_j$ y obtenemos

$$(u_j - a_{ij})^2 = (x_j - a_{ij})^2 - (u_j - x_j)^2 - 2(u_j - x_j)(a_{ij} - u_j) \quad (4.3.8)$$

Por la definición de E_x , tenemos que

$$g_x(x) = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left[\sum_{j \in L(x)} (l_j - a_{ij})^2 + \sum_{j \in I(x)} (x_j - a_{ij})^2 + \sum_{j \in U(x)} (u_j - a_{ij})^2 \right].$$

Aplicando las igualdades (4.3.7) y (4.3.8), obtenemos

$$\begin{aligned} g_x(x) &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left[\sum_{j=1}^n (x_j - a_{ij})^2 - \sum_{j \in L(x)} (l_j - x_j)^2 - \sum_{j \in U(x)} (u_j - x_j)^2 + \right. \\ &\quad \left. 2 \sum_{j \in L(x)} (l_j - x_j)(l_j - a_{ij}) - 2 \sum_{j \in U(x)} (u_j - x_j)(a_{ij} - u_j) \right] \\ &= f(x) - \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left(\sum_{j \in L(x)} (l_j - x_j)^2 + \sum_{j \in U(x)} (u_j - x_j)^2 \right) + \\ &\quad 2 \sum_{j \in L(x)} (l_j - x_j) \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (l_j - a_{ij}) - 2 \sum_{j \in U(x)} (u_j - x_j) \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (a_{ij} - u_j) \\ &= f(x) - \gamma + 2 \sum_{j \in L(x)} (l_j - x_j) \alpha_j(x) - 2 \sum_{j \in U(x)} (u_j - x_j) \beta_j(x). \end{aligned} \quad (4.3.9)$$

□

Definimos ahora una nueva función auxiliar, para $x \notin \mathcal{A}$ con $I(x) = \{i_1, \dots, i_r\} \neq \emptyset$, por $\tilde{g}_x = g_x \circ I_x : \mathbb{R}^r \rightarrow \mathbb{R}^n$.

Lema 4.3.3. *La función \tilde{g}_x es estrictamente convexa y tiene un único minimizador global en $z^* = (P_x \circ P \circ T)(x)$.*

Demostración.

$$\begin{aligned} \tilde{g}_x(z) &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \|E_x(I_x(z)) - a_i\|^2 \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \|I_x(z) - a_i\|^2 \text{ pues } E_x \circ I_x = I_x \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left[\sum_{j \in L(x)} (l_j - a_{ij})^2 + \sum_{j \in U(x)} (u_j - a_{ij})^2 + \sum_{s=1}^r (z_s - a_{i(j_s)})^2 \right]. \end{aligned}$$

Entonces

$$\nabla \tilde{g}_x(z) = 2 \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \begin{pmatrix} z_1 - a_{i(j_1)} \\ \vdots \\ z_r - a_{i(j_r)} \end{pmatrix} \quad (4.3.10)$$

y

$$\nabla^2 \tilde{g}_x(z) = 2 \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} Id_r$$

donde Id_r representa la matriz identidad en $\mathbb{R}^r \times \mathbb{R}^r$. Por lo tanto $\nabla^2 \tilde{g}_x(z)$ es una matriz simétrica y definida positiva para todo $z \in \mathbb{R}^r$ y \tilde{g}_x resulta estrictamente convexa.

Como es estrictamente convexa, la condición de optimalidad de primer orden es condición necesaria y suficiente para hallar un minimizador, es decir, basta con hallar z^* que satisfaga $\nabla \tilde{g}_x(z) = 0$. Esto es equivalente a resolver

$$\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (z_s^* - a_{i(j_s)}) = 0 \text{ para todo } s = 1, \dots, r$$

lo cual se satisface si y sólo si

$$z_s^* = \frac{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} a_{i(j_s)}}{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|}} \text{ para todo } s = 1, \dots, r \quad (4.3.11)$$

es decir, si y sólo si $z^* = (P_x \circ P \circ T)(x)$. Por ser estrictamente convexa, este minimizador es único y global. \square

Definimos \tilde{g}_x en términos de g_x , y encontramos su minimizador. Para encontrar una relación entre este minimizador y g_x , escribiremos g_x en términos de \tilde{g}_x .

Lema 4.3.4. $g_x(y) = (\tilde{g}_x \circ P_x)(y)$ para todo $y \in \mathbb{R}^n$.

Demostración. Dado $y \in \mathbb{R}^n$, por las definiciones (4.3.4) y (4.3.5), tenemos que

$$[(I_x \circ P_x)(y)]_j = \begin{cases} l_j & \text{si } j \in L(x) \\ [P_x(y)]_{j_s} = y_{j_s} & \text{si } j = j_s \in I(x) \\ u_j & \text{si } j \in U(x) \end{cases} \text{ para } j = 1, \dots, n$$

es decir que $(I_x \circ P_x)(y) = E_x(y)$. Además, sabemos que $E_x \circ I_x = I_x$ por lo que

$$(I_x \circ P_x)(y) = E_x((I_x \circ P_x)(y)).$$

Luego

$$\begin{aligned} g_x(x) &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \|E_x(y) - a_i\|^2 \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \|(I_x \circ P_x)(y) - a_i\|^2 \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \|E_x((I_x \circ P_x)(y)) - a_i\|^2 \\ &= g_x((I_x \circ P_x)(y)) \\ &= (\tilde{g}_x \circ P_x)(y). \end{aligned}$$

\square

Los siguientes resultados darán las últimas propiedades de g_x que precisaremos para probar la monotonicidad de f respecto a \tilde{T} .

Lema 4.3.5. Si $x \notin \mathcal{A}$, entonces $g_x(\tilde{T}(x)) \leq g_x(x)$. Más aún, si $I(x) \neq \emptyset$ y $P_x(\tilde{T}(x)) \neq P_x(x)$, la desigualdad es estricta.

Demostración. Si $I(x) = \emptyset$, entonces $g_x(y)$ no depende de y y por lo tanto $g_x(\tilde{T}(x)) = g_x(x)$. Si $I(x) \neq \emptyset$ y $P_x(\tilde{T}(x)) = P_x(x)$, entonces $g_x(\tilde{T}(x)) = (\tilde{g}_x \circ P_x)(P \circ T(x)) = (\tilde{g}_x \circ P_x)(x) = g_x(x)$. Por último, si $I(x) \neq \emptyset$ y $P_x(\tilde{T}(x)) \neq P_x(x)$, como $P_x(\tilde{T}(x)) = (P_x \circ P \circ T)(x)$ es el único minimizador de \tilde{g}_x , tenemos que $g_x(\tilde{T}(x)) = (\tilde{g}_x \circ P_x)(P \circ T(x)) < (\tilde{g}_x \circ P_x)(x) = g_x(x)$. \square

Lema 4.3.6. Si $x \in \Omega - \mathcal{A}$ es tal que $x \neq \tilde{T}(x)$, entonces $g_x(\tilde{T}(x)) < f(x)$.

Demostración. Si $I(x) = \emptyset$, entonces $L(x) \neq \emptyset$ o $U(x) \neq \emptyset$. Como $x \in \Omega$, tenemos que $l_j - x_j \leq 0$ y $u_j - x_j \geq 0$ para todo $j = 1, \dots, n$. Como $I(x) = \emptyset$, entonces

$$\tilde{T}_j(x) = \begin{cases} l_j & \text{si } j \in L(x) \\ u_j & \text{si } j \in U(x). \end{cases}$$

Luego, $x \neq \tilde{T}(x)$ implica que existe al menos un $j \in L(x) \cup U(x)$ tal que $l_j - x_j < 0$ o $u_j - x_j > 0$, por lo que usando la ecuación (4.3.6) obtenemos

$$g_x(\tilde{T}(x)) \leq g_x(x) = f(x) + 2 \sum_{j \in L(x)} (l_j - x_j) \alpha_j(x) - 2 \sum_{j \in U(x)} (u_j - x_j) \beta_j(x) + \gamma < f(x).$$

Para el caso en que $I(x) \neq \emptyset$ y $P_x(\tilde{T}(x)) = P_x(x)$ se da un razonamiento análogo, usando que si $P_x(\tilde{T}(x)) = P_x(x)$, entonces $\tilde{T}_j(x) = x_j$ para todo $j \in I(x)$, por lo cual $x \neq \tilde{T}(x)$ significa que $\tilde{T}_j(x) \neq x_j$ para algún $j \in L(x) \cup U(x)$.

Por último, si $I(x) \neq \emptyset$ y $P_x(\tilde{T}(x)) \neq P_x(x)$, usando que $l_j - x_j \leq 0 \leq u_j - x_j$ y aplicando el Lema 4.3.5 obtenemos

$$g_x(\tilde{T}(x)) < g_x(x) = f(x) + 2 \sum_{j \in L(x)} (l_j - x_j) \alpha_j(x) - 2 \sum_{j \in U(x)} (u_j - x_j) \beta_j(x) + \gamma \leq f(x).$$

□

El último lema que precisaremos para ver el decrecimiento de $f(x^{(k)})$ relaciona g_x con f evaluada en x y $\tilde{T}(x)$.

Lema 4.3.7. Si $x \in \Omega - \mathcal{A}$ es tal que $x \neq \tilde{T}(x)$, entonces

$$g_x(\tilde{T}(x)) \geq 2f(\tilde{T}(x)) - f(x). \quad (4.3.12)$$

Demostración. En primer lugar, notemos que

$$\begin{aligned} g_x(\tilde{T}(x)) &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left\| E_x(\tilde{T}(x)) - a_i \right\|^2 \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left[\|x - a_i\| + \left\| E_x(\tilde{T}(x)) - a_i \right\| - \|x - a_i\| \right]^2 \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left[\|x - a_i\|^2 + 2\|x - a_i\| \left(\left\| E_x(\tilde{T}(x)) - a_i \right\| - \|x - a_i\| \right) + \right. \\ &\quad \left. \left(\left\| E_x(\tilde{T}(x)) - a_i \right\| - \|x - a_i\| \right)^2 \right] \\ &= \sum_{i=1}^m w_i \|x - a_i\| + 2 \sum_{i=1}^m w_i \left(\left\| E_x(\tilde{T}(x)) - a_i \right\| - \|x - a_i\| \right) + \\ &\quad \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \left(\left\| E_x(\tilde{T}(x)) - a_i \right\| - \|x - a_i\| \right)^2, \end{aligned}$$

por lo cual

$$g_x(\tilde{T}(x)) \geq f(x) + 2[f(E_x(\tilde{T}(x))) - f(x)] = 2f(E_x(\tilde{T}(x))) - f(x).$$

Ahora, si $j \in L(x)$, entonces $T_j(x) < l_j$ y por lo tanto $\tilde{T}_j(x) = [P(\tilde{T}(x))]_j = l_j = [E_x(\tilde{T}(x))]_j$. De manera análoga puede verse que si $j \in U(x)$, $\tilde{T}_j(x) = T_j(x) = [E_x(\tilde{T}(x))]_j$, y si $j \in I(x)$, $\tilde{T}_j(x) = u_j = [E_x(\tilde{T}(x))]_j$. Entonces

$$E_x(\tilde{T}(x)) = \tilde{T}(x),$$

lo cual concluye la demostración del lema. □

Estos resultados permiten demostrar con facilidad el siguiente teorema:

Teorema 4.3.8. (*Monotonicidad de f respecto al operador \tilde{T}*)
Si $x \in \Omega - \mathcal{A}$ y $\tilde{T}(x) \neq x$, entonces

$$f(\tilde{T}(x)) < f(x).$$

Demostración. Por el Lema 4.3.6, sabemos que $g_x(\tilde{T}(x)) < f(x)$, y por el Lema 4.3.7,

$$2f(\tilde{T}(x)) - f(x) \leq g_x(\tilde{T}(x)).$$

Luego, obtenemos que

$$2f(\tilde{T}(x)) - f(x) < f(x)$$

y por lo tanto

$$f(\tilde{T}(x)) < f(x). \quad (4.3.13)$$

□

El Teorema 4.3.8 nos permite asegurar que si $x^{(0)} \in \Omega$ y $x^{(k)} \notin \mathcal{A}$ para todo $k \geq 0$, entonces la sucesión $\{f(x^{(k)})\}$ es estrictamente decreciente.

Como Ω es compacto, si $x^{(k)}$ está definido para todo k , entonces la sucesión $\{x^{(k)}\}$ tiene al menos una subsucesión que converge a algún $\tilde{x} \in \Omega$. Veamos que si la sucesión converge, lo hace a la solución del problema. Primero, veamos que el límite, si existe, es punto fijo de \tilde{T} .

Teorema 4.3.9. *Si la sucesión $\{x^{(k)}\}$ converge a $\tilde{x} \in \Omega$, entonces \tilde{x} es punto fijo de \tilde{T} .*

Demostración. Es evidente por continuidad de \tilde{T} , pues

$$\tilde{x} = \lim_{j \rightarrow \infty} x^{(k+1)} = \lim_{j \rightarrow \infty} \tilde{T}(x^{(k)}) = \tilde{T}(\tilde{x}).$$

□

Con esto, si vemos que los puntos fijos de \tilde{T} son minimizadores de f en Ω , quedará probado que si la sucesión converge lo hace a la solución del problema (4.3.1).

Teorema 4.3.10. *x^* minimiza el problema (4.3.1) si y sólo si es punto fijo de la transformación \tilde{T} .*

Demostración. Si x^* es minimizador de f sobre Ω , entonces $f(x^*) < f(y)$ para todo $y \in \Omega$, $y \neq x^*$. Por el Teorema 4.3.8, si $x^* \neq \tilde{T}(x^*)$, entonces $f(\tilde{T}(x^*)) < f(x^*)$. Pero $\tilde{T}(x^*) \in \Omega$, lo cual lleva a un absurdo. Luego, debe ser $x^* = \tilde{T}(x^*)$.

Resta ver que si x^* es punto fijo de \tilde{T} , entonces x^* es solución de (4.3.1). Sabemos que x^* es minimizador de f sobre Ω si y sólo si satisface las condiciones de optimalidad KKT, las cuales para este problema están dadas por

$$\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (x - a_i) - \sum_{j=1}^n \mu_j e_j + \sum_{j=1}^n \mu_{j+n} e_j = 0 \quad (4.3.14)$$

$$\mu_j (l_j - x_j) = 0 \quad \text{para } j = 1, \dots, n \quad (4.3.15)$$

$$\mu_{j+n} (x_j - u_j) = 0 \quad \text{para } j = 1, \dots, n \quad (4.3.16)$$

$$\mu_j \geq 0 \quad \text{para } j = 1, \dots, 2n \quad (4.3.17)$$

$$l_j - x_j \leq 0 \quad \text{para } j = 1, \dots, n \quad (4.3.18)$$

$$x_j - u_j \leq 0 \quad \text{para } j = 1, \dots, n \quad (4.3.19)$$

donde $e_j \in \mathbb{R}^n$ es el vector cuya coordenada j es $e_{jj} = 1$ y todas sus demás coordenadas son nulas.

Supongamos que $x = \tilde{T}(x)$, y veamos que se satisfacen las ecuaciones (4.3.14) a (4.3.19).

Definimos los multiplicadores

$$\mu_j = \begin{cases} \alpha_j(x) & \text{si } j \in L(x) \\ 0 & \text{si } j \notin L(x) \end{cases} \quad \text{y} \quad \mu_{j+n} = \begin{cases} \beta_j(x) & \text{si } j \in U(x) \\ 0 & \text{si } j \notin U(x) \end{cases} \quad \text{para } j = 1, \dots, n. \quad (4.3.20)$$

Por construcción, los multiplicadores satisfacen la ecuación (4.3.17). Como $x = \tilde{T}(x) = (P \circ T)(x)$, entonces $x \in \Omega$ y se satisfacen las ecuaciones (4.3.18) y (4.3.19).

Sea $j \in \{1, \dots, n\}$.

- Si $j \notin L(x)$, entonces $\mu_j = 0$.
- Si $j \in L(x)$, entonces $T_j(x) < l_j$ y por lo tanto $l_j - x_j = l_j - [P(T(x))]_j = l_j - l_j = 0$.

Luego, se satisface (4.3.15).

Análogamente, dado $j \in \{1, \dots, n\}$, tenemos que:

- Si $j \notin U(x)$, entonces $\mu_{j+n} = 0$.
- Si $j \in U(x)$, entonces $T_j(x) > u_j$ y por lo tanto $x_j - u_j = [P(T(x))]_j - u_j = u_j - u_j = 0$.

Luego, se satisface (4.3.16).

Para ver (4.3.14), analicemos la igualdad coordenada a coordenada, es decir

$$\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (x_j - a_{ij}) - \mu_j + \mu_{j+n} = 0. \quad (4.3.21)$$

- Para $j \in I(x)$, tenemos que $T_j(x) \in [l_j, u_j]$ por lo que $x = \tilde{T}(x) = P_j(T(x)) = T_j(x)$. Además, $\mu_j = \mu_{j+n} = 0$. Luego

$$\begin{aligned} \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (x_j - a_{ij}) - \mu_j + \mu_{j+n} &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (x_j - a_{ij}) \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} x_j - \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} a_{ij} \\ &= \left[\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right] \left(x_j - \frac{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} a_{ij}}{\sum_{i=1}^m \frac{w_i}{\|x - a_i\|}} \right) \\ &= \left[\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} \right] (x_j - T_j(x)) \\ &= 0. \end{aligned}$$

- Para $j \in L(x)$ tenemos que $T_j(x) < l_j$ por lo que $x = \tilde{T}(x) = l_j$. Además, $\mu_j = \alpha_j(x)$ y $\mu_{j+n} = 0$, por lo que obtenemos

$$\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (x_j - a_{ij}) - \mu_j + \mu_{j+n} = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (l_j - a_{ij}) - \alpha_j(x) = 0.$$

- Para $j \in U(x)$, $T_j(x) > u_j$ por lo que $x = \tilde{T}(x) = u_j$. Además, $\mu_j = 0$ y $\mu_{j+n} = \beta_j(x)$, por lo que obtenemos

$$\sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (x_j - a_{ij}) - \mu_j + \mu_{j+n} = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|} (u_j - a_{ij}) + \beta_j(x) = 0.$$

Por lo tanto para todo j se satisface (4.3.21), lo cual significa que vale (4.3.14). Es decir que si $x = \tilde{T}(x)$, entonces x satisface las condiciones KKT y por lo tanto x es la solución x^* de (4.3.1). \square

Capítulo 5

El método de Newton para el problema de Fermat-Weber

5.1. Desarrollo y convergencia del método

Consideremos nuevamente el problema de Fermat-Weber irrestricto en norma 2. Vimos en la Sección 3.2 que el algoritmo de Weiszfeld tiene tasa de convergencia lineal en la mayoría de los casos. En la Sección 4.2 mostramos que puede modificarse el algoritmo original, tomando

$$x^{(k+1)} = x^{(k)} - \frac{\lambda}{L(x)} \nabla f(x)$$

con $\lambda > 0$, pero aseguramos la convergencia sólo para $\lambda \leq \frac{2}{n-1}$, por lo que para $n \neq 2$ no encontramos una longitud mayor del paso, es decir que la variación sólo acelera el algoritmo original para $n = 2$.

En general, cuando se busca un método que converja rápidamente, se tiende a recurrir al método de Newton. Sin embargo, no hay muchos autores que hayan aplicado este método al problema de Fermat-Weber, posiblemente porque la función f no es diferenciable en los puntos a_1, \dots, a_m . Sin embargo, hemos visto en la Sección 3.3 que hay maneras de evitar los nodos, como por ejemplo, el método SP propuesto por Beck y Sabach [BS15]. Con esto en mente, Görner y Kanzow [GK16] probaron en 2016 la convergencia del método de Newton para resolver el problema de Fermat-Weber, partiendo de una elección adecuada del punto inicial. A lo largo de esta sección, presentaremos el algoritmo de Görner y Kanzow y la demostración de su convergencia.

De aquí en más, supondremos que ningún nodo es solución y que los nodos no son colineales, pues en ambos casos sería innecesario implementar el método de Newton.

Consideremos el punto inicial $x^{(0)}$ elegido por el método SP (Algoritmo 2). El conjunto de nivel

$$\mathcal{L}(x^{(0)}) = \left\{ x \in \mathbb{R}^n : f(x) \leq f(x^{(0)}) \right\}$$

es convexo, compacto, contiene a la solución x^* del problema de Fermat-Weber y no contiene ningún nodo. Luego, la función f resulta continuamente diferenciable en $\mathcal{L}(x^{(0)})$ y podemos aplicar el método de Newton, incorporando una búsqueda lineal de tipo Armijo para garantizar que $f(x^{(k)})$

sea no creciente y $x^{(k)} \in \mathcal{L}(x^{(0)})$ para todo $k \in \mathbb{N}$. El algoritmo propuesto en [GK16] es el siguiente:

Algoritmo 4: Método de Newton con condición de Armijo

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$; $\rho \in (0, 1)$ y $\sigma \in (0, \frac{1}{2})$

Resultado: $x^* \in \mathbb{R}^n$ que minimiza $f(x)$

Elegir $x^{(0)}$ mediante el método SP;

para $k = 0, 1, 2, \dots$ **hacer**

- 1 Calcular la dirección de Newton d_k resolviendo $H(x^{(k)})d = -\nabla f(x^{(k)})$;
 - 2 Elegir el paso t_k como el mayor $t \in \{1, \rho, \rho^2, \dots\}$ tal que
 $f(x^{(k)} + td_k) \leq f(x^{(k)}) + \sigma t \nabla f(x^{(k)})^T d_k$;
 - 3 Tomar $x^{(k+1)} \leftarrow x^{(k)} + t_k d_k$;
si se satisface algún criterio de parada entonces
 $\perp x^* \leftarrow x^{(k+1)}$
-

donde $H(x)$ denota la hessiana de f en x , es decir, $H(x) = \nabla^2 f(x)$.

El criterio de parada puede variar según la precisión que se busque. En el Capítulo 6 presentamos los criterios que utilizamos en nuestro análisis.

Veremos que el método satisface las condiciones necesarias para converger a la solución. El primer resultado que veremos será que la hessiana $H(x)$ es definida positiva en $\mathcal{L}(x^{(0)})$, lo cual garantizará que los dos primeros pasos del algoritmo están bien definidos.

Teorema 5.1.1. *Para todo $x \in \mathcal{L}(x^{(0)})$, $H(x) = \nabla^2 f(x)$ es definida positiva.*

Demostración. Recordemos que

$$\frac{\partial f}{\partial x_j}(x) = \sum_{i=1}^m w_i \frac{x_j - a_{ij}}{\|x - a_i\|}, \quad j = 1, \dots, n.$$

Dados $k, j \in \{1, \dots, n\}$ tenemos que

$$\begin{aligned} \frac{\partial^2 f}{\partial x_k \partial x_j}(x) &= \frac{\partial}{\partial x_k} \sum_{i=1}^m w_i (x_j - a_{ij}) \|x - a_i\|^{-1} \\ &= \sum_{i=1}^m w_i \left[\delta_{jk} \|x - a_i\|^{-1} + (x_j - a_{ij}) \left(-\|x - a_i\|^{-2} \frac{x_k - a_{ik}}{\|x - a_i\|} \right) \right] \\ &= \sum_{i=1}^m \frac{w_i}{\|x - a_i\|^3} \left[\delta_{jk} \|x - a_i\|^2 - (x_j - a_{ij})(x_k - a_{ik}) \right] \end{aligned}$$

por lo que

$$H(x) = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|^3} \left[\|x - a_i\|^2 Id - (x - a_i)(x - a_i)^T \right]. \quad (5.1.1)$$

Sea $y \in \mathbb{R}^n$, $y \neq 0$.

$$y^T \left[\|x - a_i\|^2 Id - (x - a_i)(x - a_i)^T \right] y = \|x - a_i\|^2 \|y\|^2 - [y^T(x - a_i)]^2$$

Por la desigualdad de Cauchy-Schwartz,

$$[y^T(x - a_i)]^2 \leq \|x - a_i\|^2 \|y\|^2$$

y la igualdad vale si y sólo si $x - a_i = c_i y$ con $c_i \in \mathbb{R}$. Luego,

$$y^T \left[\|x - a_i\|^2 Id - (x - a_i)(x - a_i)^T \right] y \geq 0$$

y la igualdad vale si y sólo si $x - a_i = c_i y$. Esto nos asegura que

$$y^T H(x)y = \sum_{i=1}^m \frac{w_i}{\|x - a_i\|^3} y^T \left[\|x - a_i\|^2 Id - (x - a_i)(x - a_i)^T \right] y \geq 0$$

donde la igualdad vale si y sólo si $x - a_i = c_i y$ para todo $i = 1, \dots, m$. Pero eso implicaría que los nodos son colineales y por hipótesis no lo son. Luego, debe ser $y^T H(x)y > 0$ para todo $y \neq 0$ y por lo tanto $H(x)$ resulta definida positiva. \square

Como $\mathcal{L}(x^{(0)})$ es compacto, y los autovalores mínimo y máximo, $\lambda_{\min}(x)$ y $\lambda_{\max}(x)$ respectivamente, de $H(x)$ son funciones continuas y positivas de $x \in \mathcal{L}(x^{(0)})$, entonces queda probado el siguiente corolario:

Corolario 5.1.2. *Existen constantes $\beta \geq \alpha > 0$ tales que*

$$\alpha \|y\|^2 \leq y^T H(x)y \leq \beta \|y\|^2$$

para todo $y \in \mathbb{R}^n$ y todo $x \in \mathcal{L}(x^{(0)})$.

El siguiente lema nos asegura que el criterio de Armijo provee una búsqueda lineal eficiente para resolver este problema. Recordemos que $\rho \in (0, 1)$ y $\sigma \in (0, \frac{1}{2})$.

Lema 5.1.3. *Existe una constante $\theta > 0$ tal que*

$$f(x^{(k)} + t_k d_k) \leq f(x^{(k)}) - \theta \left(\frac{\nabla f(x^{(k)})^T d_k}{\|d_k\|} \right)^2 \quad \text{para todo } k \in \mathbb{N}.$$

Demostración. Como d_k es la dirección de Newton, entonces $H(x^{(k)})d_k = -\nabla f(x^{(k)})$. Por el Corolario 5.1.2

$$-\frac{\nabla f(x^{(k)})^T d_k}{\|d_k\|^2} = \frac{d_k^T H(x^{(k)})d_k}{\|d_k\|^2} \leq \beta \quad \text{para todo } k \in \mathbb{N}. \quad (5.1.2)$$

Consideremos ahora $k \in \mathbb{N}$ fijo. Sea $\varphi_k(t) := f(x^{(k)} + t d_k)$. Como f es continua y acotada inferiormente, existe $\hat{t}_k > 0$ tal que es el mínimo que satisface

$$\varphi'_k(\hat{t}_k) = \sigma \varphi'_k(0).$$

Además, $x^{(k)} + \hat{t}_k d_k \in \mathcal{L}(x^{(0)})$ pues \hat{t}_k minimiza a φ_k . Como ∇f es Lipschitz continua en $\mathcal{L}(x^{(0)})$ y $(x^{(k)} + t d_k) \in \mathcal{L}(x^{(0)})$ para todo $0 < t \leq \hat{t}_k$, entonces existe $L > 0$ tal que

$$\sigma \varphi'_k(0) = \varphi'_k(\hat{t}_k) = \varphi'_k(0) + (\varphi'_k(\hat{t}_k) - \varphi'_k(0)) \leq \varphi'_k(0) + \hat{t}_k L \|d_k\|^2.$$

Entonces

$$\hat{t}_k \geq -\frac{(1 - \sigma)\varphi'_k(0)}{L \|d_k\|^2}. \quad (5.1.3)$$

Ahora, para la condición de Armijo pueden darse dos situaciones: que $t_k = 1$, o que $t_k < 1$. Supongamos que en la condición de Armijo $t_k = 1$. Entonces por (5.1.2) tenemos que

$$-\frac{1}{\beta} \frac{\nabla f(x^{(k)})^T d_k}{\|d_k\|^2} \leq 1 = t_k$$

y por lo tanto

$$f(x^{(k)} + t_k d_k) = \varphi_k(t_k) \leq \varphi_k(0) + \sigma t_k \varphi'_k(0) \leq f(x^{(k)}) - \frac{\sigma}{\beta} \left[\frac{\nabla f(x^{(k)})^T d_k}{\|d_k\|} \right]^2$$

donde $\frac{\sigma}{\beta}$ no depende de k .

Si la condición de Armijo vale con $t_k < 1$, entonces $t_k = \rho^r$ con $r \in \mathbb{N}$ y $\rho^{r-1} = \frac{t_k}{\rho}$ no satisface la condición de Armijo. Como φ_k es convexa entonces

$$\frac{\varphi_k(\hat{t}_k) - \varphi_k(0)}{\hat{t}_k} \leq \varphi'_k(\hat{t}_k).$$

Luego, por elección de \hat{t}_k obtenemos que

$$\frac{\varphi_k(\hat{t}_k) - \varphi_k(0)}{\hat{t}_k} \leq \sigma \varphi'_k(0).$$

Esto nos dice que \hat{t}_k satisface la condición de Armijo, y por lo tanto se satisface la condición de Armijo para todo t tal que $0 < t \leq \hat{t}_k$. Luego, debe cumplirse que

$$\frac{t_k}{\rho} > \hat{t}_k. \quad (5.1.4)$$

Ahora,

$$f(x^{(k)} + t_k d_k) = \varphi_k(t_k) \leq \varphi_k(0) + \sigma t_k \varphi'_k(0).$$

Usando (5.1.4) obtenemos

$$f(x^{(k)} + t_k d_k) \leq \varphi_k(0) + \sigma \rho \hat{t}_k \varphi'_k(0)$$

e incorporando (5.1.3), resulta

$$f(x^{(k)} + t_k d_k) \leq \varphi_k(0) - \frac{\sigma \rho (1 - \sigma)}{L} \left[\frac{\nabla f(x^{(k)})^T d_k}{\|d_k\|} \right]^2$$

donde $\frac{\sigma \rho (1 - \sigma)}{L}$ tampoco depende de k .

Luego, si tomamos

$$\theta := \min \left\{ \frac{\sigma}{\beta}, \frac{\sigma \rho (1 - \sigma)}{L} \right\}$$

obtenemos que para todo $k \in \mathbb{N}$ se satisface que

$$f(x^{(k)} + t_k d_k) \leq f(x^{(k)}) - \theta \left(\frac{\nabla f(x^{(k)})^T d_k}{\|d_k\|} \right)^2.$$

□

Con estos resultados, puede demostrarse la convergencia del algoritmo de Newton para el problema de Fermat-Weber, la cual puede encontrarse en el trabajo de Görner y Kanzow [GK16].

Teorema 5.1.4. *Si ningún nodo es solución, entonces la sucesión generada por el Algoritmo 4 converge a la única solución x^* del problema de Fermat-Weber, y la tasa de convergencia resulta cuadrática, es decir, existen $c > 0$ y $K \in \mathbb{N}$ tales que $\|x^{(k+1)} - x^*\| \leq c \|x^{(k)} - x^*\|^2$ para todo $k \geq K$.*

El Teorema 5.1.4 asegura que el método de Görner y Kanzow converge a la solución, y garantiza que tiene velocidad de convergencia cuadrática. En la Sección 6.1, compararemos este método con el método de Weiszfeld, y con la aceleración del mismo propuesta en la Sección 4.2.

5.2. Newton aplicado al problema con restricciones de cotas en las variables

Consideremos nuevamente el problema de Fermat-Weber con restricciones de tipo caja,

$$\text{Minimizar } f(x) = \sum_{i=1}^m w_i \|x - a_i\| \quad \text{sujeto a } l \leq x \leq u. \quad (5.2.1)$$

Así como el método de Newton permitió generar un algoritmo más eficiente que el método de Weiszfeld para el problema irrestricto, se podría buscar este resultado para el problema con restricciones. En esta sección, presentaremos los resultados obtenidos al construir un método de Newton proyectado, buscando generar un algoritmo más eficiente que el método de Weiszfeld proyectado de Pilotta y Torres [PT11]. Con el método de Newton de Görner y Kanzow [GK16] en mente, buscamos una forma de incorporar la proyección, y para ello estudiamos distintos métodos de minimización en cajas, como el método del gradiente espectral de Birgin, Martínez y Raydan [BMR00] y [BMR14] y la proyección de un paso quasi-Newton propuesta por Kim, Sra y Dhillon [KSD10], y reproducimos algunas de sus ideas aplicadas al método de Newton.

Siguiendo estas ideas, elaboramos dos propuestas, similares entre sí, que consistían en calcular la dirección de Newton, proyectar y realizar una búsqueda lineal para hallar un paso adecuado. La diferencia entre ambas propuestas es esencialmente en qué instancia se realizó la búsqueda lineal.

La primer propuesta fue tomar en cada iteración la dirección $d_k = P(x^{(k)} + d_k^{Newt}) - x^{(k)}$, donde d_k^{Newt} es la dirección dada por el método de Newton partiendo de $x^{(k)}$, y realizar una búsqueda lineal en esa dirección, dando lugar al siguiente método:

Algoritmo 5: Algoritmo de Newton proyectado - versión 1

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$, $\rho \in (0, 1)$, $\sigma \in (0, \frac{1}{2})$, $\delta > 0$

Resultado: $x^* \in \mathbb{R}^n$

Elegir $x^{(0)} \neq a_1, \dots, a_m$ tal que $x^{(0)} \in \Omega$;

para $k = 0, 1, 2, \dots$ **hacer**

Calcular la dirección de Newton d_k^{Newt} ;

$d_k = P(x^{(k)} + d_k^{Newt}) - x^{(k)}$;

$r = \text{mín}\{r \in \mathbb{N}_0 : f(x^{(k)} + \rho^r d_k) \leq f(x^{(k)}) + \sigma \rho^r \nabla f(x^{(k)})^T d_k\}$;

$\lambda_k = \rho^r$;

$x^{(k+1)} = x^{(k)} + \lambda_k d_k$;

si $\|x^{(k+1)} - x^{(k)}\| < \delta$ **entonces**

└ $x^* = x^{(k+1)}$

La siguiente propuesta fue tomar $x^{(k+1)} = P(x^{(k)} + \lambda_k d_k^{Newt})$, con λ_k elegido del siguiente modo:

Algoritmo 6: Algoritmo de Newton proyectado - versión 2

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$, $\rho \in (0, 1)$, $\sigma \in (0, \frac{1}{2})$, $\delta > 0$

Resultado: $x^* \in \mathbb{R}^n$

Elegir $x^{(0)} \neq a_1, \dots, a_m$ tal que $x^{(0)} \in \Omega$;

para $k = 0, 1, 2, \dots$ **hacer**

Calcular la dirección de Newton d_k^{Newt} ;

$r = \text{mín}\{r \in \mathbb{N}_0 : f(P(x^{(k)} + \rho^r d_k^{Newt})) \leq f(x^{(k)}) + \sigma \rho^r \nabla f(x^{(k)})^T d_k^{Newt}\}$;

$\lambda_k = \rho^r$;

$x^{(k+1)} = P(x^{(k)} + \lambda_k d_k^{Newt})$;

si $\|x^{(k+1)} - x^{(k)}\| < \delta$ **entonces**

└ $x^* = x^{(k+1)}$

De aquí nos referiremos con (NP1) a la primera versión del método de Newton proyectado, y con (NP2) a la segunda versión.

Tras aplicar en varios ejemplos ambos algoritmos, vimos que cuando converge a la solución, lo hace en pocas iteraciones y mejora los resultados del algoritmo de Weiszfeld proyectado. Sin embargo, en otras ocasiones los pasos generados por el algoritmo son muy pequeños, por lo cual no llega a la solución. Ambas situaciones serán presentadas en la Sección 6.2.

La conclusión a la que llegamos fue que el problema que enfrenta este algoritmo es que toma direcciones casi perpendiculares al gradiente, por lo que el descenso es muy pequeño y esto lo obliga a dar pasos cada vez más cortos. A continuación, presentamos un breve ejemplo para ilustrar esto, el cual mostraremos más claramente en 6.2.

Ejemplo 5.2.1. Consideremos el problema de Fermat-Weber en \mathbb{R}^2 con 4 nodos, $a_1 = (1, 0)$, $a_2 = (0, 0)$, $a_3 = (0, 1)$ y $a_4 = (1, 4)$, y pesos $w_1 = 5$, $w_2 = 3$, $w_3 = 2$ y $w_4 = 3$, e incorporemos la caja $[0, 1] \times [1.5, 3.5]$. En la figura 5.1, puede verse un gráfico aproximado de las curvas de nivel de la función f , así como también la frontera de la caja. En la figura 5.2, pueden verse la solución al problema restringido a la caja y el punto en el cual se detuvo el método (NP1). El método se detuvo por realizar pasos muy pequeños, sin llegar a la solución. Lo mismo ocurrió con el método (NP2), el cual se detuvo más lejos de la solución. Para los tres algoritmos, el criterio de parada fue que la distancia entre una iteración y la siguiente fuera menor a 10^{-5} . Los resultados fueron:

- El algoritmo de Weiszfeld proyectado encontró la solución $x^* = (0.47292, 1.5)$ y el valor de f es $f(x^*) = 21.709$.
- El algoritmo (NP1) se detuvo en $x^{(NP1)} = (0.41493, 1.5)$, con $f(x^{(NP1)}) = 21.721$.
- El algoritmo (NP2) se detuvo en $x^{(NP2)} = (0.248, 1.5)$, con $f(x^{(NP2)}) = 21.899$.

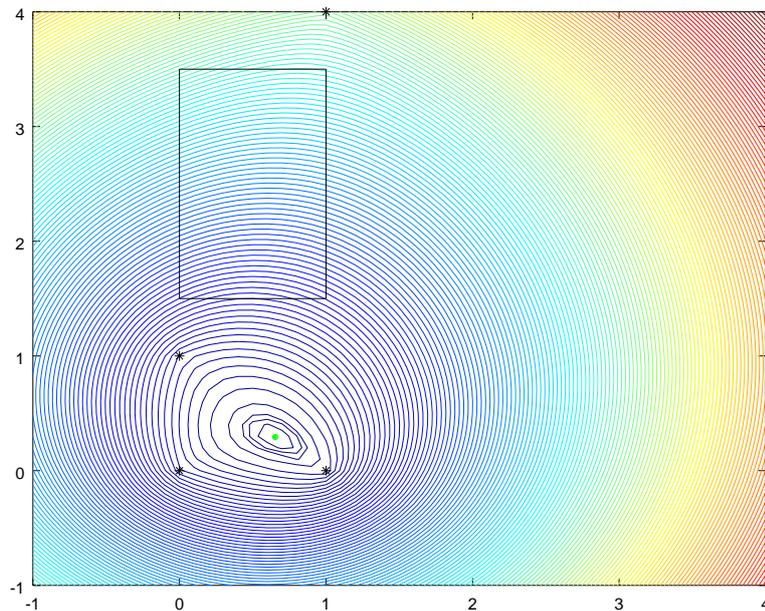


Figura 5.1: Curvas de nivel de la función dada en el ejemplo 5.2.1

Uno de los principales problemas de que el algoritmo se detenga por la longitud de paso es que no permite saber cuán lejos de la solución se detuvo. Para resolver este problema, desarrollamos un algoritmo híbrido que combina las iteraciones de (NP1) con las del método de Weiszfeld proyectado. Este nuevo método realiza las iteraciones de Newton mientras sean “buenas”, y si son “malas” realiza una iteración de Weiszfeld. El criterio usado para determinar si una iteración era buena o mala fue la longitud del paso. Para garantizar que no se interprete como mala una iteración que lleve a la solución, se incorporó el criterio del gradiente proyectado: si un punto x^* es solución del

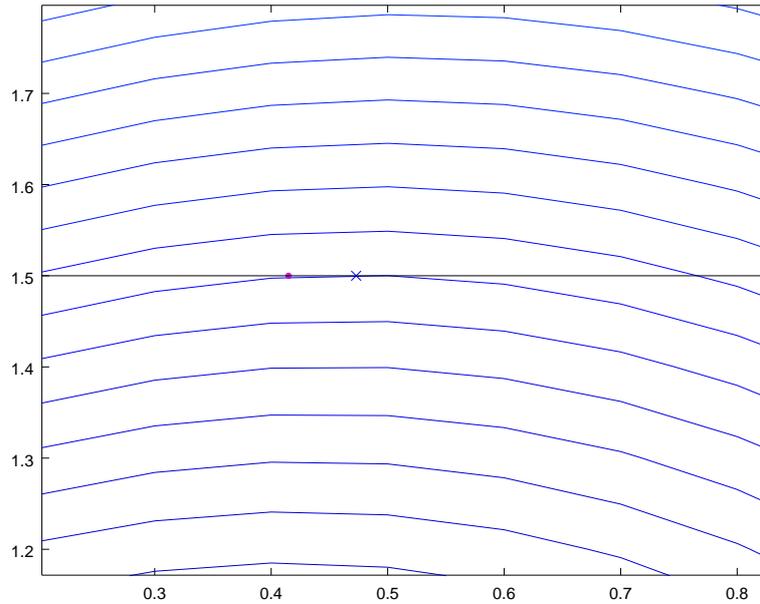


Figura 5.2: La cruz indica la solución del problema dado en el ejemplo 5.2.1, y el punto indica el resultado de la primer versión de Newton proyectado.

problema restringido a una caja Ω , entonces $g_P(x^*) = 0$, donde g_P es el gradiente proyectado de f , definido por

$$[g_P(x)]_j = \begin{cases} 0 & \text{si } x_j = l_j \text{ y } \frac{\partial f}{\partial x_j}(x) > 0 \\ 0 & \text{si } x_j = u_j \text{ y } \frac{\partial f}{\partial x_j}(x) < 0 \\ \frac{\partial f}{\partial x_j}(x) > 0 & \text{caso contrario.} \end{cases} \quad (5.2.2)$$

El algoritmo híbrido es entonces el siguiente:

Algoritmo 7: Algoritmo híbrido

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$, $\rho \in (0, 1)$, $\sigma \in (0, \frac{1}{2})$, $\varepsilon, \delta > 0$

Resultado: $x^* \in \mathbb{R}^n$ que minimiza $f(x)$

Elegir $x^{(0)} \neq a_1, \dots, a_m$ tal que $x^{(0)} \in \Omega$;

para $k = 0, 1, 2, \dots$ **hacer**

- 1 Calcular la dirección de Newton d_k^{Newt} ;
 $d_k = P(x^{(k)} + d) - x^{(k)}$;
 $r = \min\{r \in \mathbb{N}_0 : f(x^{(k)} + \rho^r d_k) \leq f(x^{(k)}) + \sigma \rho^r \nabla f(x^{(k)})^T d_k\}$;
 $\lambda_k = \rho^r$;
 $y^{(k)} = x^{(k)} + \lambda_k d_k$;
 - 2 **si** $\|g_P(y^{(k)})\| < \varepsilon$ **entonces**
 $x^* = y^{(k)}$;
Parar;
 - 3 **si** $\|y^{(k)} - x^{(k)}\| < \delta$ **entonces**
 $x^{(k+1)} = \tilde{T}(x^{(k)})$;
- en otro caso**
 $x^{(k+1)} = y^{(k)}$
-

Este algoritmo converge a la solución, pues si las iteraciones de Newton proyectado avanzan muy poco, el algoritmo realiza iteraciones de Weiszfeld proyectado que garantizan que el valor funcional decrece.

En los casos en los que (NP1) converge a la solución, el algoritmo híbrido coincide con éste método. Sin embargo, en los casos donde el (NP1) falla, el híbrido no alterna realmente entre

Newton y Weiszfeld, sino que realiza siempre las iteraciones de Weiszfeld, convirtiéndose así en un método de Weiszfeld proyectado más costoso. Esto lo mostraremos con ejemplos en la Sección 6.2.

Además, notamos que en los ejemplos analizados:

1. El método (NP1) llegaba a la solución sólo en los casos en que ésta coincidía con la solución irrestricta, es decir, cuando el minimizador del problema original x_{irr}^* se hallaba en Ω . Además, convergía en muy pocas iteraciones.
2. El método de Weiszfeld proyectado realizaba significativamente más iteraciones cuando $x^* \in \Omega$ que cuando no lo estaba.

Esto sugiere que puede ser conveniente utilizar el método de Newton irrestricto propuesto por Görner y Kanzow para determinar si $x_{irr}^* \in \Omega$, en cuyo caso la solución x_{Ω}^* del problema (4.3.1) coincide con x_{irr}^* , y si $x_{irr}^* \notin \Omega$, utilizar el método de Weiszfeld proyectado de Pilotta y Torres para hallar x_{Ω}^* .

Capítulo 6

Resultados numéricos

Implementamos todos los algoritmos utilizando el software GNU Octave, versión 4.0.0. Fueron ejecutados en una Notebook HP 250 G5, con procesador i5-6200U de 2.3GHz, 12Gb de memoria RAM, y sistema operativo Ubuntu 16.04.3 LTS de 64 bits con kernel 4.4.0-112-generic.

6.1. Problema irrestricto

A lo largo de este trabajo, presentamos tres métodos para resolver el problema de Fermat-Weber irrestricto: el método de Weiszfeld (WI), el método de Weiszfeld acelerado (WIA) propuesto por Brimberg, Chen y Chen [BCC98], y el método de Newton (NI) propuesto por Görner y Kanzow [GK16]. Los dos primeros además, pueden extenderse al problema de Fermat-Weber en norma p . A lo largo de esta sección, compararemos numéricamente estos métodos.

6.1.1. Comparación de métodos irrestrictos en norma 2

Para el caso en norma 2, se ejecutaron los tres algoritmos, (WI), (WIA) y (NI), partiendo del punto inicial $x^{(0)}$ sugerido por el método SP (Algoritmo 2, [BS15]) y con los mismos criterios de parada:

- Que se alcance una cantidad máxima de iteraciones $M = 500$.
- Que el gradiente de f sea pequeño: $\|\nabla f(x^{(k)})\| < \varepsilon = 10^{-6}$.

Para el método (WIA) se tomaron distintos valores de λ . Recordemos que cuando $p = 2$ la convergencia estaba garantizada para todo $\lambda > 0$ tal que $\lambda \leq \frac{2}{n-1}$, como indicaba la ecuación (4.2.13). Además, si $\lambda = 1$, el método coincide con el algoritmo de Weiszfeld original, (WI). Con esto en mente, los valores de λ tomados fueron $\lambda_1 = 1.25$, $\lambda_2 = 1.5$, $\lambda_3 = 1.75$, $\lambda_4 = 2$ y $\lambda_5 = \frac{2}{n-1}$. Notemos que para $n = 2$, $\lambda_5 = \lambda_4$.

Se varió la cantidad de nodos m , tomando los valores $m = 3$, $m = 10$, $m = 50$ y $m = 100$, y la dimensión n del espacio, tomando $n = 2$, $n = 3$, $n = 5$ y $n = 10$.

Para cada par (m, n) se construyeron cien ejemplos con pesos aleatorios, a los cuales se le aplicaron los siete métodos: (WI), (WIA) con los 5 valores de λ mencionados arriba y (NI). Para cada ejemplo se registró la cantidad de iteraciones y el tiempo de cada método, y luego se calcularon:

- I) El mínimo k_{min} y los cuatro cuartiles $k_{0.25}$, $k_{0.5}$, $k_{0.75}$ y k_{max} de la cantidad de iteraciones de cada método.
- II) La cantidad de veces que cada método alcanzó el máximo de iteraciones, F_{500} .
- III) La mediana $t_{0.5}$ y el tercer y cuarto cuartil $t_{0.75}$ y t_{max} de los tiempos de ejecución de cada método.

Lo mismo se realizó construyendo, para cada par (m, n) , cien ejemplos sin pesos.

Para fabricar los nodos en cada ejemplo, se utilizó el comando `unifrnd` de Octave para crear una matriz aleatoria $a \in \mathbb{R}^{m \times n}$

$$a = \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

cuyos elementos tuvieran una distribución uniforme en $[-m^2, m^2]$. En los ejemplos con pesos, se construyó $w = (w_1, \dots, w_m)$ mediante el comando `randi`, tomando valores enteros entre 1 y 10. En todos los ejemplos, se verificó que ningún nodo fuera solución.

En el Apéndice A, se encuentran las tablas con los resultados obtenidos. A continuación presentaremos algunas de ellas junto con un análisis de las mismas.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	20	51.50	87.00	166.00	500	9	0.010	0.019	0.060
(WIA)	1.25	15	40.00	68.50	132.00	500	8	0.008	0.015	0.059
	1.5	11	32.50	56.00	109.50	500	7	0.006	0.012	0.059
	1.75	8	27.00	47.50	93.00	500	7	0.005	0.010	0.061
	2	6	22.50	41.50	81.00	500	6	0.005	0.009	0.058
	$\frac{2}{n-1}$	6	22.50	41.50	81.00	500	6	0.005	0.009	0.059
(NI)	—	3	4.00	5.00	5.00	6	0	0.001	0.002	0.003

Tabla 6.1: Caso con pesos, tomando $m = 3$ y $n = 2$.

Para valores pequeños de m , en los casos con $n = 2$, los métodos (WI) y (WIA) presentaron gran variación en la cantidad de iteraciones para los distintos ejemplos. Esto puede verse por ejemplo en la Tabla 6.1, donde los cuartiles de la cantidad de iteraciones para (WI) toman valores muy diferentes, es decir que la muestra toma muchos valores distintos en el intervalo amplio entre el mínimo y el máximo. Algo similar ocurre con (WIA) para los distintos valores de λ , aunque toman valores menores, lo cual refleja un incremento en la velocidad de convergencia. El método (NI) en cambio, toma menos de diez iteraciones en todos estos ejemplos, lo cual demuestra una gran estabilidad, además de reflejar la tasa de convergencia cuadrática del método. En ese sentido, los distintos valores de λ en (WIA) presentan una aceleración de la velocidad de convergencia respecto a (WI), tanto en cantidad de iteraciones como en tiempo de ejecución. (NI) demuestra ser incluso más veloz, utilizando en casi todos los ejemplos menos iteraciones que en el mejor de los casos de los otros métodos, y disminuyendo los tiempos de ejecución.

Comparando la Tabla 6.1, la cual presenta los resultados para $m = 3$ y $n = 2$ con pesos, con la Tabla 6.2, que presenta para los mismos valores de m y n el caso sin pesos, podemos observar que los métodos (WI) y (WIA) son mucho más estables para el problema sin pesos que para el problema con pesos diferentes, en el sentido de la variación en cantidad de iteraciones. Esto puede verse notando que el tercer cuartil de la cantidad de iteraciones por método, $k_{0.75}$ para el caso sin pesos es menor que el primer cuartil, $k_{0.25}$, para el caso con pesos, en el ejemplo con $m = 3$, y $n = 2$. Esta situación se mantiene para $m = 3$ incluso variando el valor de n . Para otros valores de m , siguen obteniéndose mejores resultados para el caso sin pesos, aunque con una diferencia no tan marcada como en el caso con $m = 3$.

Para todos los valores de m y n , el algoritmo de Newton (NI) mantiene una cantidad de iteraciones entre 3 y 7, salvo en algunos casos extremos: para $m = 100$, en todos los valores de n el

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	29.50	36.00	66.50	500	2	0.004	0.007	0.060
(WIA)	1.25	17	22.50	27.50	52.50	500	1	0.003	0.006	0.056
	1.5	13	18.00	22.00	43.00	500	1	0.002	0.005	0.055
	1.75	9	14.00	18.00	36.50	500	1	0.002	0.004	0.055
	2	7	11.50	15.00	31.50	489	0	0.002	0.003	0.054
	$\frac{2}{n-1}$	7	11.50	15.00	31.50	489	0	0.002	0.003	0.055
(NI)	—	3	4.00	4.00	5.00	5	0	0.001	0.001	0.002

Tabla 6.2: Caso sin pesos, tomando $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	29.00	31.00	34.50	113	0	0.088	0.099	0.323
(WIA)	1.25	16	21.00	23.00	25.50	89	0	0.066	0.074	0.254
	1.5	12	16.00	17.50	20.00	73	0	0.051	0.057	0.210
	1.75	8	12.00	13.00	15.50	62	0	0.037	0.045	0.177
	2	5	9.00	10.00	12.50	54	0	0.029	0.036	0.154
	$\frac{2}{n-1}$	5	9.00	10.00	12.50	54	0	0.029	0.037	0.155
(NI)	—	4	4.00	4.00	4.00	500	2	0.025	0.025	37.423

Tabla 6.3: Caso con pesos, tomando $m = 100$ y $n = 2$.

método alcanzó el máximo de iteraciones al menos una vez, y lo mismo ocurrió para $m = 50$ con $n = 10$. La Tabla 6.3 muestra una de estas situaciones, para el caso $m = 100$ y $n = 2$. Podemos notar que en al menos el 75 % de los ejemplos, la cantidad de iteraciones de (NI) no supera las 4 iteraciones y el tiempo no supera tres centésimas de segundo, por lo cual significa que al menos en el 75 % de los casos es mejor que los otros métodos.

Un dato interesante que puede observarse en la Tabla 6.3 es que los métodos de Weiszfeld y de Weiszfeld acelerado funcionan mejor a medida que aumenta la complejidad del problema. Por ejemplo, en \mathbb{R}^2 el tercer cuartil de la cantidad de iteraciones de (WI) para $m = 3$ era más de cuatro veces mayor para $m = 100$, y en el caso de (WIA) con $\lambda = 2$ era más de seis veces mayor en $m = 3$ que en $m = 100$. Para $n = 2$, la mínima cantidad de iteraciones se mantuvo aproximadamente igual variando m pero el máximo disminuyó notablemente, presentándose un menor rango de valores. Variando n , también pudo observarse una mejoría de estos métodos mientras aumenta la dimensión del problema, disminuyendo la cantidad de iteraciones necesarias. La Tabla 6.4 muestra los resultados obtenidos con el método de (WI) para los casos $m = 3$ con $n = 2$, $m = 100$ con $n = 2$ y $m = 100$ con $n = 10$. Allí podemos notar cómo en $n = 2$ y $m = 3$ hay una gran amplitud de valores para la cantidad de iteraciones, mientras que aumentando la cantidad de nodos a $m = 100$, la amplitud es mucho menor, teniendo su mínimo por encima del mínimo para $m = 3$ y su valor máximo muy por debajo del tercer cuartil del caso $m = 3$. Para $m = 100$ y $n = 10$, se refleja el mismo aspecto de menor rango de valores para la cantidad de iteraciones, pero también un gran descenso en la cantidad de iteraciones, reduciéndose a aproximadamente un tercio de las iteraciones necesarias para resolver el problema con igual cantidad de nodos en $n = 2$.

En la Sección 4.2 probamos la convergencia del método (WIA), en el caso $p = 2$, para todo

n	m	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
2	3	20	51.50	87.00	166.00	500	9	0.010	0.019	0.060
2	100	22	29.00	31.00	34.50	113	0	0.088	0.099	0.323
10	100	10	10.00	11.00	11.00	12	0	0.031	0.032	0.035

Tabla 6.4: Velocidad del método de Weiszfeld para el caso con pesos, variando m y n .

$\lambda \leq \frac{2}{n-1}$. Para $n = 2$, podemos entonces asegurar la convergencia con la longitud de paso $\lambda = 2$, y generar entonces un método más rápido que el de Weiszfeld. En efecto, en todos los ejemplos de $n = 2$, el método (WIA) con $\lambda = 2$ resultó más rápido que (WI) y (WIA) con otros valores de λ . Para $n > 2$ en cambio, no podemos garantizar la convergencia para ningún valor de $\lambda > 1$, pues $\frac{2}{n-1} \leq 1$. Esto no dice que los pasos más largos vayan a fallar, pero tiene sentido esperar que lo hagan.

Observando las Tablas A.9 a A.16, que presentan los resultados para $n = 3$, pareciera que el método (WIA) converge para todos los valores de λ propuestos. Sin embargo, a medida que aumenta la cantidad de nodos, varía cuál de ellos es el más veloz, siendo $\lambda = 1.5$ en la mayoría de los casos. Lo mismo ocurre para valores mayores de n , y en algunos casos no podemos asegurar que el método efectivamente converja a la solución. Para $n = 10$ y $m = 10$ por ejemplo como podemos ver en la Tabla 6.5, el método (WIA) llegó a la solución en menos de 40 iteraciones para todos los valores de λ entre 1 y 2, tomó entre 67 y 104 iteraciones con $\lambda = \frac{2}{n-1} < 1$, pero para $\lambda = 2$ se alcanzó el máximo de iteraciones en 95 de los 100 ejemplos, lo cual no nos permite determinar si el método aproximó o no la solución.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	10	12.00	13.00	14.00	18	0	0.004	0.004	0.006
(WIA)	1.25	8	9.00	9.00	9.00	13	0	0.003	0.003	0.004
	1.5	14	15.00	16.00	16.00	17	0	0.005	0.005	0.006
	1.75	30	33.00	35.00	36.00	38	0	0.011	0.011	0.012
	2	351	500.00	500.00	500.00	500	95	0.155	0.156	0.161
	$\frac{2}{n-1}$	67	77.00	80.50	85.00	104	0	0.025	0.026	0.033
(NI)	—	3	3.00	4.00	4.00	4	0	0.003	0.003	0.003

Tabla 6.5: Caso sin pesos, tomando $m = 10$ y $n = 10$.

6.1.2. Comparación de métodos irrestrictos en norma p

Para el problema en norma p , se analizaron ejemplos variando p , m y n . Los valores de p elegidos fueron 1.1, 1.5 y 2.5, para m se eligieron los valores 3, 10 y 100, y utilizamos $n = 2$ y $n = 5$. Para cada terna (p, m, n) se construyeron cien ejemplos con pesos y cien ejemplos sin pesos, del mismo modo que para el problema en norma 2, y se verificó que ningún nodo fuera solución utilizando el resultado del Teorema 4.1.4.

Sobre cada uno de los ejemplos se ejecutaron los algoritmos de Weiszfeld para norma p (WI), visto en la Sección 4.1 y su aceleración (WIA), vista en la Sección 4.2. En cada caso se tomó como

punto inicial el punto

$$x^{(0)} = \frac{\sum_{i=1}^m w_i a_i}{\sum_{i=1}^m w_i}$$

y se usaron los siguientes criterios de parada:

- Que se alcance una cantidad máxima de iteraciones $M = 500$.
- Que el gradiente de f sea pequeño: $\|\nabla f(x^{(k)})\| < \varepsilon = 10^{-6}$.

Para el método (WIA) se tomaron los siguientes valores de λ :

$$\lambda_1 = 1.5, \quad \lambda_2 = 2, \quad \lambda_3 = \frac{2}{n(p-1)} \quad \lambda_4 = \frac{2}{(n-1)(p-1)} \quad \text{y} \quad \lambda_5 = \frac{2p}{n(p-1)}.$$

Para cada terna (p, m, n) y separando los casos con pesos y sin pesos, se registró la cantidad de iteraciones y el tiempo de cada método para cada ejemplo, y luego se calcularon:

- I) El mínimo k_{min} y los cuatro cuartiles $k_{0.25}$, $k_{0.5}$, $k_{0.75}$ y k_{max} de la cantidad de iteraciones de cada método.
- II) La cantidad de veces que cada método alcanzó el máximo de iteraciones, F_{500} .
- III) La mediana $t_{0.5}$ y el tercer y cuarto cuartil $t_{0.75}$ y t_{max} de los tiempos de ejecución de cada método.

Los resultados obtenidos se encuentran en el Apéndice B. A continuación, presentaremos algunos de los resultados más representativos, junto con un análisis de los mismos.

Análisis de los resultados para el caso $p = 1.1$

Recordemos que para el caso en que $p \geq 2$, la transformación Φ definida como en la Sección 4.1 está definida en todo punto salvo los nodos, pero para el caso $1 < p < 2$ la función es discontinua en los hiperplanos $H_{ij} = \{x \in \mathbb{R}^n : x_j = a_{ij}\}$, para todo $i = 1, \dots, m$, $j = 1, \dots, n$. Si en alguna iteración ocurre que $x^{(k)} \in H_{ij}$, las siguientes iteraciones cumplirán lo mismo, $x^{(k+1)} \in H_{ij}$, por lo cual la sucesión podría no converger o aproximarse a la solución lentamente, pues en lugar de avanzar en la dirección del gradiente debe avanzar por el hiperplano. Esto parece haberse visto reflejado en los resultados, para $p = 1.1$, principalmente para valores pequeños de m y n . La Tabla 6.6, que presenta los resultados del caso con pesos para $p = 1.1$, $m = 3$ y $n = 2$, muestra cómo los métodos (WI) y (WIA) en los casos λ_1 y λ_2 , alcanzaron el máximo de iteraciones en 30 ejemplos, y el método (WI) con λ_3 , λ_4 y λ_5 lo alcanzó en más del 80% de los ejemplos. Lo mismo ocurrió con $n = 5$.

Método	λ	Cantidad de Iteraciones					Tiempo en segundos			
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	114	144	154.5	500.0	500	29	0.031	0.098	0.106
(WIA)	1.5	74	92	101.0	500.0	500	30	0.021	0.101	0.106
	2	62	72	107.0	500.0	500	30	0.022	0.101	0.103
	10.000	8	500	500.0	500.0	500	91	0.101	0.101	0.109
	20.000	40	500	500.0	500.0	500	90	0.100	0.101	0.109
	11.000	7	500	500.0	500.0	500	86	0.101	0.101	0.107

Tabla 6.6: Caso con pesos, tomando $p = 1.1$, $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	127	500	500.0	500.0	500	99	0.100	0.102	0.110
(WIA)	1.5	79	500	500.0	500.0	500	99	0.104	0.105	0.112
	2	320	500	500.0	500.0	500	97	0.101	0.101	0.111
	10.000	323	500	500.0	500.0	500	94	0.100	0.101	0.106
	20.000	241	500	500.0	500.0	500	82	0.100	0.101	0.105
	11.000	308	500	500.0	500.0	500	83	0.100	0.101	0.107

Tabla 6.7: Caso sin pesos, tomando $p = 1.1$, $m = 3$ y $n = 2$.

Los resultados obtenidos para el caso sin pesos con $m = 3$ muestran que todos los métodos alcanzaron el máximo de iteraciones en más del 80% de los ejemplos, como podemos ver en la Tabla 6.7, que muestra el caso con $n = 2$. Esta situación no se repitió para otros valores de p o m , ni para el caso con pesos. Es decir, ocurrió sólo con los valores más pequeños de p y m elegidos, y para el caso sin pesos en el cual se toma $w_1 = \dots = w_m = 1$, a diferencia del caso ponderado en el cual $w_i \geq 1$ para todo i . Esto podría sugerir que la falla de los métodos se debe a errores numéricos causados por la presencia de números muy pequeños tomados como nulos por la máquina.

En general, el método (WIA) con los valores de λ

$$\lambda_3 = \frac{2}{n(p-1)} \quad \lambda_4 = \frac{2}{(n-1)(p-1)} \quad \text{y} \quad \lambda_5 = \frac{2p}{n(p-1)}$$

alcanzó el máximo de iteraciones en más del 75% de los ejemplos, salvo para algunos casos aislados. Esto puede deberse a que la longitud del paso es muy amplia. Los métodos (WI) y (WIA) con $\lambda_1 = 1.5$ y $\lambda_2 = 2$ fueron más eficientes. En todos los casos se encontró la solución sin llegar al máximo de iteraciones en más de la mitad de los ejemplos, como muestra por ejemplo la Tabla 6.8, y en algunos casos se alcanzó el máximo de iteraciones en menos del 25% de los casos, como muestra la Tabla 6.9, en la cual esto ocurrió sólo 9 veces.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	134	148	194.0	500.0	500	41	0.055	0.138	0.148
(WIA)	1.5	86	96	128.0	500.0	500	40	0.037	0.140	0.148
	2	63	74	109.0	500.0	500	36	0.031	0.140	0.145
	4.000	30	500	500.0	500.0	500	93	0.140	0.141	0.145
	5.000	34	500	500.0	500.0	500	93	0.140	0.141	0.143
	4.400	26	500	500.0	500.0	500	94	0.140	0.141	0.145

Tabla 6.8: Caso con pesos, tomando $p = 1.1$, $m = 3$ y $n = 5$.

Los pasos λ_3 , λ_4 y λ_5 para el método (WIA), propuestos en la Sección 4.2, no fueron eficientes pues se alcanzó el máximo de iteraciones en la mayoría de los casos. Los pasos λ_1 y λ_2 permitieron en la mayoría de los ejemplos hallar la solución, y su velocidad de convergencia fue mejor que la del método (WI).

Análisis de los resultados para el caso $p = 1.5$

Para el caso $p = 1.5$ se obtuvieron mejores resultados que para el caso $p = 1.1$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	113	144	166.5	183.5	500	9	0.352	0.395	1.143
(WIA)	1.5	80	92	105.0	120.0	500	9	0.228	0.255	1.136
	2	59	68	81.0	88.5	500	9	0.171	0.194	1.122
	10.000	8	500	500.0	500.0	500	76	1.058	1.063	1.110
	20.000	19	500	500.0	500.0	500	98	1.039	1.044	1.135
	11.000	7	500	500.0	500.0	500	80	1.056	1.060	1.145

Tabla 6.9: Caso sin pesos, tomando $p = 1.1$, $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	23	30	33.0	45.5	500	1	0.007	0.009	0.099
(WIA)	1.5	12	17	19.0	29.0	500	1	0.004	0.006	0.101
	2	8	11	13.0	21.0	500	1	0.003	0.004	0.103
	2.000	8	11	13.0	21.0	500	1	0.003	0.004	0.103
	4.000	31	500	500.0	500.0	500	91	0.101	0.102	0.107
	3.000	10	18	500.0	500.0	500	61	0.101	0.102	0.109

Tabla 6.10: Caso sin pesos, tomando $p = 1.5$, $m = 3$ y $n = 2$.

En primer lugar, podemos notar una mejoría en los resultados para (WIA) usando λ_3 , λ_4 y λ_5 . En todos los casos, $\lambda_3 = \frac{2}{n(p-1)}$ halló la solución en casi todos los ejemplos, y en algunos de los casos llegó a resolver todos los ejemplos. Esto tiene sentido, pues para ese valor de λ probamos la convergencia del algoritmo en la Sección 4.2. Los valores λ_4 y λ_5 fueron construidos a partir de aproximaciones, basadas en valores que aseguraban la convergencia en el caso $p = 2$. Luego, podría ocurrir que no generaran una sucesión convergente. En efecto, para valores pequeños de m y n el método (WIA) usando λ_5 alcanzó el máximo de iteraciones en más del 50% de los ejemplos, y usando λ_4 esto ocurrió en más del 75% de los casos, como podemos ver por ejemplo en la Tabla 6.10, con $n = 2$ y $m = 3$. Sin embargo, para valores mayores de m esta situación fue mejorando. Para $n = 2$ y $m = 100$ por ejemplo, se alcanzó el máximo de iteraciones sólo en el método (WIA), en 8 ejemplos para λ_4 y uno sólo para λ_5 , como muestra la Tabla 6.11.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	30	39	42.0	46.0	65	0	0.090	0.098	0.137
(WIA)	1.5	16	23	25.0	28.0	41	0	0.054	0.060	0.086
	2	11	15	17.0	18.5	28	0	0.036	0.039	0.059
	2.000	11	15	17.0	18.5	28	0	0.036	0.040	0.061
	4.000	11	14	18.0	23.5	500	8	0.038	0.050	1.074
	3.000	5	7	8.0	9.0	500	1	0.017	0.020	1.056

Tabla 6.11: Caso sin pesos, tomando $p = 1.5$, $m = 100$ y $n = 2$.

Para $n = 5$, los valores de λ tomados fueron $\lambda_1 = 1.5$, $\lambda_2 = 2$, $\lambda_3 = 0.8$, $\lambda_4 = 1$ y $\lambda_5 = 1.2$, que son valores pequeños. Esto se ve reflejado en los resultados, pues salvo en algunos casos particulares, la solución se halló sin llegar al máximo de iteraciones en la mayoría de los ejemplos. Para $m = 10$ por ejemplo, todos los métodos hallaron la solución para todos los ejemplos, como muestra la Tabla 6.12.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	27	33	36.0	40.0	58	0	0.022	0.024	0.034
(WIA)	1.5	14	19	21.0	24.0	36	0	0.013	0.014	0.022
	2	9	12	13.0	16.0	167	0	0.008	0.010	0.104
	0.800	36	43	47.5	52.0	74	0	0.029	0.031	0.044
	1.000	27	33	36.0	40.0	58	0	0.022	0.024	0.035
	1.200	21	26	29.0	32.0	47	0	0.017	0.019	0.028

Tabla 6.12: Caso con pesos, tomando $p = 1.5$, $m = 10$ y $n = 5$.

Al igual que en el caso con norma 2, en los ejemplos de mayor complejidad, es decir, aumentando los valores de m y n , los métodos fueron más eficientes, reduciéndose la cantidad de iteraciones necesarias en cada caso. El valor de λ para el método (WIA) que generó la convergencia más rápida fue variando al cambiar la dimensión y la cantidad de nodos. Si bien λ_5 fue más rápido en algunos casos, también falló en más casos que otras longitudes de paso. En la mayoría de los casos $\lambda_2 = 2$ probó ser un paso efectivo, alcanzando el máximo de iteraciones en tantos casos como (WI), y con una mejor velocidad de convergencia.

Análisis de los resultados para el caso $p = 2.5$

Para el problema de Fermat-Weber en norma $p > 2$, Brimberg y Love [BL92] probaron que el algoritmo puede no converger, y posteriormente Brimberg, Chen y Chen [BCC98] probaron que para ciertos valores de λ la variación dada por $x^{(k+1)} = x^{(k)} - \lambda(\Phi(x^{(k)}) - x^{(k)})$, donde Φ es la transformación de Weiszfeld, converge al minimizador de $f^{(p)}$. Esta variación es la que llamamos hasta ahora *método de Weiszfeld acelerado*, aunque los valores de λ podrían ser menores que 1.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	13	33	59.0	138.0	500	3	0.012	0.028	0.102
(WIA)	1.5	7	24	61.5	500.0	500	29	0.012	0.101	0.103
	2	10	500	500.0	500.0	500	83	0.101	0.102	0.107
	0.667	24	53	92.0	211.0	500	5	0.019	0.043	0.104
	1.333	8	23	42.5	104.0	500	2	0.009	0.021	0.101
	1.667	9	40	500.0	500.0	500	54	0.101	0.101	0.104

Tabla 6.13: Caso con pesos, tomando $p = 2.5$, $m = 3$ y $n = 2$.

En la práctica, la mayoría de los ejemplos pudieron resolverse con el método de (WI), a pesar de que en la teoría no esté asegurada la convergencia. El método (WIA) también presentó buenos resultados, salvo para los valores de λ más altos, como $\lambda_2 = 2$. Esto puede verse en la Tabla 6.13, donde se ve que alcanzó el máximo de iteraciones en 83 de los 100 ejemplos.

Para $n = 5$, los valores de λ para los cuales la convergencia de (WIA) está garantizada, λ_3 , λ_4 y λ_5 , son menores que 1. Luego, si bien estos métodos son efectivos, el método (WI) resulta más rápido. Por otro lado, los valores más grandes de λ , $\lambda_1 = 1.5$ y $\lambda_2 = 2$ no permitieron hallar la solución sin alcanzar el máximo de iteraciones. Esto puede verse por ejemplo en el caso $m = 10$ y $n = 5$ sin pesos, presentado en la Tabla 6.14.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	10	13	15.0	16.0	27	0	0.009	0.010	0.016
(WIA)	1.5	500	500	500.0	500.0	500	100	0.301	0.302	0.306
	2	500	500	500.0	500.0	500	100	0.301	0.302	0.305
	0.267	43	51	56.0	62.5	121	0	0.034	0.038	0.073
	0.333	33	40	43.0	48.5	96	0	0.026	0.029	0.058
	0.667	13	16	18.0	21.0	44	0	0.011	0.013	0.026

Tabla 6.14: Caso sin pesos, tomando $p = 2.5$, $m = 10$ y $n = 5$.

Por último, al igual que lo notado en los casos $p = 2$ y $p = 1.5$, el método pareciera ser más veloz para valores mayores de m y n , como podemos ver en la Tabla 6.15, que compara los resultados obtenidos por el método de Weiszfeld para $p = 2.5$, en distintos valores de m y n .

n	m	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
2	3	13	33	59.0	138.0	500	3	0.012	0.028	0.102
2	100	14	21	23.0	26.0	78	0	0.048	0.055	0.164
5	100	8	10	10.0	11.0	13	0	0.048	0.052	0.061

Tabla 6.15: Velocidad del método de Weiszfeld para el caso con pesos, tomando $p = 2.5$ y variando m y n .

6.2. Problema con restricciones de cajas

Para el problema con restricciones de tipo caja, el análisis realizado fue distinto. En lugar de construir ejemplos aleatorios, se diseñaron problemas que permitieran comparar los distintos tipos de situaciones que pueden darse al incorporar restricciones al problema.

Supongamos que tenemos nodos a_1, \dots, a_m , pesos w_1, \dots, w_m y la solución irrestricta x^* . Si se incorpora una caja Ω , puede darse una gran cantidad de situaciones:

- Con respecto a la cápsula convexa de los nodos, puede ocurrir que $\mathcal{C}(\mathcal{A}) \subset \Omega$, $\Omega \subset \mathcal{C}(\mathcal{A})$, $\Omega \cap \mathcal{C}(\mathcal{A}) = \emptyset$ o ninguna de ellas.
- Con respecto a los nodos, puede ocurrir que ningún nodo esté en Ω , que todos los nodos estén en Ω o que haya algunos nodos en Ω y otros fuera.
- Con respecto a la solución del problema irrestricto, puede ocurrir que $x^* \in \Omega$ o que $x^* \notin \Omega$.

Estas situaciones tienen distintos efectos en el problema restringido. Por ejemplo, si $\mathcal{A} \in \Omega$, la función $f(x) = \sum_{i=1}^m w_i \|x - a_i\|$ no es diferenciable en todo Ω , y por lo tanto podría ocurrir que una iteración coincida con un nodo y el algoritmo se detenga. Si $x^* \in \Omega$, entonces la solución

del problema restringido coincide con la solución irrestricta. Con esto en mente, generamos cuatro problemas y variamos las cajas utilizadas. Construimos

1. Un problema en \mathbb{R}^2 con pesos distintos, y uno sin pesos. Para cada uno, se aplicaron nueve cajas distintas, incluyendo el plano completo.
2. Un problema en \mathbb{R}^3 con pesos distintos, y uno sin pesos. Para cada uno, se aplicaron siete cajas distintas, incluyendo el espacio completo.

Para cada ejemplo, se aplicaron: el algoritmo de Weiszfeld proyectado (WP) propuesto por Pilotta y Torres [PT11], las dos versiones de Newton proyectado, (NP1) y (NP2) dadas por el Algoritmo 5 y el Algoritmo 6 respectivamente, y el algoritmo híbrido propuesto para resolver las fallas de Newton proyectado, dado por el Algoritmo 7.

Además, se ejecutó una variación del algoritmo de Weiszfeld proyectado, utilizando la variación propuesta por Brimberg, Chen y Chen [BCC98] para acelerar el algoritmo de Weiszfeld. Usando el operador ψ definido en la Sección 4.2 como

$$\psi(x) = x - \frac{\lambda}{L(x)} \nabla f(x)$$

el nuevo método, al cual denotaremos con (WPA), consiste en proyectar la transformación, tomando $\tilde{\psi} = P \circ \psi$. El algoritmo para este método es el siguiente:

Algoritmo 8: Aceleración del algoritmo de Weiszfeld proyectado

Datos: $a_1, \dots, a_m \in \mathbb{R}^n$; $w_1, \dots, w_m \in \mathbb{R}_{>0}$; $-\infty \leq l_j < u_j \leq +\infty$, $j = 1, \dots, n$

Resultado: $x^* \in \mathbb{R}^n$ que minimiza $f(x)$

Elegir $x^{(0)} \neq a_1, \dots, a_m$;

para $k = 0, 1, 2, \dots$ **hacer**

$y^{(k+1)} \leftarrow \psi(x^{(k)})$;
 $x^{(k+1)} \leftarrow P(y^{(k+1)})$;

Para los experimentos numéricos, se ejecutó (WPA) fijando $\lambda = 2$.

Como criterios de parada se tomaron:

- Que se alcance un máximo de iteraciones $M = 100$.
- Que la distancia entre dos iteraciones consecutivas sea pequeña, $\|x^{(k+1)} - x^{(k)}\| < 10^{-5}$.

Para cada ejemplo, se registraron la cantidad de iteraciones de cada método. Para el caso particular del método híbrido, se registró no solo el total sino también cuántas veces fue necesario aplicar una iteración de Weiszfeld. Además, sabiendo que el método de Weiszfeld proyectado encuentra la solución correcta, se compararon los resultados obtenidos por cada método con los obtenidos por (WP), calculando para cada método la diferencia entre los valores funcionales obtenidos,

$$f(x_{\text{Método}}^*) - f(x_{(WP)}^*)$$

donde $x_{\text{Método}}^*$ denota la solución hallada por cada método.

En todos los ejemplos se eligió como punto inicial la proyección a Ω del punto generado por el método SP.

A continuación, presentaremos los resultados obtenidos, primero para los problemas en \mathbb{R}^2 y luego para los problemas en \mathbb{R}^3 .

6.2.1. El problema en \mathbb{R}^2

Para \mathbb{R}^2 nos basamos en el experimento realizado por Pilotta y Torres [PT11] para mostrar la convergencia del método de Weiszfeld proyectado.

Para todos los ejemplos, utilizamos los nodos propuestos en [PT11]: $a_1 = (1, 0)$, $a_2 = (0, 0)$, $a_3 = (0, 1)$ y $a_4 = (1, 4)$. Al igual que para el caso irrestricto, analizamos por un lado el caso sin pesos y por otro lado el caso ponderado, utilizando los siguientes pesos: $w_1 = 5$, $w_2 = 3$, $w_3 = 2$ y $w_4 = 3$. Además, se construyeron 9 cajas, de distintos tamaños y ubicaciones, dadas por:

1. $\Omega_1 = \mathbb{R}^2$ que es el caso irrestricto.
2. $\Omega_2 = [0, 1] \times [1, 3.5]$
3. $\Omega_3 = [0.25, 0.75] \times [0, 1]$
4. $\Omega_4 = [0.5, 1.5] \times [3.5, 4.5]$
5. $\Omega_5 = [0, +\infty) \times [0, +\infty)$
6. $\Omega_6 = [2, 4] \times [-2, -1]$
7. $\Omega_7 = [-1, 2] \times [-1, 5]$
8. $\Omega_8 = (-\infty, -1] \times \mathbb{R}$
9. $\Omega_9 = [0.25, 0.5] \times [0.25, 0.5]$

En la Figura 6.1, pueden verse las cajas 2, 4, 6, 8 y 9, junto con los nodos, la cápsula convexa de \mathcal{A} y la solución irrestricta al problema con pesos.

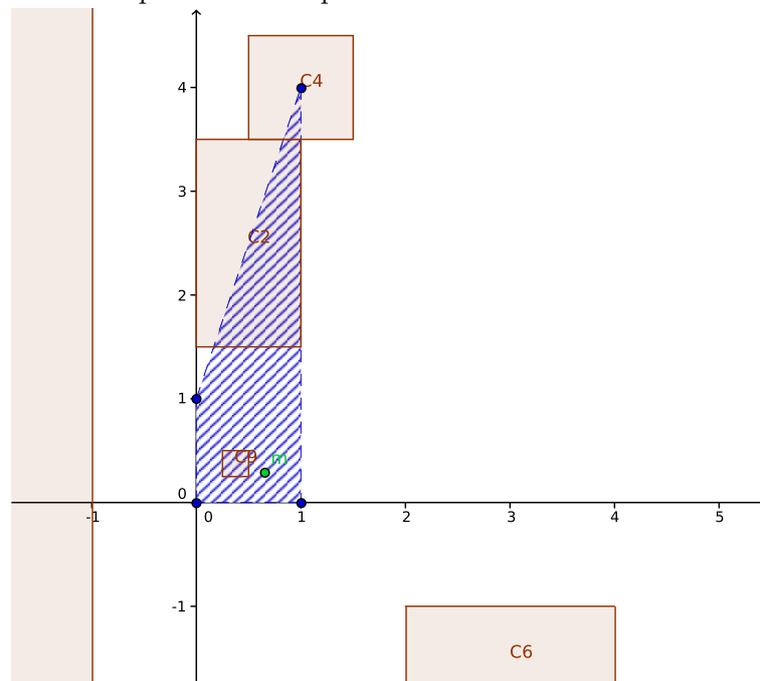


Figura 6.1: Gráfico de algunas de las cajas utilizadas. El punto llamado m denota el minimizador del problema irrestricto en el caso con pesos, y el área sombreada con rayas representa $\mathcal{C}(\mathcal{A})$.

Como podemos ver en las Tablas 6.16 y 6.17, los métodos de Newton proyectado se detuvieron en muy pocas iteraciones en todos los casos, pero para las cajas 2, 4, 8 y 9, y la caja 3, sólo en el caso sin pesos, el resultado obtenido no estuvo cerca de la solución, como podemos ver en las diferencias en valor funcional con el resultado obtenido por (WP). La solución al problema irrestricto para el problema ponderado es el punto $x_{cp}^* = (0.65398, 0.29276)$ y para el problema

sin pesos es $x_{sp}^* = (0.20003, 0.79995)$. Estudiando las cajas, podemos ver que justamente cuando $x^* \notin \Omega$ los métodos de Newton proyectado fallan, salvo para la caja 6. Para este último ejemplo, todos los métodos hallaron la solución exacta, pues la solución coincide con el punto inicial.

Ω	Cantidad de iteraciones						Diferencia con (WP) en valor funcional			
	(WP)	(WPA)	(NP1)	(NP2)	Híbrido		(WPA)	(NP1)	(NP2)	Híbrido
					Tot.	Weis.				
1	39	19	4	4	4	0	-2.90e-09	-3.91e-09	-3.91e-09	-3.91e-09
2	8	21	2	2	8	7	-3.46e-12	1.22e-02	1.90e-01	-1.36e-10
3	37	19	3	4	3	0	-2.57e-09	-2.82e-09	-2.83e-09	-2.82e-09
4	4	40	2	2	5	4	-3.42e-10	1.58e-04	1.61e-04	-4.50e-10
5	39	19	4	4	4	0	-2.90e-09	-3.91e-09	-3.91e-09	-3.91e-09
6	1	1	1	1	1	0	0.00e+00	0.00e+00	0.00e+00	0.00e+00
7	39	19	4	4	4	0	-2.90e-09	-3.91e-09	-3.91e-09	-3.91e-09
8	9	20	3	3	10	8	3.11e-12	2.61e-01	2.61e-01	-2.04e-11
9	11	7	5	5	14	10	-2.32e-10	7.20e-03	7.20e-03	-3.62e-10

Tabla 6.16: Resultados obtenidos para el caso con pesos

Cuando la solución irrestricta coincide con la solución sobre Ω , los métodos de Newton proyectado son mucho mas eficientes que el método de Weiszfeld proyectado, pero esto se debe a que ambos algoritmos realizan iteraciones del método irrestricto, en el cual ya demostramos la mayor velocidad de el método de Newton. Cuando el método de Newton se detiene antes de tiempo, el algoritmo híbrido encuentra la solución utilizando iteraciones de Weiszfeld proyectado en aproximadamente el total de las iteraciones, y demora al menos tantas iteraciones como (WP). Luego, el híbrido resulta una versión costosa de (WP) cuando (NP1) falla, y coincide con (NP1) cuando éste funciona.

Ω	Cantidad de iteraciones						Diferencia con (WP) en valor funcional			
	(WP)	(WPA)	(NP1)	(NP2)	Híbrido		(WPA)	(NP1)	(NP2)	Híbrido
					Tot.	Weis.				
1	47	24	4	4	4	0	-1.16e-09	-1.44e-09	-1.44e-09	-1.44e-09
2	7	25	1	1	7	7	-3.39e-12	5.04e-02	5.04e-02	-3.27e-11
3	18	6	7	7	22	16	-4.93e-10	1.54e-03	1.54e-03	-4.07e-10
4	7	34	1	1	7	6	2.15e-11	7.03e-02	7.03e-02	0.00e+00
5	47	24	4	4	4	0	-1.16e-09	-1.44e-09	-1.44e-09	-1.44e-09
6	1	1	1	1	1	0	0.00e+00	0.00e+00	0.00e+00	0.00e+00
7	47	24	4	4	4	0	-1.16e-09	-1.44e-09	-1.44e-09	-1.44e-09
8	8	16	2	2	8	7	2.72e-12	2.35e-03	2.35e-03	-1.36e-11
9	9	9	3	3	11	9	-5.52e-11	2.40e-03	2.40e-03	-8.46e-11

Tabla 6.17: Resultados obtenidos para el caso sin pesos

Entre los métodos (WP) y (WPA), es difícil destacar uno por sobre el otro. Ambos métodos encuentran la solución independientemente de la caja. La única diferencia notable entre ambos es

que el método de Weiszfeld proyectado es más veloz cuando la solución irrestricta no coincide con la solución sobre Ω , mientras que el método acelerado es más veloz cuando la solución coincide con la irrestricta. Promediando la cantidad de iteraciones para cada ejemplo de cada método, obtenemos que para el caso ponderado, (WP) toma en promedio 20.778 iteraciones y (WPA) 18.333 iteraciones, y en el caso sin pesos, el promedio es 21.222 para (WP) y 18.111 para (WPA). Esto pareciera señalar que el método acelerado es en general mejor que el método de Weiszfeld proyectado. Sin embargo, debido a que se ejecutaron muy pocos ejemplos, no podemos asegurar que esto sea así.

6.2.2. El problema en \mathbb{R}^3

Para todos los ejemplos en \mathbb{R}^3 , utilizamos los siguientes nodos:

$$a_1 = (0, 0, 0), \quad a_2 = (0, 0, 1), \quad a_3 = (0, 1, 0), \quad a_4 = (0, 1, 2),$$

$$a_5 = (1, 0, 0), \quad a_6 = (1, 0, 1), \quad a_7 = (1, 1, 0) \quad \text{y} \quad a_8 = (1, 1, 3).$$

En la Figura 6.2, puede verse una representación de los nodos y su cápsula convexa.

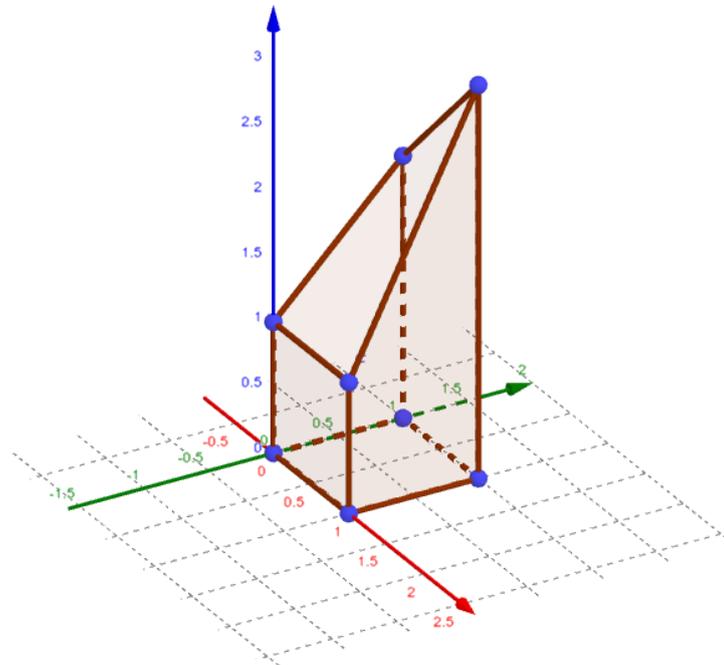


Figura 6.2: Representación de los nodos del problema en \mathbb{R}^3 .

Al igual que para el caso irrestricto, analizamos por un lado el caso sin pesos y por otro lado el caso ponderado, utilizando los siguientes pesos: $w_1 = 10$, $w_2 = 2$, $w_3 = 5$, $w_4 = 3$, $w_5 = 4$, $w_6 = 8$, $w_7 = 1$ y $w_8 = 7$. Además, se construyeron 7 cajas, de distintos tamaños y ubicaciones, dadas por:

1. $\Omega_1 = \mathbb{R}^3$ que es el caso irrestricto.
2. $\Omega_2 = [0.1, 0, 9] \times [0.1, 0, 9] \times [0.1, 0, 9]$
3. $\Omega_3 = [-2, 2] \times [-1, 5] \times [-3, 5]$
4. $\Omega_4 = [0, 1] \times [0, 1] \times [0, 0.5]$

5. $\Omega_5 = [2, 3] \times [1, 2] \times [3, 5]$
6. $\Omega_6 = [0.25, 0.75] \times [0.25, 0.75] \times [1, 2]$
7. $\Omega_7 = [2, \infty) \times (-\infty, \infty) \times (-\infty, \infty)$

Al igual que en el caso en \mathbb{R}^2 , los métodos (NP1) y (NP2) se detuvieron en pocas iteraciones pero no en todos los casos encontraron la solución. En las cajas 4 a 7 la solución irrestricta no se encontraba dentro de la caja. Para la quinta caja, la solución sobre Ω coincidía con el punto inicial, por lo que todos los métodos hallaron la solución exacta. Sin embargo en los otros tres casos, los métodos (NP1) y (NP2) no hallaron la solución, como podemos ver en la comparación en valor funcional con el método (WP). En estos casos, el método híbrido realizó al menos tantas iteraciones como (WP), utilizando iteraciones de Weiszfeld proyectado en la mayoría de los casos. Al igual que para el caso en \mathbb{R}^2 , esto nos lleva a concluir que el híbrido una versión costosa de (WP) cuando (NP1) falla, y coincide con (NP1) cuando éste funciona.

Ω	Cantidad de iteraciones						Diferencia con (WP) en valor funcional			
	(WP)	(WPA)	(NP1)	(NP2)	Híbrido		(WPA)	(NP1)	(NP2)	Híbrido
					Tot.	Weis.				
1	21	27	3	4	3	0	-1.64e-09	-2.16e-09	-2.16e-09	-2.16e-09
2	21	27	3	4	3	0	-1.64e-09	-2.16e-09	-2.16e-09	-2.16e-09
3	21	27	3	4	3	0	-1.64e-09	-2.16e-09	-2.16e-09	-2.16e-09
4	11	18	4	4	13	10	-1.56e-09	2.93e-02	2.93e-02	-1.81e-09
5	1	1	1	1	1	0	0.00e+00	0.00e+00	0.00e+00	0.00e+00
6	12	5	2	2	12	11	-9.29e-10	7.74e-03	7.74e-03	-1.45e-09
7	7	77	1	1	7	7	9.48e-11	8.65e-02	8.65e-02	-1.45e-10

Tabla 6.18: Resultados obtenidos para el caso con pesos

Ω	Cantidad de iteraciones						Diferencia con (WP) en valor funcional			
	(WP)	(WPA)	(NP1)	(NP2)	Híbrido		(WPA)	(NP1)	(NP2)	Híbrido
					Tot.	Weis.				
1	13	13	2	3	2	0	-9.17e-11	-1.43e-10	-1.48e-10	-1.43e-10
2	13	13	2	3	2	0	-9.17e-11	-1.43e-10	-1.48e-10	-1.43e-10
3	13	13	2	3	2	0	-9.17e-11	-1.43e-10	-1.48e-10	-1.43e-10
4	9	13	2	2	8	7	-2.08e-10	4.50e-04	4.50e-04	-2.58e-10
5	1	1	1	1	1	0	0.00e+00	0.00e+00	0.00e+00	0.00e+00
6	10	12	4	4	11	8	-9.55e-11	8.73e-04	8.73e-04	-1.43e-10
7	7	45	2	2	7	6	-1.88e-11	1.69e-02	1.69e-02	-6.81e-11

Tabla 6.19: Resultados obtenidos para el caso sin pesos

En cuanto al método (WPA), probó ser eficiente en todos los ejemplos, resolviéndolos en pocas iteraciones. Sin embargo, en general el método (WP) fue más veloz. Algo llamativo es que si bien en la mayoría de los ejemplos ambos métodos tuvieron rendimientos similares, en la caja $\Omega_7 = [2, \infty) \times (-\infty, \infty) \times (-\infty, \infty)$ el método (WPA) tomó 77 iteraciones para el caso ponderado

y 45 en el caso sin pesos, contra 7 iteraciones que precisó (WP). Hacer pasos más largos hacia la solución irrestricta podría causar que el paso proyectado se exceda de la solución sobre Ω , debiendo realizar más pasos para corregir esta situación, lo cual podría justificar que en este caso aumente tanto la cantidad de iteraciones.

Capítulo 7

Conclusiones

En este trabajo se estudiaron distintas variaciones del problema de Fermat-Weber, y se extendieron los resultados de convergencia del método acelerado de Brimberg, Chen y Chen [BCC98] a \mathbb{R}^n . Además se analizó la efectividad de distintos métodos para resolver estas variantes: el algoritmo de Weiszfeld, el algoritmo acelerado, el método de Newton sugerido por Görner y Kanzow [GK16] y las modificaciones de los mismos para resolver las distintas variantes del problema.

El algoritmo de Weiszfeld es el método más utilizado para resolver el problema de Fermat-Weber, pues es el método más antiguo y ha sido estudiado en profundidad por diversos autores. Además, es un método sencillo de programar y aplicar, y su convergencia puede garantizarse realizando pequeñas modificaciones, por lo cual es una manera segura de hallar la solución. Otra gran ventaja de este método es que puede generalizarse para resolver otros problemas similares, como el problema de Fermat-Weber con restricciones de tipo caja y el problema en norma $p < 2$. Sin embargo para este último resulta necesario el desarrollo de una estrategia similar a la dada por el método SP de Beck y Sabach [BS15], que permita evitar que las iteraciones del método coincidan con puntos de los hiperplanos H_{ij} .

Por otro lado, la variación propuesta por Brimberg, Chen y Chen mostró ser efectiva para acelerar el método en norma $p \leq 2$ y para generar un método convergente para el problema en norma $p > 2$. Además, el método acelerado para norma 2 demostró ser compatible con la proyección a cajas, permitiendo usarlo para acelerar el método de Weiszfeld proyectado. Para esta aplicación sin embargo, resta realizar una demostración de convergencia y en todas las aplicaciones es necesario estudiar en mayor profundidad la elección de los valores de λ .

El algoritmo de Newton propuesto por Görner y Kanzow es digno de resaltar, pues es un método eficiente y con una muy buena tasa de convergencia. Además, gracias a las características de f , como que sea estrictamente convexa y que f , ∇f y $\nabla^2 f$ tengan estructuras similares, el algoritmo no consume tanta memoria ni requiere tantas operaciones como el método de Newton aplicado a otras funciones. Un posible trabajo a futuro podría ser generalizar el método de Newton al problema en norma p .

Los algoritmos desarrollados mediante la proyección del método de Newton para el problema con restricciones no fueron efectivos, pues en la mayoría de los casos los pasos realizados fueron muy cortos generando que el método se detenga. El método híbrido formulado para resolver este problema no fue eficiente, pues no realizó una combinación real entre ambos métodos sino que realizó en algunos casos Newton irrestricto y en otros casos Weiszfeld proyectado. Algo a destacar de los experimentos realizados es que se descubrió que el método de Weiszfeld proyectado es más veloz cuando la solución irrestricta no se encuentra en la caja. Esto sugiere una posible reformulación para el método híbrido que consista en hallar la solución irrestricta mediante el algoritmo de Newton irrestricto, y si esta no se encuentra en la caja, utilizar el método de Weiszfeld proyectado para hallar la solución restringida a la caja. De este modo, el método híbrido tomaría alrededor de cuatro iteraciones si la solución irrestricta coincide con la solución restringida, como el método de Newton, y cuando esto no ocurre toma sólo alrededor de cuatro iteraciones más que el método de Weiszfeld

proyectado. Esto generaría en general un método más veloz, pero sólo para valores pequeños de m , pues vimos que Weiszfeld es más veloz para valores grandes de m . Luego, aún resta encontrar una variación eficiente del método de Newton para hallar la solución del problema de Fermat-Weber restringido a cajas.

Apéndice A

Tablas de resultados para el problema irrestricto en norma 2

A.1. Resultados para $n = 2$

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	20	51.50	87.00	166.00	500	9	0.010	0.019	0.060
(WIA)	1.25	15	40.00	68.50	132.00	500	8	0.008	0.015	0.059
	1.5	11	32.50	56.00	109.50	500	7	0.006	0.012	0.059
	1.75	8	27.00	47.50	93.00	500	7	0.005	0.010	0.061
	2	6	22.50	41.50	81.00	500	6	0.005	0.009	0.058
	$\frac{2}{n-1}$	6	22.50	41.50	81.00	500	6	0.005	0.009	0.059
(NI)	–	3	4.00	5.00	5.00	6	0	0.001	0.002	0.003

Tabla A.1: Caso con pesos, tomando $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	22	29.50	36.00	66.50	500	2	0.004	0.007	0.060
(WIA)	1.25	17	22.50	27.50	52.50	500	1	0.003	0.006	0.056
	1.5	13	18.00	22.00	43.00	500	1	0.002	0.005	0.055
	1.75	9	14.00	18.00	36.50	500	1	0.002	0.004	0.055
	2	7	11.50	15.00	31.50	489	0	0.002	0.003	0.054
	$\frac{2}{n-1}$	7	11.50	15.00	31.50	489	0	0.002	0.003	0.055
(NI)	–	3	4.00	4.00	5.00	5	0	0.001	0.001	0.002

Tabla A.2: Caso sin pesos, tomando $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	20	35.00	42.50	59.00	164	0	0.013	0.018	0.051
(WIA)	1.25	14	27.00	33.00	46.00	131	0	0.010	0.014	0.040
	1.5	11	21.00	26.00	37.00	108	0	0.008	0.012	0.033
	1.75	8	17.00	21.00	31.00	92	0	0.007	0.010	0.029
	2	6	13.00	18.00	26.00	80	0	0.006	0.008	0.026
	$\frac{2}{n-1}$	6	13.00	18.00	26.00	80	0	0.006	0.008	0.025
(NI)	—	3	4.00	4.00	5.00	6	0	0.003	0.004	0.004

Tabla A.3: Caso con pesos, tomando $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	30.00	35.50	46.00	193	0	0.011	0.014	0.061
(WIA)	1.25	17	23.00	27.00	35.00	154	0	0.009	0.011	0.052
	1.5	13	18.00	21.50	28.50	127	0	0.007	0.009	0.041
	1.75	9	15.00	17.00	23.50	109	0	0.006	0.007	0.038
	2	6	12.00	14.00	20.00	95	0	0.004	0.006	0.030
	$\frac{2}{n-1}$	6	12.00	14.00	20.00	95	0	0.004	0.006	0.030
(NI)	—	3	4.00	4.00	5.00	5	0	0.003	0.004	0.008

Tabla A.4: Caso sin pesos, tomando $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	24	30.00	33.00	37.00	268	0	0.048	0.054	0.388
(WIA)	1.25	18	22.00	25.00	28.00	213	0	0.036	0.041	0.307
	1.5	11	17.00	19.00	22.00	177	0	0.027	0.032	0.259
	1.75	10	13.00	14.50	18.00	151	0	0.022	0.026	0.218
	2	6	10.00	11.50	14.00	132	0	0.017	0.020	0.193
	$\frac{2}{n-1}$	6	10.00	11.50	14.00	132	0	0.017	0.020	0.192
(NI)	—	4	4.00	4.00	4.00	500	1	0.013	0.013	16.187

Tabla A.5: Caso con pesos, tomando $m = 50$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	20	26.50	29.00	32.00	71	0	0.043	0.047	0.107
(WIA)	1.25	14	20.00	22.00	24.00	56	0	0.032	0.036	0.082
	1.5	11	15.00	17.00	19.00	46	0	0.024	0.028	0.067
	1.75	8	11.00	13.00	15.00	39	0	0.019	0.022	0.057
	2	5	8.00	10.00	12.00	33	0	0.015	0.017	0.048
	$\frac{2}{n-1}$	5	8.00	10.00	12.00	33	0	0.015	0.017	0.048
(NI)	—	4	4.00	4.00	4.00	5	0	0.013	0.013	0.017

Tabla A.6: Caso sin pesos, tomando $m = 50$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	29.00	31.00	34.50	113	0	0.088	0.099	0.323
(WIA)	1.25	16	21.00	23.00	25.50	89	0	0.066	0.074	0.254
	1.5	12	16.00	17.50	20.00	73	0	0.051	0.057	0.210
	1.75	8	12.00	13.00	15.50	62	0	0.037	0.045	0.177
	2	5	9.00	10.00	12.50	54	0	0.029	0.036	0.154
	$\frac{2}{n-1}$	5	9.00	10.00	12.50	54	0	0.029	0.037	0.155
(NI)	—	4	4.00	4.00	4.00	500	2	0.025	0.025	37.423

Tabla A.7: Caso con pesos, tomando $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	26.00	28.00	30.00	81	0	0.080	0.086	0.232
(WIA)	1.25	16	19.00	20.50	22.00	64	0	0.059	0.064	0.185
	1.5	11	14.00	16.00	17.00	53	0	0.046	0.049	0.151
	1.75	8	11.00	12.00	13.00	45	0	0.034	0.038	0.128
	2	5	8.00	9.00	10.00	39	0	0.026	0.029	0.113
	$\frac{2}{n-1}$	5	8.00	9.00	10.00	39	0	0.026	0.029	0.112
(NI)	—	3	4.00	4.00	4.00	4	0	0.025	0.025	0.026

Tabla A.8: Caso sin pesos, tomando $m = 100$ y $n = 2$.

A.2. Resultados para $n = 3$

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	26	53.00	78.50	191.00	500	6	0.009	0.021	0.057
(WIA)	1.25	20	41.00	61.50	152.00	500	4	0.007	0.017	0.057
	1.5	15	33.00	50.00	125.50	500	4	0.006	0.014	0.057
	1.75	12	27.50	42.00	107.00	500	4	0.005	0.012	0.057
	2	9	23.00	36.50	93.50	500	3	0.004	0.011	0.059
	$\frac{2}{n-1}$	26	53.00	78.50	191.00	500	6	0.009	0.021	0.056
(NI)	—	3	5.00	5.00	6.00	7	0	0.001	0.002	0.002

Tabla A.9: Caso con pesos, tomando $m = 3$ y $n = 3$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	27.00	32.00	59.50	500	1	0.004	0.007	0.055
(WIA)	1.25	16	21.00	24.00	47.00	500	1	0.003	0.005	0.055
	1.5	12	16.00	19.00	38.00	500	1	0.002	0.004	0.057
	1.75	9	13.00	15.50	32.00	500	1	0.002	0.004	0.056
	2	6	10.00	13.00	28.00	500	1	0.001	0.003	0.056
	$\frac{2}{n-1}$	22	27.00	32.00	59.50	500	1	0.004	0.007	0.055
(NI)	—	3	4.00	4.00	5.00	5	0	0.001	0.001	0.002

Tabla A.10: Caso sin pesos, tomando $m = 3$ y $n = 3$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	17	25.00	29.00	38.50	112	0	0.009	0.012	0.035
(WIA)	1.25	12	18.00	21.50	29.50	89	0	0.007	0.009	0.029
	1.5	9	14.00	17.00	23.50	73	0	0.005	0.007	0.024
	1.75	11	18.00	22.00	28.00	62	0	0.007	0.009	0.020
	2	18	33.00	48.50	80.00	500	1	0.015	0.025	0.155
	$\frac{2}{n-1}$	17	25.00	29.00	38.50	112	0	0.009	0.012	0.036
(NI)	—	4	4.00	4.00	4.00	5	0	0.003	0.003	0.004

Tabla A.11: Caso con pesos, tomando $m = 10$ y $n = 3$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	15	20.00	23.00	28.00	65	0	0.007	0.009	0.020
(WIA)	1.25	10	14.00	17.00	21.00	51	0	0.005	0.007	0.016
	1.5	6	11.00	13.00	16.00	42	0	0.004	0.005	0.013
	1.75	10	13.50	16.00	19.50	35	0	0.005	0.006	0.011
	2	15	23.00	30.00	42.50	119	0	0.010	0.014	0.037
	$\frac{2}{n-1}$	15	20.00	23.00	28.00	65	0	0.007	0.009	0.020
(NI)	—	4	4.00	4.00	4.00	5	0	0.003	0.003	0.005

Tabla A.12: Caso sin pesos, tomando $m = 10$ y $n = 3$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	17	20.00	21.50	23.00	57	0	0.032	0.033	0.083
(WIA)	1.25	12	14.00	15.00	16.00	45	0	0.022	0.024	0.065
	1.5	7	9.00	11.00	11.00	36	0	0.016	0.016	0.052
	1.75	10	13.00	14.00	16.00	30	0	0.021	0.023	0.043
	2	16	21.00	23.50	27.00	69	0	0.034	0.039	0.101
	$\frac{2}{n-1}$	17	20.00	21.50	23.00	57	0	0.032	0.033	0.082
(NI)	—	4	4.00	4.00	4.00	4	0	0.013	0.013	0.015

Tabla A.13: Caso con pesos, tomando $m = 50$ y $n = 3$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	15	18.00	19.00	20.00	26	0	0.027	0.029	0.038
(WIA)	1.25	10	12.00	13.00	14.00	19	0	0.019	0.020	0.027
	1.5	6	8.00	9.00	10.00	14	0	0.013	0.014	0.020
	1.75	9	11.00	12.00	14.00	16	0	0.019	0.020	0.023
	2	14	18.00	20.00	22.00	31	0	0.029	0.032	0.045
	$\frac{2}{n-1}$	15	18.00	19.00	20.00	26	0	0.027	0.029	0.037
(NI)	—	3	4.00	4.00	4.00	4	0	0.013	0.013	0.014

Tabla A.14: Caso sin pesos, tomando $m = 50$ y $n = 3$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	16	19.00	20.00	21.00	26	0	0.057	0.060	0.074
(WIA)	1.25	11	13.00	14.00	15.00	19	0	0.040	0.043	0.054
	1.5	6	8.00	9.00	10.00	15	0	0.026	0.029	0.043
	1.75	11	13.00	13.00	14.50	19	0	0.039	0.042	0.054
	2	17	20.00	21.00	24.00	33	0	0.063	0.069	0.094
	$\frac{2}{n-1}$	16	19.00	20.00	21.00	26	0	0.058	0.060	0.076
(NI)	—	4	4.00	4.00	4.00	500	8	0.025	0.026	37.459

Tabla A.15: Caso con pesos, tomando $m = 100$ y $n = 3$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	15	17.00	18.00	18.00	23	0	0.052	0.054	0.070
(WIA)	1.25	10	11.00	12.00	13.00	17	0	0.034	0.037	0.052
	1.5	6	7.00	8.00	9.00	13	0	0.023	0.026	0.040
	1.75	9	11.00	12.00	12.50	15	0	0.034	0.037	0.046
	2	15	17.00	18.00	20.00	25	0	0.053	0.057	0.073
	$\frac{2}{n-1}$	15	17.00	18.00	18.00	23	0	0.052	0.054	0.070
(NI)	—	3	4.00	4.00	4.00	500	1	0.025	0.026	32.554

Tabla A.16: Caso sin pesos, tomando $m = 100$ y $n = 3$.

A.3. Resultados para $n = 5$

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	23	53.00	92.00	212.50	500	10	0.010	0.024	0.058
(WIA)	1.25	17	41.00	72.50	169.00	500	10	0.008	0.019	0.057
	1.5	13	33.00	59.50	140.00	500	9	0.007	0.016	0.058
	1.75	10	27.00	50.00	119.50	500	7	0.006	0.013	0.059
	2	7	23.00	43.00	103.50	500	7	0.005	0.012	0.056
	$\frac{2}{n-1}$	52	112.00	189.00	429.50	500	20	0.022	0.049	0.059
(NI)	—	3	5.00	5.00	5.00	7	0	0.001	0.002	0.002

Tabla A.17: Caso con pesos, tomando $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	25.00	28.00	32.00	231	0	0.003	0.004	0.026
(WIA)	1.25	16	19.00	21.00	24.00	185	0	0.002	0.003	0.021
	1.5	12	15.00	16.50	19.00	153	0	0.002	0.002	0.017
	1.75	9	11.50	13.00	16.00	131	0	0.001	0.002	0.015
	2	6	9.00	11.00	13.00	116	0	0.001	0.001	0.013
	$\frac{2}{n-1}$	49	56.00	61.00	69.00	466	0	0.007	0.008	0.052
(NI)	—	4	4.00	5.00	5.00	5	0	0.001	0.002	0.002

Tabla A.18: Caso sin pesos, tomando $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	14	18.00	21.00	23.50	44	0	0.007	0.007	0.014
(WIA)	1.25	9	13.00	15.00	17.00	34	0	0.005	0.005	0.011
	1.5	12	16.00	17.00	18.00	27	0	0.005	0.006	0.008
	1.75	21	31.00	34.50	38.00	41	0	0.011	0.012	0.013
	2	58	129.50	210.50	339.00	500	18	0.065	0.106	0.161
	$\frac{2}{n-1}$	35	44.00	49.00	54.50	94	0	0.015	0.017	0.029
(NI)	—	4	4.00	4.00	4.00	5	0	0.003	0.003	0.004

Tabla A.19: Caso con pesos, tomando $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	12	15.00	17.00	18.50	31	0	0.005	0.006	0.010
(WIA)	1.25	8	10.00	12.00	13.00	24	0	0.004	0.004	0.007
	1.5	11	13.00	14.00	15.00	19	0	0.004	0.005	0.006
	1.75	21	25.00	28.00	31.00	37	0	0.009	0.010	0.013
	2	52	83.00	116.50	194.00	500	4	0.037	0.061	0.166
	$\frac{2}{n-1}$	32	37.00	40.00	43.50	66	0	0.013	0.014	0.022
(NI)	—	3	4.00	4.00	4.00	4	0	0.003	0.003	0.004

Tabla A.20: Caso sin pesos, tomando $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	12	14.00	15.00	16.00	19	0	0.022	0.023	0.028
(WIA)	1.25	7	8.00	9.00	10.00	13	0	0.013	0.014	0.019
	1.5	12	14.00	14.00	15.00	17	0	0.020	0.022	0.025
	1.75	21	24.00	25.00	27.00	33	0	0.036	0.040	0.051
	2	40	49.00	55.00	63.00	89	0	0.080	0.092	0.128
	$\frac{2}{n-1}$	33	37.50	39.00	41.00	47	0	0.057	0.059	0.068
(NI)	—	3	4.00	4.00	4.00	500	3	0.013	0.013	19.038

Tabla A.21: Caso con pesos, tomando $m = 50$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	11	13.00	13.00	14.00	16	0	0.020	0.021	0.025
(WIA)	1.25	6	8.00	8.00	9.00	11	0	0.012	0.013	0.016
	1.5	10	12.00	13.00	13.00	14	0	0.019	0.020	0.022
	1.75	18	20.00	22.00	23.00	26	0	0.032	0.035	0.040
	2	35	41.50	46.00	51.00	64	0	0.068	0.077	0.092
	$\frac{2}{n-1}$	31	34.00	35.00	36.00	40	0	0.052	0.054	0.060
(NI)	—	3	3.00	4.00	4.00	500	1	0.013	0.014	17.862

Tabla A.22: Caso sin pesos, tomando $m = 50$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	12	14.00	14.00	15.00	16	0	0.040	0.043	0.046
(WIA)	1.25	7	8.00	8.00	9.00	10	0	0.023	0.026	0.030
	1.5	12	14.00	14.00	15.00	16	0	0.040	0.043	0.046
	1.75	20	23.00	25.00	26.00	28	0	0.071	0.074	0.080
	2	39	46.00	49.50	54.00	62	0	0.142	0.155	0.179
	$\frac{2}{n-1}$	34	38.00	39.00	40.00	42	0	0.112	0.115	0.121
(NI)	—	3	4.00	4.00	4.00	500	11	0.025	0.026	37.362

Tabla A.23: Caso con pesos, tomando $m = 100$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	12	12.00	13.00	13.00	15	0	0.037	0.038	0.047
(WIA)	1.25	6	7.00	7.00	8.00	9	0	0.021	0.023	0.028
	1.5	11	12.00	12.00	13.00	14	0	0.035	0.037	0.042
	1.75	18	20.00	21.00	22.00	25	0	0.060	0.065	0.072
	2	34	38.00	41.00	45.00	52	0	0.118	0.129	0.154
	$\frac{2}{n-1}$	32	34.00	35.00	36.00	38	0	0.100	0.103	0.122
(NI)	—	3	3.00	3.00	4.00	500	2	0.019	0.025	37.871

Tabla A.24: Caso sin pesos, tomando $m = 100$ y $n = 5$.

A.4. Resultados para $n = 10$

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	24	47.50	76.50	212.50	500	8	0.009	0.024	0.058
(WIA)	1.25	18	36.50	60.50	168.50	500	6	0.007	0.019	0.058
	1.5	14	29.50	49.50	140.00	500	5	0.006	0.016	0.058
	1.75	10	24.00	41.50	119.00	500	4	0.005	0.013	0.056
	2	7	20.00	35.50	104.00	500	4	0.004	0.012	0.056
	$\frac{2}{n-1}$	130	234.00	365.00	500.00	500	37	0.041	0.056	0.058
(NI)	—	4	5.00	5.00	5.00	6	0	0.002	0.002	0.002

Tabla A.25: Caso con pesos, tomando $m = 3$ y $n = 10$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	25.00	27.00	29.50	57	0	0.003	0.003	0.007
(WIA)	1.25	16	19.00	20.00	22.50	45	0	0.002	0.003	0.005
	1.5	12	14.00	16.00	18.00	37	0	0.002	0.002	0.004
	1.75	9	11.00	12.00	14.00	31	0	0.001	0.002	0.004
	2	6	9.00	10.00	11.50	27	0	0.001	0.001	0.003
	$\frac{2}{n-1}$	119	131.00	138.00	150.00	273	0	0.016	0.017	0.031
(NI)	—	4	4.00	5.00	5.00	5	0	0.001	0.002	0.003

Tabla A.26: Caso sin pesos, tomando $m = 3$ y $n = 10$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	12	15.00	16.00	19.00	29	0	0.005	0.006	0.009
(WIA)	1.25	9	10.00	11.00	13.00	22	0	0.003	0.004	0.007
	1.5	15	17.00	18.00	18.00	20	0	0.006	0.006	0.007
	1.75	33	38.00	39.00	40.00	44	0	0.012	0.013	0.015
	2	500	500.00	500.00	500.00	500	100	0.156	0.157	0.160
	$\frac{2}{n-1}$	79	91.50	99.50	108.50	153	0	0.031	0.034	0.048
(NI)	—	3	4.00	4.00	4.00	4	0	0.003	0.003	0.004

Tabla A.27: Caso con pesos, tomando $m = 10$ y $n = 10$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	10	12.00	13.00	14.00	18	0	0.004	0.004	0.006
(WIA)	1.25	8	9.00	9.00	9.00	13	0	0.003	0.003	0.004
	1.5	14	15.00	16.00	16.00	17	0	0.005	0.005	0.006
	1.75	30	33.00	35.00	36.00	38	0	0.011	0.011	0.012
	2	351	500.00	500.00	500.00	500	95	0.155	0.156	0.161
	$\frac{2}{n-1}$	67	77.00	80.50	85.00	104	0	0.025	0.026	0.033
(NI)	—	3	3.00	4.00	4.00	4	0	0.003	0.003	0.003

Tabla A.28: Caso sin pesos, tomando $m = 10$ y $n = 10$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	10	10.00	11.00	11.00	13	0	0.016	0.017	0.019
(WIA)	1.25	9	10.00	10.00	10.00	11	0	0.014	0.015	0.016
	1.5	16	18.00	18.00	19.00	20	0	0.026	0.027	0.029
	1.75	33	36.00	37.00	39.00	42	0	0.054	0.056	0.061
	2	116	159.50	180.00	209.50	280	0	0.262	0.303	0.410
	$\frac{2}{n-1}$	75	79.00	82.00	84.00	89	0	0.119	0.122	0.131
(NI)	—	3	3.00	3.00	4.00	500	2	0.010	0.013	17.756

Tabla A.29: Caso con pesos, tomando $m = 50$ y $n = 10$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	9	9.00	10.00	10.00	11	0	0.014	0.015	0.016
(WIA)	1.25	8	9.00	9.00	9.00	9	0	0.013	0.013	0.014
	1.5	15	16.00	16.00	16.00	17	0	0.023	0.024	0.025
	1.75	29	30.50	32.00	32.00	34	0	0.046	0.047	0.050
	2	97	121.50	133.00	154.50	229	0	0.193	0.223	0.333
	$\frac{2}{n-1}$	68	71.00	73.00	74.00	77	0	0.106	0.108	0.113
(NI)	—	3	3.00	3.00	3.00	4	0	0.009	0.010	0.013

Tabla A.30: Caso sin pesos, tomando $m = 50$ y $n = 10$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	10	10.00	11.00	11.00	12	0	0.031	0.032	0.035
(WIA)	1.25	9	10.00	10.00	10.00	11	0	0.029	0.030	0.033
	1.5	17	18.00	19.00	19.00	20	0	0.054	0.055	0.058
	1.75	33	35.00	37.00	38.00	40	0	0.106	0.109	0.116
	2	98	123.50	136.00	143.00	172	0	0.391	0.412	0.496
	$\frac{2}{n-1}$	77	81.00	82.00	84.00	89	0	0.238	0.242	0.257
(NI)	—	3	3.00	3.00	4.00	500	1	0.019	0.025	32.633

Tabla A.31: Caso con pesos, tomando $m = 100$ y $n = 10$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	9	9.00	9.00	10.00	10	0	0.026	0.029	0.030
(WIA)	1.25	8	9.00	9.00	9.00	10	0	0.026	0.026	0.030
	1.5	15	16.00	17.00	17.00	18	0	0.049	0.049	0.051
	1.75	28	31.00	32.00	33.00	35	0	0.092	0.094	0.100
	2	80	100.00	107.00	114.00	146	0	0.308	0.329	0.419
	$\frac{2}{n-1}$	71	72.00	74.00	75.00	78	0	0.212	0.216	0.226
(NI)	—	3	3.00	3.00	3.00	4	0	0.019	0.019	0.025

Tabla A.32: Caso sin pesos, tomando $m = 100$ y $n = 10$.

Apéndice B

Tablas de resultados para el problema irrestricto en norma p

B.1. Resultados para $p = 1.1$

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	114	144	154.5	500.0	500	29	0.031	0.098	0.106
(WIA)	1.5	74	92	101.0	500.0	500	30	0.021	0.101	0.106
	2	62	72	107.0	500.0	500	30	0.022	0.101	0.103
	10.000	8	500	500.0	500.0	500	91	0.101	0.101	0.109
	20.000	40	500	500.0	500.0	500	90	0.100	0.101	0.109
	11.000	7	500	500.0	500.0	500	86	0.101	0.101	0.107

Tabla B.1: Caso con pesos, tomando $p = 1.1$, $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	127	500	500.0	500.0	500	99	0.100	0.102	0.110
(WIA)	1.5	79	500	500.0	500.0	500	99	0.104	0.105	0.112
	2	320	500	500.0	500.0	500	97	0.101	0.101	0.111
	10.000	323	500	500.0	500.0	500	94	0.100	0.101	0.106
	20.000	241	500	500.0	500.0	500	82	0.100	0.101	0.105
	11.000	308	500	500.0	500.0	500	83	0.100	0.101	0.107

Tabla B.2: Caso sin pesos, tomando $p = 1.1$, $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	133	148	172.0	489.0	500	25	0.058	0.164	0.173
(WIA)	1.5	81	96	110.5	406.0	500	24	0.038	0.138	0.175
	2	64	73	96.0	302.0	500	24	0.033	0.103	0.179
	10.000	11	500	500.0	500.0	500	96	0.169	0.169	0.178
	20.000	40	500	500.0	500.0	500	97	0.168	0.169	0.177
	11.000	9	500	500.0	500.0	500	94	0.169	0.169	0.178

Tabla B.3: Caso con pesos, tomando $p = 1.1$, $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	109	165	206.5	258.0	500	4	0.070	0.088	0.169
(WIA)	1.5	70	102	132.5	169.0	500	3	0.045	0.057	0.172
	2	57	74	97.0	125.5	500	2	0.033	0.043	0.171
	10.000	11	500	500.0	500.0	500	83	0.168	0.170	0.186
	20.000	11	500	500.0	500.0	500	90	0.168	0.169	0.182
	11.000	9	500	500.0	500.0	500	86	0.168	0.169	0.182

Tabla B.4: Caso sin pesos, tomando $p = 1.1$, $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	141	164	184.5	214.0	500	10	0.390	0.456	1.082
(WIA)	1.5	91	104	119.0	140.5	500	10	0.252	0.299	1.088
	2	66	78	90.0	103.0	500	10	0.190	0.220	1.080
	10.000	8	500	500.0	500.0	500	82	1.056	1.059	1.109
	20.000	500	500	500.0	500.0	500	100	1.037	1.039	1.087
	11.000	7	500	500.0	500.0	500	90	1.045	1.057	1.113

Tabla B.5: Caso con pesos, tomando $p = 1.1$, $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	113	144	166.5	183.5	500	9	0.352	0.395	1.143
(WIA)	1.5	80	92	105.0	120.0	500	9	0.228	0.255	1.136
	2	59	68	81.0	88.5	500	9	0.171	0.194	1.122
	10.000	8	500	500.0	500.0	500	76	1.058	1.063	1.110
	20.000	19	500	500.0	500.0	500	98	1.039	1.044	1.135
	11.000	7	500	500.0	500.0	500	80	1.056	1.060	1.145

Tabla B.6: Caso sin pesos, tomando $p = 1.1$, $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	134	148	194.0	500.0	500	41	0.055	0.138	0.148
(WIA)	1.5	86	96	128.0	500.0	500	40	0.037	0.140	0.148
	2	63	74	109.0	500.0	500	36	0.031	0.140	0.145
	4.000	30	500	500.0	500.0	500	93	0.140	0.141	0.145
	5.000	34	500	500.0	500.0	500	93	0.140	0.141	0.143
	4.400	26	500	500.0	500.0	500	94	0.140	0.141	0.145

Tabla B.7: Caso con pesos, tomando $p = 1.1$, $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	126	500	500.0	500.0	500	88	0.141	0.142	0.151
(WIA)	1.5	81	500	500.0	500.0	500	88	0.143	0.145	0.153
	2	95	500	500.0	500.0	500	89	0.140	0.141	0.150
	4.000	500	500	500.0	500.0	500	100	0.140	0.141	0.148
	5.000	500	500	500.0	500.0	500	100	0.140	0.141	0.150
	4.400	500	500	500.0	500.0	500	100	0.140	0.141	0.150

Tabla B.8: Caso sin pesos, tomando $p = 1.1$, $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	142	161	176.5	500.0	500	32	0.106	0.297	0.306
(WIA)	1.5	87	104	116.5	500.0	500	34	0.070	0.300	0.306
	2	66	78	90.5	500.0	500	31	0.055	0.300	0.305
	4.000	34	500	500.0	500.0	500	94	0.301	0.302	0.306
	5.000	27	500	500.0	500.0	500	98	0.301	0.302	0.305
	4.400	29	500	500.0	500.0	500	95	0.301	0.301	0.308

Tabla B.9: Caso con pesos, tomando $p = 1.1$, $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	126	138	146.0	155.0	274	0	0.088	0.094	0.164
(WIA)	1.5	79	88	95.0	101.0	183	0	0.057	0.062	0.114
	2	59	65	69.5	74.5	137	0	0.042	0.045	0.082
	4.000	28	35	500.0	500.0	500	59	0.300	0.301	0.321
	5.000	24	38	500.0	500.0	500	63	0.300	0.301	0.322
	4.400	26	32	500.0	500.0	500	65	0.300	0.302	0.318

Tabla B.10: Caso sin pesos, tomando $p = 1.1$, $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	154	165	169.0	500.0	500	29	0.798	2.343	2.376
(WIA)	1.5	97	107	110.0	500.0	500	28	0.521	2.344	2.386
	2	70	78	80.5	500.0	500	26	0.381	2.342	2.370
	4.000	32	500	500.0	500.0	500	77	2.348	2.355	2.458
	5.000	26	500	500.0	500.0	500	86	2.351	2.361	2.505
	4.400	29	500	500.0	500.0	500	82	2.351	2.356	2.500

Tabla B.11: Caso con pesos, tomando $p = 1.1$, $m = 100$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	135	145	152.0	161.5	500	23	0.717	0.764	2.380
(WIA)	1.5	85	94	98.0	105.0	500	23	0.463	0.498	2.453
	2	61	70	73.0	115.0	500	20	0.346	0.541	2.408
	4.000	29	500	500.0	500.0	500	85	2.354	2.363	2.496
	5.000	23	500	500.0	500.0	500	89	2.355	2.364	2.515
	4.400	26	500	500.0	500.0	500	88	2.355	2.366	2.499

Tabla B.12: Caso sin pesos, tomando $p = 1.1$, $m = 100$ y $n = 5$.

B.2. Problema en norma $p = 1.5$

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	24	46	70.5	126.0	500	6	0.014	0.025	0.100
(WIA)	1.5	15	29	44.5	81.0	500	2	0.009	0.016	0.103
	2	9	20	31.5	58.5	500	1	0.006	0.012	0.101
	2.000	9	20	31.5	58.5	500	1	0.006	0.012	0.101
	4.000	12	57	500.0	500.0	500	55	0.101	0.102	0.106
	3.000	8	17	33.0	500.0	500	29	0.007	0.101	0.105

Tabla B.13: Caso con pesos, tomando $p = 1.5$, $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	23	30	33.0	45.5	500	1	0.007	0.009	0.099
(WIA)	1.5	12	17	19.0	29.0	500	1	0.004	0.006	0.101
	2	8	11	13.0	21.0	500	1	0.003	0.004	0.103
	2.000	8	11	13.0	21.0	500	1	0.003	0.004	0.103
	4.000	31	500	500.0	500.0	500	91	0.101	0.102	0.107
	3.000	10	18	500.0	500.0	500	61	0.101	0.102	0.109

Tabla B.14: Caso sin pesos, tomando $p = 1.5$, $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	31	47	54.5	72.0	199	0	0.019	0.024	0.067
(WIA)	1.5	17	28	34.0	45.5	130	0	0.012	0.016	0.044
	2	11	19	23.0	32.0	96	0	0.008	0.011	0.033
	2.000	11	19	23.0	32.0	96	0	0.008	0.011	0.033
	4.000	9	28	79.0	500.0	500	46	0.027	0.171	0.175
	3.000	7	12	17.0	23.0	500	12	0.006	0.008	0.172

Tabla B.15: Caso con pesos, tomando $p = 1.5$, $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	25	40	49.5	67.0	262	0	0.017	0.023	0.089
(WIA)	1.5	14	24	31.0	42.0	173	0	0.011	0.014	0.059
	2	9	17	21.5	30.0	128	0	0.007	0.010	0.044
	2.000	9	17	21.5	30.0	128	0	0.007	0.010	0.044
	4.000	7	20	37.0	500.0	500	39	0.013	0.171	0.176
	3.000	6	10	13.0	22.5	500	12	0.004	0.008	0.173

Tabla B.16: Caso sin pesos, tomando $p = 1.5$, $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	35	44	49.0	54.0	500	1	0.104	0.116	1.052
(WIA)	1.5	20	26	29.0	33.0	500	1	0.063	0.070	1.052
	2	12	17	20.0	22.0	472	0	0.042	0.047	0.997
	2.000	12	17	20.0	22.0	472	0	0.042	0.047	0.997
	4.000	12	16	19.0	25.0	500	12	0.041	0.053	1.131
	3.000	5	8	9.0	11.0	500	1	0.019	0.024	1.054

Tabla B.17: Caso con pesos, tomando $p = 1.5$, $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	30	39	42.0	46.0	65	0	0.090	0.098	0.137
(WIA)	1.5	16	23	25.0	28.0	41	0	0.054	0.060	0.086
	2	11	15	17.0	18.5	28	0	0.036	0.039	0.059
	2.000	11	15	17.0	18.5	28	0	0.036	0.040	0.061
	4.000	11	14	18.0	23.5	500	8	0.038	0.050	1.074
	3.000	5	7	8.0	9.0	500	1	0.017	0.020	1.056

Tabla B.18: Caso sin pesos, tomando $p = 1.5$, $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	35	51	80.0	187.5	500	8	0.022	0.051	0.142
(WIA)	1.5	19	31	51.0	122.5	500	4	0.014	0.034	0.145
	2	14	22	42.5	90.5	500	4	0.012	0.026	0.141
	0.800	45	65	102.5	236.5	500	10	0.029	0.067	0.142
	1.000	35	51	80.0	187.5	500	8	0.022	0.053	0.144
	1.200	27	41	65.5	155.0	500	7	0.019	0.044	0.142

Tabla B.19: Caso con pesos, tomando $p = 1.5$, $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	22	29	34.5	41.5	113	0	0.010	0.011	0.032
(WIA)	1.5	12	17	21.0	25.0	72	0	0.006	0.007	0.021
	2	8	11	14.0	19.5	111	0	0.004	0.006	0.032
	0.800	29	38	45.0	54.0	143	0	0.013	0.015	0.040
	1.000	22	29	34.5	41.5	113	0	0.010	0.012	0.033
	1.200	18	23	27.5	33.0	93	0	0.008	0.010	0.026

Tabla B.20: Caso sin pesos, tomando $p = 1.5$, $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	27	33	36.0	40.0	58	0	0.022	0.024	0.034
(WIA)	1.5	14	19	21.0	24.0	36	0	0.013	0.014	0.022
	2	9	12	13.0	16.0	167	0	0.008	0.010	0.104
	0.800	36	43	47.5	52.0	74	0	0.029	0.031	0.044
	1.000	27	33	36.0	40.0	58	0	0.022	0.024	0.035
	1.200	21	26	29.0	32.0	47	0	0.017	0.019	0.028

Tabla B.21: Caso con pesos, tomando $p = 1.5$, $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	24	28	30.0	34.0	48	0	0.018	0.020	0.030
(WIA)	1.5	13	16	17.0	20.0	30	0	0.011	0.012	0.019
	2	8	10	11.0	13.0	92	0	0.007	0.008	0.057
	0.800	31	37	39.0	44.5	62	0	0.024	0.027	0.037
	1.000	24	28	30.0	34.0	48	0	0.018	0.020	0.029
	1.200	18	22	24.0	27.0	39	0	0.014	0.016	0.024

Tabla B.22: Caso sin pesos, tomando $p = 1.5$, $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	28	30	31.0	32.0	34	0	0.147	0.151	0.161
(WIA)	1.5	16	17	17.0	18.0	19	0	0.081	0.085	0.090
	2	8	9	10.0	10.0	12	0	0.047	0.048	0.056
	0.800	38	41	41.5	42.0	45	0	0.197	0.199	0.213
	1.000	28	30	31.0	32.0	34	0	0.147	0.151	0.161
	1.200	22	24	24.0	25.0	27	0	0.115	0.118	0.127

Tabla B.23: Caso con pesos, tomando $p = 1.5$, $m = 100$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	25	27	28.0	28.0	30	0	0.131	0.133	0.141
(WIA)	1.5	14	15	15.0	16.0	17	0	0.072	0.075	0.081
	2	7	8	9.0	9.0	10	0	0.042	0.043	0.047
	0.800	33	36	37.0	38.0	39	0	0.174	0.178	0.184
	1.000	25	27	28.0	28.0	30	0	0.131	0.133	0.141
	1.200	19	21	22.0	22.0	23	0	0.103	0.104	0.111

Tabla B.24: Caso sin pesos, tomando $p = 1.5$, $m = 100$ y $n = 5$.

B.3. Problema en norma $p = 2.5$

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	13	33	59.0	138.0	500	3	0.012	0.028	0.102
(WIA)	1.5	7	24	61.5	500.0	500	29	0.012	0.101	0.103
	2	10	500	500.0	500.0	500	83	0.101	0.102	0.107
	0.667	24	53	92.0	211.0	500	5	0.019	0.043	0.104
	1.333	8	23	42.5	104.0	500	2	0.009	0.021	0.101
	1.667	9	40	500.0	500.0	500	54	0.101	0.101	0.104

Tabla B.25: Caso con pesos, tomando $p = 2.5$, $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	12	19	23.5	43.0	500	2	0.005	0.008	0.101
(WIA)	1.5	6	12	18.0	37.5	500	10	0.004	0.008	0.103
	2	15	37	224.0	500.0	500	47	0.045	0.101	0.107
	0.667	22	32	39.0	67.5	500	2	0.008	0.014	0.103
	1.333	7	12	16.0	30.0	500	2	0.003	0.006	0.106
	1.667	8	14	26.5	107.0	500	20	0.005	0.022	0.106

Tabla B.26: Caso sin pesos, tomando $p = 2.5$, $m = 3$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	16	26	33.5	53.5	331	0	0.012	0.018	0.112
(WIA)	1.5	7	16	22.0	37.5	500	3	0.007	0.013	0.176
	2	7	30	117.0	500.0	500	46	0.040	0.171	0.174
	0.667	27	43	54.0	84.5	489	0	0.019	0.029	0.167
	1.333	9	17	23.0	38.0	247	0	0.008	0.013	0.084
	1.667	6	16	26.0	71.5	500	16	0.009	0.024	0.173

Tabla B.27: Caso con pesos, tomando $p = 2.5$, $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	13	22	30.0	43.0	192	0	0.010	0.015	0.066
(WIA)	1.5	5	13	18.0	28.0	500	3	0.006	0.009	0.178
	2	9	27	92.5	500.0	500	31	0.031	0.170	0.181
	0.667	24	36	48.5	68.0	292	0	0.017	0.023	0.100
	1.333	7	14	20.0	31.0	142	0	0.007	0.011	0.048
	1.667	6	14	23.5	39.0	500	8	0.008	0.013	0.178

Tabla B.28: Caso sin pesos, tomando $p = 2.5$, $m = 10$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	14	21	23.0	26.0	78	0	0.048	0.055	0.164
(WIA)	1.5	7	10	11.0	14.0	97	0	0.024	0.030	0.204
	2	11	16	19.0	27.5	500	3	0.041	0.058	1.061
	0.667	27	36	39.0	44.0	120	0	0.083	0.093	0.253
	1.333	9	13	14.0	17.0	57	0	0.030	0.036	0.122
	1.667	4	9	11.0	14.0	500	1	0.023	0.030	1.055

Tabla B.29: Caso con pesos, tomando $p = 2.5$, $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	14	18	20.0	21.5	95	0	0.042	0.046	0.202
(WIA)	1.5	6	8	10.0	11.0	500	1	0.021	0.024	1.056
	2	9	14	17.0	24.0	500	4	0.036	0.051	1.079
	0.667	25	32	34.0	37.0	146	0	0.072	0.078	0.307
	1.333	8	11	12.5	14.0	69	0	0.027	0.030	0.145
	1.667	4	8	9.0	12.0	500	2	0.019	0.025	1.061

Tabla B.30: Caso sin pesos, tomando $p = 2.5$, $m = 100$ y $n = 2$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	19	37	73.0	168.0	500	7	0.020	0.046	0.140
(WIA)	1.5	500	500	500.0	500.0	500	100	0.141	0.141	0.144
	2	500	500	500.0	500.0	500	100	0.140	0.141	0.147
	0.267	82	156	292.0	500.0	500	32	0.082	0.140	0.145
	0.333	65	124	232.5	500.0	500	26	0.065	0.140	0.143
	0.667	29	59	113.0	255.0	500	9	0.032	0.072	0.142

Tabla B.31: Caso con pesos, tomando $p = 2.5$, $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	16	18	19.5	27.5	500	1	0.005	0.008	0.140
(WIA)	1.5	500	500	500.0	500.0	500	100	0.141	0.141	0.144
	2	500	500	500.0	500.0	500	100	0.140	0.141	0.144
	0.267	54	70	90.0	122.5	500	2	0.025	0.035	0.140
	0.333	43	56	71.0	97.0	500	2	0.020	0.027	0.141
	0.667	19	24	32.0	45.0	500	2	0.009	0.013	0.141

Tabla B.32: Caso sin pesos, tomando $p = 2.5$, $m = 3$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	12	16	18.0	19.0	49	0	0.011	0.012	0.029
(WIA)	1.5	500	500	500.0	500.0	500	100	0.301	0.302	0.308
	2	500	500	500.0	500.0	500	100	0.301	0.302	0.312
	0.267	42	63	69.5	86.0	208	0	0.042	0.052	0.125
	0.333	33	49	54.0	67.5	164	0	0.032	0.041	0.099
	0.667	13	20	23.0	30.0	78	0	0.014	0.018	0.048

Tabla B.33: Caso con pesos, tomando $p = 2.5$, $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	10	13	15.0	16.0	27	0	0.009	0.010	0.016
(WIA)	1.5	500	500	500.0	500.0	500	100	0.301	0.302	0.306
	2	500	500	500.0	500.0	500	100	0.301	0.302	0.305
	0.267	43	51	56.0	62.5	121	0	0.034	0.038	0.073
	0.333	33	40	43.0	48.5	96	0	0.026	0.029	0.058
	0.667	13	16	18.0	21.0	44	0	0.011	0.013	0.026

Tabla B.34: Caso sin pesos, tomando $p = 2.5$, $m = 10$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	—	8	10	10.0	11.0	13	0	0.048	0.052	0.061
(WIA)	1.5	52	72	81.0	95.0	162	0	0.384	0.449	0.767
	2	500	500	500.0	500.0	500	100	2.353	2.357	2.437
	0.267	43	49	51.0	52.5	58	0	0.240	0.250	0.274
	0.333	33	38	39.0	40.0	45	0	0.184	0.191	0.212
	0.667	12	14	15.0	15.0	17	0	0.070	0.074	0.080

Tabla B.35: Caso con pesos, tomando $p = 2.5$, $m = 100$ y $n = 5$.

Método	λ	Cantidad de Iteraciones						Tiempo en segundos		
		k_{min}	$k_{0.25}$	$k_{0.5}$	$k_{0.75}$	k_{max}	F_{500}	$t_{0.5}$	$t_{0.75}$	t_{max}
(WI)	–	7	8	9.0	10.0	11	0	0.042	0.047	0.052
(WIA)	1.5	42	56	66.0	75.5	137	0	0.310	0.359	0.644
	2	500	500	500.0	500.0	500	100	2.352	2.356	2.467
	0.267	37	43	45.0	46.0	49	0	0.211	0.216	0.233
	0.333	28	33	34.0	35.0	37	0	0.160	0.165	0.178
	0.667	10	12	13.0	13.0	14	0	0.061	0.061	0.067

Tabla B.36: Caso sin pesos, tomando $p = 2.5$, $m = 100$ y $n = 5$.

Bibliografía

- [BB71] Edwin F. Beckenbach and Richard Bellman. *Inequalities / E.F. Beckenbach, R. Bellman*. Springer-Verlag Berlin [Germany : West], 1971.
- [BCC98] Jack Brimberg, Reuven Chen, and Doron Chen. Accelerating convergence in the Fermat-Weber location problem. *Oper. Res. Lett.*, 22(4-5):151–157, 1998.
- [BL92] Jack Brimberg and Robert F. Love. Local convergence in a generalized Fermat-Weber problem. *Annals OR*, 40(1):33–66, 1992.
- [BL93] Jack Brimberg and Robert F. Love. Global Convergence of a Generalized Iterative Procedure for the Minisum Location Problem with lp Distances. *Operations Research*, 41(6):1153–1163, 1993.
- [BL99] J. Brimberg and R.F. Love. Local convexity results in a generalized fermat-weber problem. *Computers and Mathematics with Applications*, 37(8):87 – 97, 1999.
- [BMR00] Ernesto G. Birgin, José Mario Martínez, and Marcos Raydan. Nonmonotone Spectral Projected Gradient Methods on Convex Sets. *SIAM Journal on Optimization*, 10(4):1196–1211, 2000.
- [BMR14] Ernesto G. Birgin, José Mario Martínez, and Marcos Raydan. Spectral projected gradient methods: Review and perspectives. *Journal of Statistical Software*, 60(3):1–21, 2014.
- [Bri95] Jack Brimberg. The Fermat-Weber location problem revisited. *Math. Program.*, 71:71–76, 1995.
- [Bri03] Jack Brimberg. Further Notes on Convergence of the Weiszfeld Algorithm. *Yugoslav Journal of Operations Research*, 13(2):199–206, 2003.
- [BS15] Amir Beck and Shoham Sabach. Weiszfeld’s Method: Old and New Results. *J. Optim. Theory Appl.*, 164(1):1–40, January 2015.
- [Buz13] Gustavo Daniel Buzai. Location-allocation models applied to urban public services. spatial analysis of primary health care centers in the city of luján, argentina. *Hungarian Geographical Bulletin*, 62(4):387–408, 2013.
- [CCM02] Lázaro Cánovas, Roberto Cañavate, and Alfredo Marín. On the convergence of the Weiszfeld algorithm. *Math. Program.*, 93(2):327–330, 2002.
- [CT89] Ramaswamy Chandrasekaran and Arie Tamir. Open questions concerning Weiszfeld’s algorithm for the Fermat-Weber location problem. *Math. Program.*, 44(1-3):293–295, 1989.
- [Dre92] Zvi Drezner. A note on the weber location problem. *Annals of Operations Research*, 40(1):153–161, Dec 1992.
- [Dre95] Zvi Drezner. A note on accelerating de Weiszfeld procedure. *Locarion Science*, 3(4-5):275–279, 1995.

- [GK16] Simone Görner and Christian Kanzow. On Newton's Method for the Fermat-Weber Location Problem. *J. Optimization Theory and Applications*, 170(1):107–118, 2016.
- [Gue07] Guerrero García, Carlos. *Un modelo multicriterio de localización de centros no deseados con pesos*. Junta de Andalucía, Consejo Económico y Social de Andalucía, 2007.
- [JL81] Henrik Juel and Robert F. Love. Fixed Point Optimality Criteria for the Location Problem with Arbitrary Norms. *Journal of the Operational Research Society*, 32(10):891–897, Oct 1981.
- [Kat74] I. Norman Katz. Local convergence in Fermat's problem. *Math. Program.*, 6(1):89–104, 1974.
- [KK62] Harold W. Kuhn and Robert E. Kuenne. An efficient algorithm for the numerical solution of the generalized weber problem in spatial economics. *Journal of Regional Science*, 4(2):21–33, 1962.
- [KSD10] Dongmin Kim, Suvrit Sra, and Inderjit S. Dhillon. Tackling Box-Constrained Optimization via a New Projected Quasi-Newton Approach. *SIAM J. Scientific Computing*, 32(6):3548–3563, 2010.
- [Kuh73] Harold W. Kuhn. A note on Fermat's problem. *Math. Program.*, 4(1):98–107, 1973.
- [Kwa11] Samuel Kwarteng. Location of a library in sunyani municipality using the planar k-centra single-facility euclidean location algorithm. Master's thesis, Kwame Nkrumah University of Science and Technology, 2011.
- [LM72] R. F. Love and J. G. Morris. Modelling Inter-City Road Distances by Mathematical Functions. *Operational Research Quarterly (1970-1977)*, 23(1):61–71, 1972.
- [LM79] Robert F. Love and James G. Morris. Mathematical Models of Road Travel Distances. *Manage. Sci.*, 25(2):130–139, February 1979.
- [LMW88] Robert F. Love, James G. Morris, and George O. Wesolowsky. *Facilities location*. Number 7 in Publications in operations research series. North-Holland, New York [u.a.], 1988.
- [OJ78] Lawrence M. Ostresh Jr. On the Convergence of a Class of Iterative Methods for Solving the Weber Location Problem. *Operations Research*, 26(4):597–609, 1978.
- [OR70] J.M. Ortega and W.C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1970.
- [PT11] Elvio A. Pilotta and Germán Ariel Torres. A projected Weiszfeld algorithm for the box-constrained Weber location problem. *Applied Mathematics and Computation*, 218(6):2932–2943, 2011.
- [RBS01] Liliana Ramírez and Joaquín Bosque Sendra. Localización de hospitales: analogías y diferencias del uso del modelo p-mediano en sig raster y vectorial. In *Anales de geografía de la Universidad Complutense*, volume 21, pages 53–79, 2001.
- [VZ01] Yehuda Vardi and Cun-Hui Zhang. A modified Weiszfeld algorithm for the Fermat-Weber location problem. *Math. Program.*, 90(3):559–566, 2001.
- [Wei37] Endre Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tôhoku Mathematical Journal*, 43:355–386, 1937.
- [WF29] Alfred Weber and Carl J. Friedrich. *Alfred Weber's theory of the location of industries*. Materials for the study of business. The University of Chicago Press, 1929.

Los abajo firmantes, miembros del Tribunal de Evaluación de tesis, damos Fe que el presente ejemplar impreso, se corresponde con el aprobado por éste Tribunal.

Dr. Elvio A. Pilotta Dra. Aldana M. González Montoro Dr. Damián R. Fernández Ferreyra