ALGUNAS CUESTIONES TEÓRICAS SOBRE LA VALIDEZ DEL ALGORITMO DE WEISZFELD PARA EL PROBLEMA DE WEBER

Roberto Javier Cañavate Bernal
Universidad Politécnica de Cartagena

RESUMEN

El problema económico de Weber se describe matemáticamente como la localización de un punto que minimice la suma ponderada de las distancias a una serie de puntos dados. El algoritmo de Weiszfeld es el método más utilizado para la resolución del problema de Weber a pesar de que puede existir un conjunto de puntos iniciales, denominado el conjunto de Kuhn, para los que el algoritmo no converja al óptimo. Chandrasekaran y Tamir conjeturaron que si los puntos dados no están contenidos en ningún hiperplano del espacio entonces el conjunto de Kuhn es numerable. Dicha conjetura fue probada por Brimberg en 1995, pero la utilización de argumentos incorrectos en la demostración reabrió de nuevo la conjetura de Chandrasekaran y Tamir. En este trabajo se muestra que, en las condiciones de Chandrasekaran y Tamir, el conjunto de Kuhn es de medida cero. Este resultado permite establecer la validez teórica del algoritmo de Weiszfeld (el objetivo de la conjetura de Chandrasekaran y Tamir) puesto que demuestra que es prácticamente nula la probabilidad de que al elegir al azar un punto inicial para el algoritmo de Weiszfeld, éste no converja al óptimo del problema de Weber.

1. INTRODUCCIÓN

A principios del siglo XX, en su influyente libro sobre localización de industrias (véase [Weber, 1909]), el economista Alfred Weber formuló el siguiente problema: *Se necesita establecer la ubicación de una fábrica que se abastecerá de materia prima de dos almacenes, y cuyo producto se venderá en un cierto mercado. La distancia desde la fábrica a los almacenes y al mercado supone un coste para la fábrica que puede ser estimado y que se considera proporcional a dicha distancia. ¿Cuál es la localización que genera un menor coste económico para la empresa?* Este problema, cuya resolución requiere la búsqueda de un punto que minimice la suma de las distancias ponderadas (por el factor de coste unitario de las distancias) a los tres puntos conocidos (los dos almacenes y el mercado), y al que se denomina *el problema de Weber*, es el origen de una gran corriente de aplicaciones prácticas, adaptaciones y generalizaciones que perduran hasta nuestros días, y que lo convierten en un campo especialmente fructífero. Es importante observar que aunque el problema original de Weber fue formulado en el ámbito de la localización de industrias para tres puntos conocidos, las aplicaciones prácticas conllevan normalmente una gran cantidad de puntos, y abarcan campos tan diversos como la ergonomía o la robótica.

En la literatura se pueden encontrar varios procedimientos para resolver el problema de Weber, aunque el que ha tenido más éxito es el denominado algoritmo de Weiszfeld. Sin embargo, es conocido que en determinados casos existen problemas en la convergencia de dicho algoritmo al óptimo. Su amplia utilización práctica ha favorecido a lo largo de los últimos años estudios acerca de la validez teórica del algoritmo, pero sin resultados definitivos hasta el momento.

En el presente trabajo se mostrará un resultado que apoya la validez del algoritmo de Weiszfeld, respaldando así su utilización práctica, al demostrar que si se elige al azar un punto para comenzar el algoritmo la probabilidad de que éste no conduzca al óptimo del problema de Weber es nula.

2. EL PROBLEMA DE WEBER Y EL ALGORITMO DE WEISZFELD

El *problema de Weber* se puede formular matemáticamente del siguiente modo: dada una cantidad finita de puntos distintos $a_1,\ldots,a_m$ de $\Re^n$, con pesos asociados $w_i>0$, hallar el punto de $\Re^n$ que minimiza la función

$$W(x)=\sum_{i=1}^{m} w_i \left\| x-a_i \right\| \qquad \forall x\in\Re^n,$$

donde $\left\| x-a_i \right\|$ es la distancia euclídea entre el punto x y el punto a_i, y w_i es el coste o ponderación que corresponde al punto a_i. Se denominan *vértices* a los puntos $a_1,\ldots,a_m$, *pesos* a los valores reales positivos $w_1,\ldots,w_m$, y *función de Weber* a la aplicación W anterior.

El problema es trivial si los vértices están contenidos en una recta de $\Re^n$ (véase [Cañavate, 2001]), por lo que en adelante se supondrá en todos los casos que los puntos $a_1,\ldots,a_m$ no están alineados. Bajo esta suposición la función de Weber es estrictamente convexa, con lo cual el problema de Weber posee un único óptimo. Además, es conocido que dicho óptimo se encuentra en la envolvente convexa de los vértices (véase [Wendell y Hurter, 1973]).

Mediante un sencillo procedimiento es posible examinar la optimalidad de cada uno de los vértices (véase [Love et. al., 1988]), por lo que centraremos nuestra atención en el resto de puntos de $\Re^n$. Dado $x\neq a_i$, $\forall i=1,\ldots,m,$ puesto que W es de clase C^1 y estrictamente convexa en el conjunto $\Re^n-\{a_1,\ldots,a_m\}$, se puede afirmar que x es el óptimo del problema de Weber si, y sólo si, anula el gradiente de W:

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

Despejando parcialmente el punto x se obtiene la expresión

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

a partir de la cual (véase [Weiszfeld, 1937]) se propuso un método de resolución basado en la creación de una sucesión iterativa de puntos que converge al óptimo del problema de Weber.

Definición. Se llama *función de Weiszfeld* a la aplicación $T : \Re^n \longrightarrow \Re^n$ definida como

$$T(x) = \begin{cases} \dfrac{\sum_{i=1}^{m} \frac{w_i}{\|x - a_i\|} \cdot a_i}{\sum_{j=1}^{m} \frac{w_j}{\|x - a_j\|}} & si\ x \neq a_1, \ldots, a_m \\ a_i & si\ x = a_i\ \ para\ cierto\ i = 1, \ldots, m \end{cases}$$

Definición. Dado un punto $x_0 \in \Re^n$, se denomina *sucesión de Weiszfeld* a la sucesión de puntos $(x_r)_{r \in N}$ definida por

$$x_r = T(x_{r-1}) \quad \forall r \in N, r \geq 1.$$

El resultado siguiente pone de manifiesto la relación entre la sucesión de Weiszfeld y el óptimo del problema de Weber.

Teorema. Dada una sucesión de Weiszfeld:

1. Si la sucesión de Weiszfeld no contiene ningún vértice, entonces converge al óptimo del problema de Weber. En particular, si dos términos consecutivos de ésta coinciden, $x_r = x_{r+1}$, entonces x_r es el óptimo del problema de Weber.
2. Si un término de la sucesión de Weiszfeld es un vértice del problema de Weber, $x_r = a_k$, entonces la sucesión se estabiliza en dicho vértice, esto es, $a_k = x_r = x_{r+1} = \Lambda$.

El *algoritmo de Weiszfeld* consiste en la generación de una sucesión de Weiszfeld y la aplicación del teorema anterior. Sin embargo, este resultado sólo permite asegurar la convergencia al óptimo cuando la sucesión de Weiszfeld creada no contiene vértices del problema. Esta salvedad es muy importante ya que si a_j es un vértice no óptimo y ocurriera que $x_r=a_j$, entonces la sucesión de Weiszfeld convergería a a_j, un punto diferente al óptimo.

Este detalle fundamental, ignorado por Weiszfeld, fue puesto de manifiesto en primera instancia por Kuhn (véase [Kuhn, 1973]), mediante un ejemplo en el que se producía este hecho en una sucesión de Weiszfeld generada a partir de un punto x_0 distinto de los vértices.

3. EL CONJUNTO DE KUHN

Tras el error inicial de Weiszfeld y el ejemplo posterior de Kuhn, se inició un debate sobre la *cantidad* de puntos x_0 iniciales para los cuales el algoritmo de Weiszfeld no converge al óptimo, esto es, sobre la validez del algoritmo de Weiszfeld. En el trabajo citado, Kuhn presentó un teorema el que se intentaba dar respuesta a la cuestión sobre el *tamaño* del conjunto de puntos iniciales para los que el algoritmo no converge al óptimo.

Definición. Dada un problema de Weber, se denomina *conjunto de Kuhn* y se denota por *K(T)*, al conjunto de puntos iniciales que originan una sucesión de Weiszfeld que contiene algún vértice:

$$K(T) = \bigcup_{i=1}^{m} \left(\bigcup_{r=1}^{+\infty} T^{-r}(a_i) \right).$$

Los puntos pertenecientes al conjunto de Kuhn generan sucesiones de Weiszfeld que contienen algún vértice del problema de Weber, de modo que no es posible asegurar la convergencia del algoritmo de Weiszfeld al óptimo del problema. En otras palabras, *K(T)* es el conjunto de *malos* puntos iniciales para el algoritmo de Weiszfeld.

Teorema (Kuhn). El conjunto de Kuhn es numerable.

Con este resultado se puede afirmar que si se elige al azar un punto inicial para originar la sucesión de Weiszfeld, la probabilidad de que dicho punto pertenezca al conjunto de Kuhn, esto es, que no pueda garantizarse su convergencia al óptimo del problema, es nula.

El teorema de Kuhn tuvo que ser rechazado a partir de los contraejemplos presentados por Chandrasekaran y Tamir (véase [Chandrasekaran y Tamir, 1989]). Sin embargo, éstos conjeturaron que si se añadía la condición de que los vértices no estuvieran contenidos en ningún hiperplano de $\Re^n$ entonces el teorema de Kuhn sería válido.

Conjetura (Chandrasekaran y Tamir, 1989). Si los vértices del problema de Weber no están contenidos en ningún hiperplano de $\Re^n$, el conjunto de Kuhn es numerable.

Algunos años más tarde se publicó un trabajo (véase [Brimberg, 1995]) en el que se aseguraba haber demostrado la validez de la conjetura de Chandrasekaran y Tamir, dando así por terminada la controversia sobre la convergencia del algoritmo de Weiszfeld. De hecho, el resultado principal del documento mejoraba la conjetura de Chandrasekaran y Tamir al afirmar que la condición de que los vértices no estén contenidos en ningún hiperplano, esto es, la hipótesis de la conjetura, supone una caracterización de la numerabilidad del conjunto de Kuhn.

Teorema (Brimberg, 1995). El conjunto de Kuhn es numerable si, y sólo si, el conjunto de vértices no está contenido en ningún hiperplano de $\Re^n$.

Investigaciones posteriores (véase [Cánovas et. al., 2002]) evidenciaron la existencia de importantes errores en el teorema de Brimberg: una de las dos implicaciones es falsa y la otra, la que coincide con la conjetura de Chandrasekaran y Tamir, había sido demostrada con argumentos erróneos. De este modo quedaba nuevamente abierta la conjetura y, lo que es más importante, la cuestión acerca del tamaño del conjunto de Kuhn.

4. EL TAMAÑO DEL CONJUNTO KUHN

La conclusión más importante que se puede extraer de la sección previa es que aún después del esfuerzo realizado por diferentes investigadores para alcanzar teóricamente un resultado que respalde la validez del algoritmo de Weiszfeld, sigue siendo necesaria la acotación del tamaño del conjunto de Kuhn para avalar la utilización empírica que se realiza del algoritmo desde hace años. En esta sección abordaremos esa tarea, sólo que en lugar de hacerlo a través de la conjetura de Chandrasekaran y Tamir se llevará a cabo desde una perspectiva más general, mostrando que el conjunto de Kuhn es un conjunto de medida cero, es decir, *pequeño*.

Para simplificar la escritura de los razonamientos posteriores estableceremos la siguiente notación:

$$R(f) = \left\{ x \in \Re^n \;/\; f \text{ es de clase } C^1 \text{ en } x \text{ y } |Jf(x)| \neq 0 \right\}$$
$$D(f) = \left\{ x \in \Re^n \;/\; f \text{ es de clase } C^1 \text{ en } x \text{ y } |Jf(x)| = 0 \right\}$$
$$C(f) = \left\{ x \in \Re^n \;/\; f \text{ no es de clase } C^1 \text{ en } x \right\}$$

donde $|Jf(x)|$ es el determinante de la matriz jacobiana de *f* en *x*.

En el caso específico de la función de Weiszfeld (*T*), de sus características se deduce inmediatamente que $C(T) \subseteq \{a_1, ..., a_m\}$, con lo que *C(T)* es un conjunto finito. Además, está demostrado (véase [Brimberg, 1995]) que *D(T)* es un conjunto de medida cero.

La conclusión principal de este trabajo se obtiene a partir del estudio del conjunto de antiimágenes de los vértices mediante la función de Weiszfeld, para lo cual se empleará el siguiente resultado (véase [Cañavate, 2001]).

Teorema. Sea A un conjunto abierto de $\Re^n$ y sea $f : A \subseteq \Re^n \longrightarrow \Re^n$ una función de clase C^1 en A. Si $M \subseteq \Re^n$ es un conjunto de medida cero, entonces $f^{-1}(M) \cap R(f)$ también tiene medida cero.

Teorema. En las condiciones de la conjetura de Chandrasekaran y Tamir, el conjunto de Kuhn tiene medida cero.

Demostración: El conjunto de Kuhn de un cierto problema de Weber viene dado por

$$K(T) = \bigcup_{i=1}^{m} \left(\bigcup_{r=1}^{+\infty} T^{-r}(a_i) \right).$$

Para cada índice $i \in \{1,...,m\}$ probaremos a continuación, mediante inducción en el índice *r*, que $T^{-r}(a_i)$ es un conjunto de medida cero. Esto finalizará la demostración ya que en ese caso $K(T)$ es la unión finita de uniones numerables de conjuntos de medida cero, que de nuevo tiene medida cero.

- **Caso $r = 1$:** Puesto que $\Re^n = R(T) \cup D(T) \cup C(T)$, entonces

  $$T^{-1}(a_i) = \left[T^{-1}(a_i) \cap R(T)\right] \cup \left[T^{-1}(a_i) \cap D(T)\right] \cup \left[T^{-1}(a_i) \cap C(T)\right],$$

  y es evidente que $T^{-1}(a_i) \cap D(T)$ es un conjunto de medida cero y $T^{-1}(a_i) \cap C(T)$ un conjunto finito.

  En el caso del conjunto $T^{-1}(a_i) \cap R(T)$, puesto que $M = \{a_i\}$ es un conjunto de medida cero, parece natural la utilización del teorema anterior. Sin embargo, no es posible su aplicación directa dado que la función de Weiszfeld no es de clase C^1 en todo $\Re^n$. Ahora, *T* sí es una función de clase C^1 en $\Re^n - \{a_1,...,a_m\}$, de modo que denominando $\widetilde{T}$ a la restricción de *T* a $\Re^n - \{a_1,...,a_m\}$ y aplicando el teorema se obtiene que $\widetilde{T}^{-1}(a_i) \cap R(\widetilde{T})$ es un conjunto de medida cero. Por último, puesto que la máxima diferencia que puede existir entre los conjuntos $T^{-1}(a_i)$ y $\widetilde{T}^{-1}(a_i)$ es el conjunto $\{a_1,...,a_m\}$, debemos concluir que también $T^{-1}(a_i) \cap R(T)$ es un conjunto de medida cero.

  De este desarrollo se deduce que $T^{-1}(a_i)$ es un conjunto de medida cero por ser la unión de un conjunto finito y dos conjuntos de medida cero.

- **Caso general**: Sea $r > 1$. La hipótesis de inducción afirma que $T^{-r}(a_i)$ es un conjunto de medida cero y resta por demostrar que $T^{-(r+1)}(a_i)$ también lo es. Repitiendo el proceso, se tiene que

$$T^{-(r+1)}(a_i) = \left[T^{-(r+1)}(a_i) \cap R(T)\right] \cup \left[T^{-(r+1)}(a_i) \cap D(T)\right] \cup \left[T^{-(r+1)}(a_i) \cap C(T)\right],$$

y es evidente que $T^{-(r+1)}(a_i) \cap D(T)$ es un conjunto de medida cero y $T^{-(r+1)}(a_i) \cap C(T)$ un conjunto finito.

Pero dado que $T^{-(r+1)}(a_i) = T^{-1}\left[T^{-r}(a_i)\right]$ y $T^{-r}(a_i)$ es un conjunto de medida cero, se pueden repetir para $M = T^{-r}(a_i)$ los argumentos realizados en el caso anterior para el conjunto $M = \{a_i\}$, con la conclusión de que $T^{-(r+1)}(a_i) \cap R(T)$ tiene medida cero.

Finalmente se concluye que $T^{-(r+1)}(a_i)$ es un conjunto de medida cero por ser la unión de un conjunto finito y dos de medida nula.

5. CONCLUSIONES

En este trabajo se ha mostrado que la validez teórica del algoritmo de Weiszfeld permanece intacta, ya que el conjunto de puntos iniciales para los que falla (el denominado conjunto de Kuhn) tiene medida cero, con lo cual la probabilidad de que el punto inicial escogido para el algoritmo de Weiszfeld pertenezca a este conjunto es nula. Por tanto, aunque la conjetura de Chandrasekaran y Tamir permanece abierta, el objetivo de ésta, es decir, el sostenimiento teórico del algoritmo de Weiszfeld, queda cubierto por los resultados del trabajo.

6. REFERENCIAS BIBLIOGRÁFICAS

- BRIMBERG, J. (1995). "The Fermat-Weber location problem revisited". Mathematical Programming, 71, pp. 71-76.
- CÁNOVAS, L., CAÑAVATE, R. Y MARÍN, A. (2002). "On the convergente of the Weiszfeld algorithm". Mathematical Programming, 93, pp. 327-330.
- CAÑAVATE, R. (2001). "El algoritmo de Weiszfeld para el problema de Weber". Tesina de Licenciatura. Universidad de Murcia.
- CHANDRASEKARAN, R. y TAMIR, A. (1989). "Open cuestiones concerning Weiszfeld's algorithm for the Fermat-Weber location problem". Mathematical Programming, 44, pp. 293-295.
- KUHN, H.W. (1973). "A note on Fermat's problem". Mathematical Programming, 4, pp. 98-107.
- LOVE, R.F., MORRIS, J.G. y WESOLOWSKY, G.O. (1988). "Facilities location: Models & methods". Amsterdam. North-Holland.
- WEBER, A. (1909). "Über den standort der industrien". Tübingen. (Traducción de FRIEDERICH, C.J. (1929). "Theory of the location of industries". University of Chicago Press.
- WEISZFELD, E.V. (1937). "Sur le point pour lequal la somme des distances de n points donnés est minimum". The Tohoku Mathematical Journal, 43, pp. 335-386.
- WENDELL, R.E. y HURTER A.P., Jr. (1973). "Location theory, dominance and convexity". Operations Research, 21 (1), pp. 314-320.