ÍNDICE

Estudios de Economía Aplicada
Nº 11, 1999. 143-160

Aplicación del método del conjunto activo al equilibrio estático del usuario en redes

PEDREIRA ANDRADE, L. P. *
SEIJAS MACÍAS, J. A.*
Universidad de A Coruña

Esta versión incluye todas las correcciones sugeridas por el evaluador, las cuales nos han parecido oportunas y por las que le quedamos muy agradecidos.

RESUMEN

El objetivo central de este trabajo es comprobar el comportamiento del método del conjunto activo como medio de asignación óptima en redes de tráfico con presencia de congestión. Diversos estudios proponen la utilización de métodos basados en la idea del conjunto activo, o en la dirección del gradiente a la hora de establecer el nivel de tráfico óptimo en los arcos de una red. En este trabajo se plantea una propuesta similar. Consideramos una red de tráfico donde las funciones de coste de los arcos son del tipo BPR, la demanda de flujo entre los orígenes y destinos está dada y es constante. No existen valoraciones temporales; es, por tanto, una aproximación puramente estática al problema. Tampoco hemos considerado la posible aproximación estocástica. En el óptimo, la red asignará todo el tráfico a los diversos arcos de forma tal que el coste global de funcionamiento de la red será el mínimo; asimismo, el coste de todas las rutas posibles será igual, de modo que ningún usuario estará dispuesto a modificar su ruta de forma individual puesto que no le supondrá una reducción el coste del viaje.

Palabras Clave: Tráfico, Optimización de Redes, Conjunto Activo.

ABSTRACT

In this work we study the behavior of the active set method as a means of optimum assignment in traffic networks with congestion. There are several studies about the use of methods based on the

* Este trabajo se ha realizado con la ayuda financiera del Ministerio de Fomento, Ayudas a la Investigación 1997, proyecto: "Estimación del peaje óptimo asociado al uso de infraestructuras viarias".

idea of the active set, or in the direction of the gradient to establish the optimum traffic level in the links of a network. In this work is outlined a similar proposal. We consider a traffic network where the cost functions of the links are of the type BPR, the flow demand between the origins and destinations is given and constant. There isn't temporary valuations; it is, by so much, an approximation purely statics to the problem. We haven't considered the possible stochastic view. In the optimum, the network will assign all the traffic to the several links in a way such that the global operation cost of the network will be the minimum; also, the cost of all the possible routes will be equal, and nobody will modify her route, as an individual way, since will not suppose to her a reduction the cost of the trip.

Keywords: Traffic, Network Optimization, Active Set Method.

Códigos UNESCO:120710,120711,531212.

Artículo recibido en septiembre de 1997. Revisado en noviembre de 1998.

ÍNDICE

1. Optimización de la Red

El problema de optimización estática del tráfico de una red consiste en la mimización del coste total asociado a la asignación de tráfico de la red, de forma tal que todos los usuarios alcancen sus destinos desde sus respectivos orígenes. La asignación óptima resultante se denomina asignación de equilibrio del usuario, cuando ningún usuario puede mejorar el coste de su viaje mediante un cambio unilateral en su ruta.

Representamos la red como un grafo dirigido *G(N,A),* donde *N* es el conjunto de nodos y *A* es el conjunto de arcos, que interconectan dichos nodos. No existen arcos con entrada y salida en el mismo nodo (lazos). Dentro del conjunto de nodos podemos distinguir tres tipos: a) Nodos origen: Son aquéllos de donde parten los usuarios para dirigirse a un destino. Son los únicos nodos de la red que pueden generar tráfico. b) Nodos destino: Serán los nodos hacia donde se dirige el flujo de tráfico que existe en la red. Son receptores de tráfico. c) Nodos intermedios: Se ubican dentro de la red entre los nodos origen y destino. Actúan como distribuidores del tráfico existente, posibilitando la existencia de diversos caminos alternativos entre los nodos orgien y los nodos destino. El grafo ha de ser conexo con respecto a los nodos origen y destino, de tal forma que se garantiza que existan rutas que los conecten. En el Gráfico 1 se muestra un ejemplo de red de tráfico.

La variable de optimización es el flujo de usuarios que circula por la red. A la hora de hablar de flujo en una red podemos considerar dos visiones: en un arco y en una ruta. Arco es simplemente un nexo directo entre dos nodos, por lo tanto, el flujo que sale del nodo cabecera del arco será el mismo que el que llega al nodo objetivo de

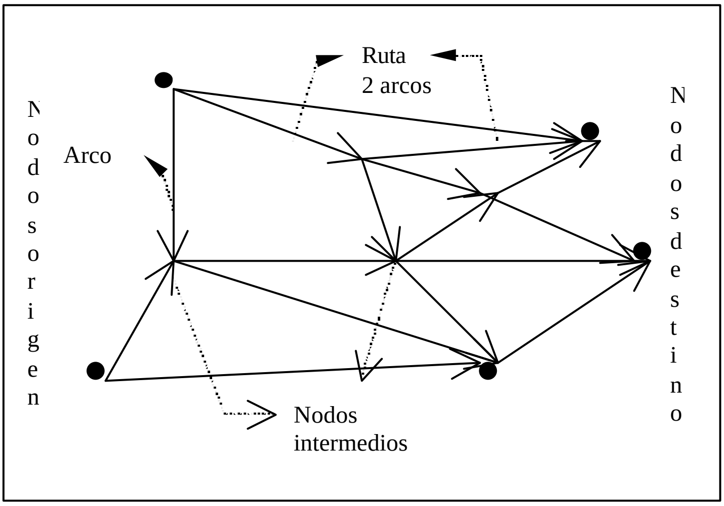


Ilustración 1: Red de Tráfico

ÍNDICE

dicho arco; este principio se conoce como *principio de conservación del flujo de arco*. Ruta es un nexo entre un nodo origen y un nodo destino, y está consitituida por un conjunto de arcos, y un conjunto de nodos intermedios[1].

Definimos el flujo x_i como el total de usuarios que circulan a través de un arco *i*. En base al principio de conservación del flujo podemos enunciar la siguiente proposición que establece una relación entre el flujo de ruta y el flujo de arco, permitiéndonos expresar el problema en cualquiera de las dos variables.

Proposición: El flujo de un arco será igual a la suma de los flujos de todas la rutas que utilicen dicho arco. Sea f_J^{ab} el flujo que circula entre los conjuntos de nodos origen y destino a y b, respectivamente, a través de la J-ésima ruta. Entonces dado el principio de conservación del flujo, podemos escribir el flujo a través del i-ésimo arco de la ruta J-ésima ($i \in J$), donde J es el conjunto de índices de los arcos que componen dicha ruta , como:

$$x_i = \begin{cases} \sum_J f_J^{ab} d_J^{ab} & \text{si solo existe un origen y un destino} \\ \sum_A \sum_B \sum_J f_J^{ab} d_J^{ab} & \text{si existen multiples origenes y destinos} \end{cases}$$

donde δ_J es una función que vale 1 si el arco x_i forma parte de la J-ésima ruta entre el origen y el destino considerado, y 0 en otro caso.

Las restricciones del problema se determinan en base al principio de conservación del flujo de arco, y el *principio de conservación del flujo de nodos intermedios* establece que: El flujo de los arcos de entrada a un nodo intermedio de la red de transporte ha de ser igual al flujo de los arcos de salida de dicho nodo. Asimismo, también se ha de cumplir: 1) El flujo de arcos de entrada a un nodo destino de una red de transporte ha de ser igual a la suma de usuarios que demanda ese nodo como destino de sus viajes. 2) El flujo de arcos de salida de un nodo origen de una red de transporte ha de ser igual a la suma de usuarios que oferta ese nodo como origen de sus viajes. Las restricciones resultantes son funciones lineales y puden ser dos tipos: a) restricciones de no negatividad de los flujos ($xi \geq 0$, $\forall i \in A$) y b) restricciones de igualdad. Las segundas representan el principio de conservación del flujo en nodos intermedios y los principios de flujo de salida igual a oferta de usuarios para los nodos origen y de flujo de entrada igual a demanda de usuarios para los nodos de demanda. El conjunto de valores del lado derecho de la restricciones se recoge en el vector *q*.

1. En el caso de no existir nodos intermedios solo existe un arco y por lo tanto, no existirá diferencia entre arco y ruta.

$$q = \left\{ q_i, i \in N / q_i = \begin{cases} 0 \text{ si } i \text{ es un nodo intermedio} \\ \text{Oferta}_i \text{ si es un nodo origen} \\ \text{Demanda}_i \text{ si es un nodo destino} \end{cases} \right\}.$$

La matriz R de restricciones de igualdad estará formada por 0, 1, ó -1; siendo una matriz poco densa, lo que facilita su factorización[2].

$$R_{ij} = \begin{cases} 0 \text{ si el arco } i \text{ no incide a / desde el nodo } j \\ 1 \text{ si el arco } i \text{ es incidente desde un nodo intermedio } j \text{ o} \\ \text{incidente a un nodo destino } j \text{ o incidente desde un nodo origen } j. \\ -1 \text{ si el arco } i \text{ es incidente a un nodo intermedio } j \text{ o} \\ \text{incidente desde un nodo destino } j \text{ o incidente a un nodo origen } j. \end{cases}$$

El número de filas será igual al número de nodos del problema, *N*, y el número de columnas vendrá determinado por el número de arcos, *A*, que presente la red. Para redes de tamaño grande, este número suele ser un múltiplo del número de nodos, así pues, el orden de la matriz *R* será $N \times (a \times N)$, donde $a > 1$. El número de columnas será, habitualmente, "*a*" veces superior al número de filas; lo que nos garantiza la aplicabilidad del método del conjunto activo, puesto que en general podremos determinar el espacio nulo.

La función objetivo mide el coste asociado al funcionamiento de la red. Para ello utilizamos una función que recoja el coste que soportan los usuarios por circular a través de los distintos arcos. Esta función de coste de arco, c_i, será no lineal, continua, y creciente; dependiente directamente del flujo por dicho arco, x_i, asimismo deberá ser asintótica y recoger de este modo la existencia de congestión en el arco para un determinado nivel de flujo/capacidad. De las diversas posibilidades de implementación de dicha función, una de las más frecuentes en la literatura es la establecida por el Bureau of Public Roads de Washington (E.U.A.) en 1964 y cuya formulación es $c_i = tf_i(1 + a(x_i / K_i)^b)$, donde tf_i es el tiempo necesario para atravesar el arco i si no existiese ningún grado de congestión[3], x_i es el flujo que circula a través del i-ésimo arco, K_i es la capacidad máxima asociada a dicho arco, y *a* y *b* son dos constantes cuyos valores son: $a = 0.15$ y $b = 4$. Este tipo de función ha sido ampliamente utilizada y no suele presentar problemas a la hora de implementarla en los algoritmos de asignación habituales. La función cumple las características deseadas:

---

2. Véase Matstoms (1994) para la implementación de un método de factorización QR de matrices no densas en MATLAB.

3 También denominado tiempo con tráfico libre.

a) Es una función no lineal y continua en todo $\Re$, si $K_i \neq 0$.

b) Es una función creciente: $\frac{\partial c_i}{\partial x_i} = \frac{abtf_i x_i^{b-1}}{K_i^b} \succ 0$, para x_i, tf_i y $K_i \geq 0$.

c) Es una función convexa: $\frac{\partial^2 c_i}{\partial x_i^2} = \frac{a(b-1)btf_i x_i^{b-2}}{K_i^b} \succ 0$,para $x_i \geq 0$, y Ki, $tf_i > 0$.

Una de las principales críticas establecidas con respecto a esta función es que no es asintótica. Davidson (1966) propone la utilización de una función alternativa, asintótica, que tienda a infinito según nos aproximemos a la capacidad máxima. Nosotros nos decantaremos por la utilización de la función BPR, puesto que permite la existencia de flujos superiores a la capacidad, y por tanto no implica que el algoritmo de optimización tenga que controlar las cotas superiores de las variables. En general, la utilización de la función de Davidson supone una mayor complejidad de gestión del algoritmo y no parece implicar ningún tipo de mejora de la precisión del mismo.

2. Problema de Optimización

La introducción del problema de minimización del coste asociado a una determinada asignación de flujo en una red de tráfico se conoce como equilibrio estático en redes. La determinación de dicho equilibrio ha sido objeto de numerosos estudios y aporximaciones (Sheffi, 1985). Las primeras aproximaciones al problema se deben a Beckaman *et al*. (1955), que establecen la formulación clásica del problema para funciones de coste separables, donde el coste en cada arco es independiente del flujo en los otros. Este modelo ha sido ampliamente aplicado al análisis del transporte y se han desarrollado numerosos algoritmos de resolución (Thomas 1991, Patriksson, 1995).

Posteriores reformas introducen puntos de vista alternativos, principalmente destaca la introducción de funciones de coste no separables, donde existen interacciòne entre flujos y costes de arcos diferentes, y para los cuales la formulación de equlibrio se realiza como un problema de desigualdades variacionales (Nagurney, 1993).

Consideramos dos posibles criterios de optimización: 1) Minimizar el coste para cada usuario. Este es, por ejemplo, el criterio que se sigue en el algoritmo de determinación del camino más corto de una red. 2) Minimizar el coste total para todos los usuarios. En un problema de optimización de redes, y suponiendo que el coste de arco es función del flujo que los atraviesa, se demuestra que ambos criterios se cumplen para la función objetivo adecuada[4].

Si buscamos responder sólo al segundo criterio, el problema de optimización es:

$$\operatorname*{Min}_{x \in \Re^n} \quad \mathrm{CT} = \sum_i \mathrm{x}_i c_i(x_i)$$

Sujeto a:

$$Rx = q$$
$$x_i \geq 0, \quad \forall i = 1, \ldots, n.$$

Si calculamos :

$$\frac{\partial CT}{\partial x_i} = tf_i(1 + a(x_i / K_i)^b) + tf_i ab(x_i / K_i)^b \qquad (1)$$

el primer término de (1) es el coste marginal que soporta un nuevo viajero que se incorpore al arco i de la red; mientras que el segundo término viene a representar el coste marginal que el nuevo viajero impone al resto de usuarios de la red, y evidentemente, a sí mismo.

En el óptimo todas las rutas tendrán el mismo coste, puesto que si en alguno de ellos el valor del coste fuese menor, se produciría un desplazamiento de usuarios hacia dicha ruta. Por tanto en el equilibrio, el coste de atravesar todos los arcos que conecten de forma directa el origen y el destino ha de ser el mismo. Podemos expresar esta condición de equilibrio como:

$$\begin{array}{l} c^*_i(x_i) = c^*, \quad \forall i, x_i > 0 \\ \forall i, x_i = 0 \quad \Rightarrow \quad c^*_i(0) = c_i(0) \geq c.^* \end{array} \qquad (2)$$

Vemos que, asimismo, si el flujo de un arco es igual a cero, entonces es que el coste de atravesar dicho arco con flujo cero, esto es el valor de tf_i, es superior al valor de coste de equilibrio c*; de ahí que el arco permanezca sin flujo en la solución óptima. Se puede demostrar, Newell (1980), que este criterio de optimización no es apropiado para la resolución del problema puesto que sólo nos produce una respuesta óptima para el segundo criterio de asignación: el valor óptimo minimiza el coste total de funcionamiento de la red, pero no supone que todos los viajeros sigan la ruta de coste mínimo.

4. Véase Newell (1980, pp. 137-138)

Beckman et al. (1955) demuestran que empleando un función objetivo adecuada podemos responder de forma precisa a ambos criterios de asignación. Su aproximación se basa en el planteamiento del siguiente problema de asignación de equilibrio del usuario (AEU):

$$\underset{x \in \Re^n}{\text{Min}} \quad \text{CT} = \sum_{\text{i}} \int_0^{x_i} c_i(w)dw$$

sujeto a:

$$Rx = q$$

$$x_i \geq 0, \quad \forall i = 1, \ldots, n.$$

Si al igual que antes, calculamos la derivada de CT con respecto a x_i, obtendremos que ésta es igual c_i. Dado este resultado, la condición de equilibrio es:

$$\begin{aligned} c_i(x_i) &= c \quad \text{en todas las rutas con } x_i \succ 0 \\ c_i(0) &\geq c \quad \text{en todas las rutas con } x_i = 0. \end{aligned} \tag{3}$$

Bajo esta condición los costes de todas las rutas utilizadas entre el origen y el destino serán iguales. En esta situación y para esta función objetivo, los dos criterios de asignación se cumplen; esto es, todos los usuarios seguirán la ruta de coste mínimo, y se minimiza el coste total de todos los viajeros. Por lo tanto, de acuerdo con el equilibrio de usuario ningún viajer podrá reducir el coste de su viaje mediante un cambio unilateral de su ruta.

La formulación del problema AEU sólo es válida cuando las funciones de coste son separables. Su resolución habital es a través del algoritmo de las combinaciones convexas, planteado por Frank-Wolfe (1954); Lupi (1985) demostró que dicho algoritmo era superior a la mayoría de las alternativas planteadas. No obstante, el algoritmo original ha sido objeto de diversas mejoras en los últimos años, Arezki & van Vliet (1990) introducen el modelo PARTAN, Weintraub (1985) realiza mejoras en la determinación de la longitud de paso, Larsson & Patriksson (1992) emplean la descomposición simplicial.

Asimismo, paralelamente se han desarrollado otro tipo de aproximaciones algoritmícas, entre las cuales destancan la aplicación de métodos del conjunto activo. Entre otro podemos citar a Toint & Tuyttens (1990), Wu (1993) y Nabona & Heredia (1992). Siguiendo sus planteamientos nuestra aproximación también desarrolla un método del conjunto activo con una aproximación quasi-newton.

3. Aplicación Basada en el Método del Conjunto Activo

Dado nuestro problema de optimización AEU como:

$$\underset{x \in \Re^n}{\mathrm{Min}} \quad CT(x) = \sum_i \int_0^{x_i} c_i(w)dw$$

sujeto a: (4)

$$R\breve{x} = q$$
$$x_i \geq 0, \quad \forall i = 1, \ldots, n.$$

Tenemos una función objetivo no lineal y un conjunto de restricciones lineales de igualdad junto a las restricciones de no negatividad de las variables. Siguiendo los planteamientos de Fletcher (1987) y Gill et al. (1981, 1991) para métodos del conjunto activo, abordamos este problema mediante el siguiente algoritmo[5].

Paso 0: Inicialización del Algoritmo y estimación del punto inicial

Dada la función objetivo CT(x), calculamos el gradiente de dicha función $g(x) = \nabla CT(x)$; y una matriz simétrica B como aproximación a la Hessiana de CT(x), $G = \nabla^2 CT(x)$. Inicialmente establecemos $B^0 = I_N$, donde el superíndice indica el número de iteración.

Calculamos un punto inicial x^0, vector de flujos iniciales de acuerdo con el criterio de asignar flujo a los arcos de forma tal que presenten valores superiores a cero[6]. En el caso de que esto no fuera factible, el algoritmo del conjunto activo no plantea problemas, siempre y cuando consideremos las restricciones de no negatividad pertinentes dentro del conjunto de restricciones activas, *W*.

Factorizamos la matriz *W*, de restricciones activas, que coincide, en principio, con la matriz de restricciones de sigualdad *R*, para lo cual utilizamos la factorización LQ. Dado que *W* es una matriz de orden, como mínimo, *N* x *A*, la factorización LQ, supone que *Q* es una matriz *A* x *A*, y por tanto, *L* es una matriz *N* x *N* triangular inferior, produciéndose asimismo una matriz de ceros 0, de orden *N* x (*A*-*N*). El resultado es: *W* = (*L* 0) *Q*. A partir de la matriz *Q*, determinamos la matriz *Z*, como matriz del espacio nulo de *R*, formada por las *A*-*N* últimas columnas de la matriz *Q*.

Paso 1: Cálculo de la dirección factible

Una vez establecido un punto de partida debemos dotarnos de una dirección de búsqueda factible para proceder a avanzar a lo largo la misma. Al ser un problema

5. Aproximaciones al conjunto activo se pueden encontrar en: Fletcher (1987), Gill et. al (1981, 1991), Nash & Sofer (1996), Pierre (1969) y Stweart (1973).

6. Seguimos una estrategia opuesta a la habitual asignación “todo-o-nada” donde sólo una ruta, la de menor coste, recibe todo el flujo.

de minimización, ésta dirección será descendente, de forma tal que nos lleve a un menor valor de la función objetivo. La dirección de búsqueda en la i-ésima iteración es:

$$y^i = -Z^t g(x^i) / Z^t B^i Z, \qquad (5)$$

donde el numerador es el denominado gradiente reducido y el denominador es la aproximación de la Hessiana reducida.

Si $Z^t B Z$ es una matriz definida positiva, entonces la dirección obtenida a través de la resolución de (5) es una dirección de búsqueda descendente puesto que

.

$$g^t Zy = -(Z^t g)^t (Z^t BZ)^{-1} Z^t g < 0$$

Paso 2: Longitud del paso y nueva solución

La longitud de paso debe controlar que ninguna de las variables tome valores negativos, y por tanto, no se incumplan las condiciones de no negatividad, así como el cumplimiento de las restricciones de igualdad. Podemos entonces calcular la nueva solución como: $x^i = x^{i-1} + \alpha y^i$, donde α representa la longitud máxima de paso, e y^i representa el vector de dirección de búsqueda obtenida en el i-ésima iteración, siendo x^{i-1} la solución anterior.

El valor de la longitud de paso, supuesto que $y \neq 0$ es:

$$\alpha = \underset{j}{Min}\left\{\left|\frac{x_j^i}{y_j^i}\right|, 1\right\} \qquad (6)$$

De esta forma tenemos garantizado el cumplimiento de las condiciones de no negatividad en todo momento. Una cota genérica más cuidadosa, que sólo consideraría aquellos casos donde el valor de y_j es negativo, sería:

$$\alpha = \underset{\substack{j \\ y_j^i < 0}}{Min}\left\{\left|\frac{x_j^i}{y_j^i}\right|, 1\right\} \qquad (7)$$

Si tenemos algún flujo igual a cero para alguno de los arcos, esto supone que algunas de las restricciones de no negatividad formarán parte del conjunto activo; entonces podremos utilizar (7) ligeramente modificado para calcular la longitud de paso:

ÍNDICE

$$\alpha = \underset{\substack{j \\ y_j^i<0 \\ x_j^i \neq 0}}{Min}\left\{\left|\frac{x_j^i}{y_j^i}\right|,1\right\} \qquad (8)$$

Paso 3: Criterio de Optimalidad

El algoritmo deberá finalizar en el momento en el que se haya alcanzado el punto óptimo. Las condiciones de primer orden para problemas de optimización de funciones no lineales con restricciones lineales son:

$$Z^t g(x^*) = 0, \qquad (9)$$

donde g(x*) es el valor del gradiente de la función objetivo en el punto óptimo. Entonces, dado que siempre existirán restricciones activas, se calcula el valor de los multiplicadores de Lagrange, $l = ((W^T W)^{-1} W^T)^T . g(x)$, en el punto óptimo, y si el resultado es mayor o igual que cero para todos ellos, entonces el punto x* es un punto óptimo local: $l^* \geq 0$. Si alguno de los valores de λ (para las restricciones de desigualdad) es negativo, entonces indica que la función objetivo puede decrecer moviéndonos a lo largo de la dirección de la restricción cuyo multiplicador sea negativo. Esto supone que hemos considerado nula alguna de las variables cuya restricción se ha introducido en el conjunto activo, y por tanto, podemos mejorar la solución dando un valor a dicha variable.

La condición de segundo orden para que x* sea un óptimo global es que la matriz Hessiana proyectada sea definida positva. En nuestro caso, como trabajamos con la estimación quasi-newton de dicha matriz esta condición supone que $Z^t B^K Z$ debe ser una matriz definida positiva. Podemos también calcular los estimadores de segundo orden de los multiplicadores de Lagrange a través de la fórmula:

$$h = (WW^t)^{-1} W(g(x^*) + B^i\ y^i) .$$

En el óptimo global, los dos estimadores de los multiplicadores de Lagrange coinciden y son positivos o iguales a cero.

El supuesto más beneficioso para nuestro algoritmo sería encontrar la solución óptima y detener el proceso iterativo. No obstante, en todo proceso algorítmico la sucesión de iteraciones converge hacia el óptimo, aunque no necesariamente se alcanza en un número relativamente pequeño de iteraciones. En este estudio consideramos que el mejor criterio de convergencia consiste en controlar la optimalidad de la solución a través del criterio del gradiente proyectado $Z^t g(x^i) \rightarrow 0$, donde tenemos garantizada la convergencia local al óptimo, y dado el carácter monótono de la función objetivo, dicha convergencia es global. Cuando el valor de este criterio disminuya por debajo de un $\varepsilon > 0$, pequeño, entonces detendremos nuestro algoritmo.

Por último, existen otras situaciones en las cuales el algoritmo acaba puesto que no es posible continuar avanzando: Si $g(x^{i+1}) = g(x^i)$ no es posible calcular la nueva matriz auxiliar B^{i+1}, y no podemos establecer una dirección de búsqueda. Si $y^i = 0$ ó $\alpha = 0$, en este caso la dirección de búsqueda vale cero y, por lo tanto, tenemos dos soluciones sucesivas iguales.

Paso 4: Determinar si alguna variable se ha tomado valor nulo

Puede ocurrir que durante el desarrollo del algoritmo alguna de las variables se anule y, por tanto, habría que incluir la restricción de no negatividad correspondiente a dicha variable dentro del conjunto de trabajo, lo que afectará a la matriz *Z*, puesto que el espacio nulo se modifica. La incorporación de una nueva restricción al conjunto activo supone la necesidad de realizar una adaptación del mismo, así como de la factorización de dicho conjunto y de la matriz del espacio nulo, *Z*, obtenida para el primer conjunto activo. Para ello, debemos o bien calcular la nueva factorización de la nueva matriz del conjunto activo, o adaptar la factorización existente. El proceso de adaptación se basa en calcular una matriz householder *H* que anule las t+2 componentes de *L*, donde t es el número de restricciones del conjunto activo antes de incorporar la nueva restricción. Si *W* es el conjunto activo anterior y w la nueva restricción:

$$\overline{W} = \begin{pmatrix} W \\ w \end{pmatrix} = \begin{pmatrix} L & 0 \\ \hat{r} & Q \end{pmatrix} Q = (\overline{L} \quad 0)\overline{Q}$$

donde

$$\overline{Q} = QH$$

La matriz del espacio nulo *Z*, vendría dada por las *A*-(*N*+*l*) últimas columnas de $\overline{Q}$, donde *l* es el número de restricciones de no negatividad incorporadas al conjunto de activo inicial.

Paso 5: Adaptación de la matriz B

Se procede a la actualización de la matriz de aproximación de la Hessiana, *B*, de acuerdo con Broyden - Fletcher - Goldfard - Shanno (BFGS) que proponen como forma de cálculo de la actualización de la matriz B en la i+1-ésima iteración:

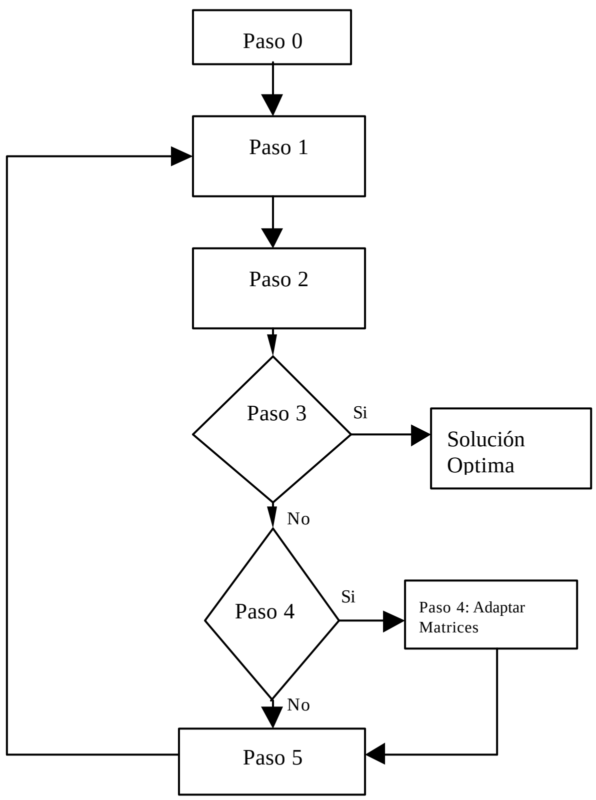
$$B^{i+1} = B^i - \left(\frac{1}{s^t B^i s} B^i s s^t B^i \right) + \frac{1}{p^t s} p p^t$$

donde $p = g^i - g^{i-1}$, donde g^i indica el gradiente evaluado en x^i, y $s = x^i - x^{i-1}$.

ÍNDICE

DESCRIPCIÓN DEL ALGORITMO.

1) Paso 0: Calcular el gradiente de la función objetivo, *g*, la matriz de restricciones activas, *W* y su factorización, la matriz del espacio nulo resultante *Z*, y la solucción inicial, x^0. Sea $B = I_N$.
2) Paso 1: Calcular la dirección de búsqueda, y^i.
3) Paso 2: Calcular la longitud de paso, α, y la nueva solución, x^i.
4) Paso 3: Calcular el gradiente proyectado. Si $Z^t g(x^i) \to 0$, entonces la nueva solución es óptima, $x^* = x^i$. Stop.
   En caso contrario Paso 4.
5) Paso 4: Determinar si alguna variable a tomado valor nulo en la nueva solución. En caso positivo, adaptar matriz de restricciones activas, *R* y Paso 5. En caso contrario, Paso 5.
6) Paso 5: Calcular el gradiente en la nueva solución, adaptar la matriz aproximación a la Hessiana *B*, volver al Paso 1.



El algoritmo propuesto se ha implementado dentro de los procedimientos del programa de cálculo simbólico Mathematica, versión 2.1.

El algoritmo resultante es convergente globalmente al óptimo, esto es, si encuentra una solución óptima, dicha solución es un óptimo global.

Consideremos el programa de optimización definido en (4) como un problema de minimización de una función convexa en un conjunto convexo *K*, donde

$$K = \{x / Wx = q, x_i \geq 0, i = 1,2,\ldots,n\}.$$

En el primer apartado de este trabajo se ha mostrado que la función objetivo es convexa. El conjunto de restricciones es un conjunto convexo puesto que se deriva del hecho de que las restricciones son funciones cóncavas. En

ÍNDICE

esta situación y para el problema de optimización definido en (4) se demuestran los dos teoremas siguientes:

Teorema 1:

Toda solución local al problema de optimización convexa (4) es una solución global, y el conjunto de soluciones globales es convexo[7].

El siguiente corolario del teorema anterior establece la unicidad de la solución.

Corolario: Si la función CT(x) es estrictamente convexa en K entonces cualquier solución global es única.

Otra ventaja del problema de programación convexa definida es que nos garantiza que si un punto solución satisface las condiciones necesarias de primer orden de Kuhn-Tucker (Fletcher, 1987), entonces dicha solución es un óptimo global.

Teorema 2:

Sea el problema de optimización definido en (4), si la función objetivo CT(x) y las funciones de las restricciones definidas por $R\,x = q$ son funciones de clase uno en K y si existe un punto x* que cumple las condiciones:

$$Z^T g(x^*) = 0,$$

$$l^* \geq 0, \text{ para las restricciones de desigualdad,}$$

entonces dicho punto es una solución global del problema planteado en (4).

Demostración: Sean $c_j(x) = w_j x - q_j, \quad \forall j = 1,\ldots,m,$ funciones cóncavas y sea x' $\in$ K tal que x' $\neq$ x*.

Entonces dado que $l^* = ((W^T W)^{-1} W^T)^T g(x^*) \geq 0$ y $c'_j = w_j x' - q_j = 0,$ y siendo la función objetivo CT convexa, se cumple:

$$CT(x') \geq CT(x') - \sum_{j=1}^{m} l^*_j c'_j \geq CT(x^*) + (x' - x^*)^T g(x^*) - \sum l^*_j (c^*_j + (x' - x^*)^T w_j).$$

Entonces, puesto que $l^*_j c^W_j = 0, \forall j,$ aplicando el lema de Farkas, $g(x^*) = \sum_j w^W_j l^W_j,$ queda demostrado que $CT(x') \geq CT(x^*)$, y por tanto, x* es una solución global.

7. La demostración del teorema figura en Fletcher (1987), p.67.

4. Resultados Obtenidos

Para proceder a una evaluación hemos tomado tres ejemplos de redes sencillas y, hemos calculado el óptimo utilizando el método del conjunto activo[8]. Los resultados obtenidos son comparados con los que se obtienen a través del método de Frank-Wolfe, tal y como aparece recogido en Sheffi (1985).

Ejemplo 1. Este ejemplo está tomado de Sheffi (1985). Consideramos una red formada por un nodo origen y un nodo destino conectados por tres arcos. No existirán nodos intermedios por lo que cada arco será equivalente a una ruta entre el origen y el destino. Las funciones de coste de cada arco están basadas en la función de flujo-coste BPR. La demanda total de usuarios en la red será de 10. Así pues, sólo existirá una restricción de igualdad en el conjunto activo.

Ejemplo 2. El siguiente ejemplo está tomado de Hearn & Ribera (1980) y considera una red con 18 arcos y 9 nodos, estableciendo dos nodos origen y dos nodos destino, lo que supone un total de 4 demandas fijas de tránsito entre orígenes y destinos. Los arcos presentan funciones de coste del tipo BPR.

Ejemplo 3. Este ejemplo se ha realizado en función del que plantean Nguyen & Dupuis (1984). Consideramos una red de 13 nodos y 19 arcos, y 4 pares origen-destino. Al igual, que en los casos anteriores trabajaremos con funciones de coste del tipo BPR, y la demanda entre los diversos pares de nodos origen-destino se considerará fija. La principal característica diferencial de este ejemplo es que las capacidades de los arcos son altas en relación al flujo global que circula por la red, de forma tal, que ésta no tiene porque aparecer fuertemente congestionada en la solución óptima, como sucedía en la red del ejemplo 2, donde las capacidades de los arcos apenas llegaban para soportar el flujo circulante entre los diversos orígenes y destinos.

8. Para la realización de todos los cálculos se ha usado un ordenador PC clónico con procesador Pentium de Intel a 100 MHz. Se ha utilizado la versión 2.1 del programa de Mathematica para Windows 3.11.

ÍNDICE

Resultados

	Ejemplo 1		*Ejemplo 2*		*Ejemplo 3*	
	F-W	*C.A.*	*F-W*	*C.A.*	*F-W*	*C.A.*
Nº Iteraciones	11	10	14	14	22	22
Funcion Objetivo	189,332047	189,3320416	1894,78	1856,498199	67945,8	67836,37941
Arcos Flujo Nulo	0	0	4	7	0	2
Arcos Congestión	2	2	8	8	3	4
Convergencia a la solución óptima	0,00	0,00	2,06	0,00	0,16	0,00

Las soluciones finales obtenidas por ambos métodos son diferentes en aquellas redes más complejas frente a las redes más sencillas donde existe una mayor proximidad de resultados; un análisis más pormenorizado de los mismos pone de relieve dos características:

a) El método del conjunto activo obtiene una solución mucho mejor que la obtenida a través del método de Frank-Wolfe. En los tres ejemplos el valor de la función objetivo es menor o igual empleando el método del conjunto activo que el de Frank-Wolfe; las diferencias son más significativas en el segundo, propiciadas por el mayor grado de congestión existente.

b) La velocidad de aproximación a la solución final es superior en las primeras etapas mediante el empleo del algoritmo de Frank-Wolfe; no obstante, según va aumentado el número de etapas tiende a disminuir su velocidad de convergencia, mientras que la del algoritmo del conjunto activo se mantiene durante un mayor número de etapas, lo que permite una mejor aproximación al optimo.

5. Conclusiones

El análisis de la aplicación del método del conjunto activo pone de manifiesto:

1.º La solución alcanzada por el método del conjunto activo en redes es una solución óptima global, proporcionando un valor de la función objetivo inferior al obtenido mediante el método de Frank-Wolfe. El método del conjunto activo se muestra más lento a la hora de aproximarse a la solución óptima en las primeras etapas, propiciado por el empleo de una línea de búsqueda quasi-newton; en estas primeras iteraciones el método de Frank-Wolfe muestra resultados superiores. No obstante, éste tiende a sufrir una cierta ralentización según van

ÍNDICE

avanzando las iteraciones propiciada por la propiedad descendente de la longitud de paso, lo que le impide aproximarse a buen ritmo al mínimo de la función objetivo, siendo ampliamente superado en velocidad por el método del conjunto activo.

2.ª Desde el punto de vista operativo hay que significar la mayor eficiencia del método del conjunto activo, puesto que tan sólo requiere una asignación inicial, frente al método de Frank-Wolfe que implica la necesidad de realizar una asignación "todo-o-nada" en cada una de las iteraciones. Este tipo de asignaciones se vuelven muy complicadas cuando existen numerosos arcos que posibilitan la existencia de un número grande de rutas. Este proceso es realmente complejo y difícil de implementar en un ordenador. Por su parte, el algoritmo del conjunto activo obliga a trabajar con matrices grandes, aunque no densas, lo que supone que se necesitará un cierto espacio de memoria para almacenamiento de las mismas, con el cual deberemos contar en el procesador. Lógicamente, cuanto mayor sea el número de arcos del problema, será superior el orden de las matrices con las cuales trabajamos y se incrementarán de forma importante nuestros requisitos de memoria.

En general, la comparación entre ambos algoritmos favorece al método del conjunto activo, puesto que, en general, obtiene mejores soluciones utilizando un número de iterantes menor o similar.

Bibliografía

AREZKI, Y. & VAN VLIET, D. (1990), *A Full Analytical Implementation of the PARTAN/Frank-Wolfe Algorithm for Equilibrium Assignment.* en "Transportation Science", 24, pp. 58-62.

BECKMAN, M., McGUIRE, C.B. & WINSTEN, C.B. (1955), *Studies in The Economics of Transportation.* Yale University Press, New Haven.

DAFERMOS, S.C. (1971), *An Extended Traffic Assignment Model with Applications to Two-Way Traffic*, en "Transportation Science", 5, pp. 366-389.

DAVIDSON, (1966), *A Flow Travel Time Relationship for Use in Transportation Planning*, en "Procedings of the Australian Road Research Board", 3, pp. 183-194.

FLETCHER, R. (1987), *Practical Methods of Optimization*, J. Wiley & Sons, New York.

FRANK, M. & WOLFE, P. (1956), *An Algorithm for Quadratic Programming*, en "Naval Research Logistics Quaterly", 3, pp. 95-110.

GILL, P; MURRAY, W. & WRIGHT, M. (1981), *Practical Optimization*, Academic Press, New York.

ÍNDICE

GILL, P; MURRAY, W. & WRIGHT, M. (1991), *Numerical Linear Algebra and Optimization*, Addison-Wesley, New York.

HEARN, D.W. & RIBERA, J. (1980), *Bounded Flow Equilibrium Problems by Penalty Methods*, en "Proceedings of the 1980 IEEE International Conference on Circuits and Compueters".

LARSSON, T. & PATRIKSSON, M. (1992), *Simplicial Decomposition with Disaggregated Represtation for the Traffic Assignment Problem.* en "Transportation Science", 26, pp. 4-17.

LUPI, M. (1983), *Convergence of the Frank-Wolfe Algorithm in Transportation Neteworks*, en "Civil Engineeering Systems", 19, pp. 7-15.

MATSTOMS, P. (1994), *Sparse QR Factorization in MATLAB*, en "ACM Transactions on Mathematical Software", 20, pp. 136-159.

NABONA, N. & HEREDIA, F.J. (1992), *Numerical Implementation and Computational Results of Nonlinear Network Optimization with Linear Side Constraints.*, en KALL, P. (1992), *System Modelling and Optimization*, Springer-Verlag, Berlin., pp. 301-310.

NASH, S. & SOFER, A. (1996), *Linear and Nonlinear Programming*, McGraw-Hill, New York.

NAGURNEY, A. (1993), "Network Economics", Kluwer, Dordecht.

NEWELL, G. (1980), *Traffic Flow on Transportation Networks,* MIT Press, Cambridge, Mass.

NGUYEN, S. & DUPUIS, C. (1984). *An Efficient Method for Computing Traffic Equilibria in Network with Asymetric Transportation Costs*, "Transportation Science", 18, pp. 185-202.

PATRIKSSON, M. (1995), "The Traffic Assignment Problem". VSP. Holanda.

PIERRE, D. (1969), *Optimization Theory with Applications,* Dover, New York.

SALTER, R. & HOUNSELL, N. (1996), *Highway Traffic Analysis and Design*, MacMillan Press, London.

SHEFFI, Y. (1985), *Urban Transportation Networks*. Prentice-Hall, Englewood Cliffs, N.J.

STWEART, G.W. (1973), *Introduction to Matrix Computations*, Academic Press, London.

THOMAS, R. (1991), *Traffic Assignament Techniques,* Avebury Press, Aldershot, U.K.

TOINT, P. & TUYTTENS, D. (1990), *On Large Scale Nonlinear Network Optimization*, en "Mathematical Programming", 48, pp. 125-159.

WEINTRAUB, A.; ORTIZ, C. & GONZALES, J. (1985), *Accelerating Convergence of Frank-Wolfe Algorithm*, en "Transportation Research B", 19, pp. 113-122.

WU, C. (1993), "Solving Large-Scale Nonlinear Network Problems with Relaxation and Decomposition Algorithms", Tesis Doctoral, Pennsilvanya State University.