

SOBRE LA AMPLITUD DE PASO MULTIVARIANTE EN PROGRAMACION NO-LINEAL CON CONDICIONES LINEALES

L. F. ESCUDERO

CENTRO DE INVESTIGACION UAM-IBM MADRID

En este trabajo se describe un nuevo método para la obtención de la amplitud de paso de la dirección de búsqueda en programación no-lineal con condiciones lineales. Tradicionalmente, se considera la amplitud de paso como un escalar con valor estrictamente positivo, tal que el nuevo punto también sea factible y suficientemente descendente. En su lugar, se propone en este trabajo una amplitud de paso multivariante tal que se limita, independientemente, la amplitud de cada elemento superbásico de la dirección de búsqueda; como resultado se permite la activación de más de una variable superbásica en la misma iteración.

Keywords. BASIC, SUPERBASIC AND NON-BASIC VARIABLES, DIRECTION OF SEARCH MULTIVARIATE STEPSIZE.

1. INTRODUCCION:

Sea el problema de programación no-lineal sin condiciones

$$\min\{F(X) : X \in R \subset R^n\} \tag{1.1}$$

donde

$$R = \{X | AX = b, U \leq X \leq L\} \tag{1.2}$$

tal que A es la matriz m.n de condiciones, b es el vector independiente, U y L son los -- vectores de acotación del vector solución X, y F es una función de clase ℓ^2 en R^n , al menos para el conjunto R de puntos factibles, tal que el nivel $L(\bar{X}) = \{X \in R : F(X) \leq F(\bar{X})\}$ para $\forall \bar{X} \in R$ está finitamente acotado y cerrado.

Sean M y J los conjuntos de condiciones variables, respectivamente. Sea $\bar{B}\bar{S}$ el conjunto de variables no-activas para el punto factible \bar{X} , y \bar{N} el conjunto de variables activas (i.e., $j \in \bar{N}$ implica que $\bar{X}_j = L_j$ ó $\bar{X}_j = U_j$ en la resolución del problema (1.1)-(1.2) en el espacio -- definido por el conjunto $\bar{B}\bar{S}$). Sea X^* un punto mínimo local débil en (1.1)-(1.2); i.e., X^* es un punto en el conjunto R para el que $\exists \delta > 0$ tal que $F(X^*) \leq F(X) \forall X \in R : \|X - X^*\| \leq \delta$. (El punto X^* será un mínimo local fuerte, si la rela-

ción anterior es de estricta desigualdad para $X^* \neq X$). Nota: $\|\underline{a}\|$ es la raíz cuadrada de la distancia euclídea del vector \underline{a} .

Dado que el conjunto R está definido por el sistema lineal de ecuaciones (1.2), se puede efectuar la caracterización de todos los movimientos factibles activos desde un punto -- factible, sea \bar{X} dadas las propiedades de los subespacios lineales. Sea el movimiento factible desde \bar{X} en el espacio definido por el conjunto $\bar{B}\bar{S}$; dado que $A\bar{X} = b$, $AX = b$, $X = \bar{X} + d$, y $X_j = \bar{X}_j \forall j \in \bar{N}$, resulta que

$$Ad = 0, \quad \bar{I}d = 0 \tag{1.3}$$

donde d es la dirección de búsqueda desde \bar{X} a X, e \bar{I} es la matriz $\bar{n} \cdot \bar{n}$ de límites activos en \bar{X} para $\bar{n} = |\bar{N}|$, tal que \bar{I} es la matriz unitaria n.n de la que se ha eliminado la fila relativa a la j-ésima variable para $\forall j \in \bar{N}$.

Dada la condición (1.3) para cualquier dirección factible activa, el vector d puede expresarse tal que

$$d = \bar{Z}d_s \tag{1.4}$$

- L.F. Escudero - Centro de Investigación UAM-IBM - Paseo de la Castellana, 4 - Madrid
 - Article rebut el març de 1986.

donde \bar{Z} es una matriz n.s para $s=n-m-\bar{n}$ de rango completo en columnas, tal que éstas forman la base del espacio nulo del rango de la matriz $(A^t, \bar{I}^t)^t$ (i.e., $A\bar{Z}=0, \bar{I}\bar{Z}=0$ y d_s es un vector (con dimensión s) denominado dirección de búsqueda superbásica o reducida. Es preciso notar que cualquier vector d_s produce una dirección factible activa d ; por otra parte, un vector d que no pueda expresarse por (1.4) no es una dirección factible activa en el espacio definido por el conjunto $\bar{B}\bar{S}$.

Sea \bar{h} el vector (con dimensión s) del gradiente reducido tal que

$$\bar{h} = \bar{Z}^t \bar{g} \quad (1.5)$$

donde $\bar{g} = g(\bar{X})$ es el vector (con dimensión n) del gradiente de F evaluado en \bar{X} . Es preciso notar que (ver /6/, /14/, /16/ entre otros) una condición necesaria para que el punto \bar{X} sea un mínimo local débil consiste en que éste sea un punto estacionario restringido; i.e., $\|\bar{h}\| = 0$; esta condición es equivalente a indicar que el gradiente \bar{g} se expresa como una combinación lineal de las filas de las matrices A e \bar{I} , tal que

$$\bar{g} = A^t \mu + \bar{I}^t \lambda \quad (1.6)$$

para los vectores μ y λ (con dimensiones m y \bar{n} , respectivamente) denominados multiplicadores de Lagrange de las condiciones y límites activos respectivamente. En base a (1.6), resulta que

$$\lambda = \bar{g}_N - N^t \mu \quad (1.7)$$

donde \bar{g}_N es el gradiente del conjunto \bar{N} de variables activas, y N es la matriz m. \bar{n} relativa a dicho conjunto en la matriz A, tal que $g^{-t} = (\bar{g}_{BS}^t, \bar{g}_N^t)$ y $A = (BS \ N)$ donde BS es la matriz m.(n- \bar{n}) relativa al conjunto $\bar{B}\bar{S}$. El vector μ debe satisfacer el sistema lineal

$$\bar{g}_{BS} = (BS)^t \mu \quad (1.8)$$

Sea $\bar{G} = G(\bar{X})$ la matriz n.n Hessiana de F evaluada en el punto \bar{X} . Se denomina matriz Hessiana reducida a la matriz, sea \bar{H} (con dimensión s.s) tal que

$$\bar{H} = \bar{Z}^t \bar{G} \bar{Z} \quad (1.9)$$

El conjunto suficiente de condiciones a satisfacer por el punto \bar{X} para ser un punto mínimo local débil en el problema (1.1)-(1.2) es el siguiente para casos no-degenerados -- (ver referencias arriba indicadas):

- (i) \bar{X} es factible; i.e., \bar{X} satisface (1.2)
- (ii) \bar{X} es un punto estacionario restringido, i.e., \bar{X} satisface (1.6) y, por tanto, $\|\bar{h}\| = 0$.
- (iii) La matriz \bar{G} es positiva definida para toda dirección factible activa, d en el espacio definido por el conjunto $\bar{B}\bar{S}$. Dada la expresión (1.4), esta condición es equivalente a indicar que la matriz Hessiana reducida \bar{H} (1.9) debe ser positiva definida.
- (iv) Los elementos del vector λ de los multiplicadores de Lagrange deben ser $\lambda_j > 0$ para $j \in \bar{N} | \bar{X}_j = L_j$ y $\lambda_j < 0$ para $j \in \bar{N} | \bar{X}_j = U_j$.

La estructura de este trabajo es la siguiente. La Sección 2 está dedicada a la motivación de la amplitud de paso multivariante. La sección 3 está dedicada a la caracterización de la matriz Z en problemas de gran escala. En las Secciones 4 y 5 se describe el algoritmo para la obtención de la amplitud de paso multivariante. La Sección 6 discute la convergencia. La Sección 7 presenta algunos resultados y, finalmente, en la última sección se discute el campo de aplicación.

2. MOTIVACION.

La metodología general para obtener X^* a partir de un punto factible \bar{X} y su conjunto asociado $\bar{B}\bar{S}$, es la siguiente:

- (1) Obtener una dirección superbásica, d_s , suficientemente descendente; i.e., d_s debe satisfacer la condición

$$(-\bar{h}^t d_s) / (\|\bar{h}\| \|d_s\|) \geq \sigma \quad (2.1)$$

donde σ es una tolerancia positiva que impide que el gradiente y la dirección sean cuasi-ortogonales. Ver en /3/, /5/, /7/, /8/, /9/, /10/, /14/, /16/ entre otros, los diversos métodos para obtener la dirección d_s que satisface la condición (2.1).

(2) Obtener la dirección de búsqueda d , utilizando la expresión (1.4)

(3) Obtener el nuevo punto factible para la iteración dada, sea k tal que

$$X^{(k)} = \bar{X} + \alpha d \quad (2.2)$$

donde α es la amplitud de paso. Tradicionalmente, se ha considerado que α es un escalar que:

(i) Impide que el punto $X=X^{(k)}$ (aún siendo factible en el sistema $AX=b$) viole los vectores límite U y L del conjunto $\bar{B}S$; es preciso notar que $X_j = \bar{X}_j \quad \forall j \in \bar{N}$. Dado que α debe permitir que αd sea descendente, resulta de ambas condiciones que

$$0 < \alpha \leq \alpha_M \quad (2.3)$$

donde α_M es el límite máximo para no violar la acotación de las variables; ver (2.8).

(ii) Reduce suficientemente el valor de la -- función objetivo, tal que el método es globalmente convergente; i.e., $||\bar{h}|| \rightarrow 0, k \rightarrow \infty$. Ver más abajo las condiciones que ha de satisfacer la amplitud de paso.

(4) Si $\alpha < \alpha_M$, se comienza en el paso (1) con la nueva iteración; si $\alpha = \alpha_M$, se activa la variable que ha producido la reducción en la amplitud de paso y, por tanto, se incluye en el conjunto \bar{N} y se reduce en una unidad el cardinal del conjunto $\bar{B}S$.

(5) Si el punto es óptimo para el conjunto $\bar{B}S$, se analiza el vector λ y, por tanto, el conjunto \bar{N} . Si $\lambda_j < 0$ para $j \in \bar{N} | \bar{X}_j = L_j$, ó si $\lambda_j > 0$ para $j \in \bar{N} | \bar{X}_j = U_j$, se desactiva dicha variable y, por tanto, se incluye en el conjunto $\bar{B}S$ y se reduce en una unidad el cardinal del conjunto \bar{N} ; se comienza en el paso (1) con la nueva iteración. En caso contrario, se considera que el punto \bar{X} es el punto X^* .

La metodología arriba descrita es globalmente convergente si la dirección de búsqueda d satisface la condición (2.1) y, además, la amplitud de paso (2.3) reduce suficientemente F , e.g. si se cumplen /4/ las condiciones -- GPW /18/, /22/.

$$(i) |g(\bar{X} + \alpha d)^T d| \leq -\eta g(\bar{X})^T d \quad (2.4a)$$

o, alternativamente /17/, caso de que el - gradiente sea difícil o costoso de calcular,

$$\frac{|F(\bar{X} + \alpha d) - F(\bar{X} + \nu d)|}{\alpha - \nu} \leq -\eta g(\bar{X})^T d \quad (2.4b)$$

donde ν es un escalar tal que $0 < \nu < \alpha$, y $0 \leq \eta < 1$.

$$(ii) F(\bar{X} + \alpha d) \leq F(\bar{X}) + (\mu g(\bar{X})^T d) \alpha$$

para $0 < \mu \leq 0.5$, tal que si $\mu \leq \eta$, $\exists \alpha$ que cumple las condiciones GPW en programación no-lineal sin restricciones. Valores típicos, $\eta = 0.9$ y $\mu = 10E-04$. Las condiciones (2.4a) y (2.4b) -- evitan que α tome un valor excesivamente pequeño; la condición (2.4c) impide que tome un valor excesivamente grande. Ver figura 1.

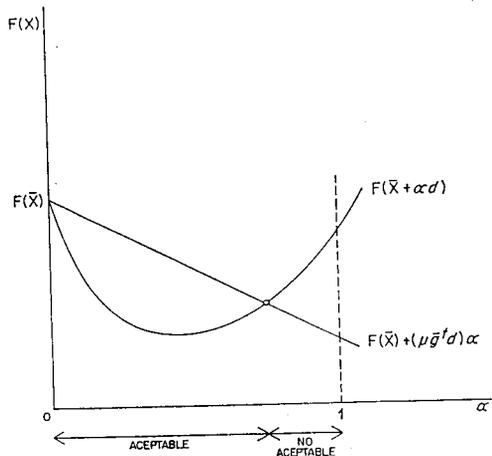


Figura 1- LÍMITES DE LA AMPLITUD DE PASO

Nota: $\alpha = 1$ para $\gamma = 0$; $\alpha' < \alpha''$ para $\gamma' > \gamma''$

Existen diversos métodos para la obtención - de la amplitud de paso que satisfaga las con- diciones GPW. La metodología denominada lí-- nea de búsqueda exacta consiste en obtener - el valor de α tal que se minimice $F(X)$ en la dirección d ,

$$\alpha = \arg \min \{F(\bar{X} + \alpha d) : 0 < \alpha \leq \alpha_M\} \quad (2.5)$$

Es preciso notar que el valor de α que satisfa- face las condiciones GPW para $\alpha < \infty, \eta = 0, \mu = 0$ es precisamente el valor de α obtenido en (2.5). para d_M suficientemente grande. La resolución del problema (2.5), aunque no es tan compleja como la minimización de F , todavía tiene una gran dificultad, excepto en casos muy particu- lares.

Alternativamente a la obtención de α mediante métodos exactos, existen los métodos denominados línea de búsqueda aproximada. La tipología más frecuentemente utilizada se basa en la obtención de α mediante ajustes polinomiales y su posterior modificación mediante técnicas heurísticas; ver una amplia panorámica en /4/, y el método más utilizado en -- /15/. Otro tipo de método que, aunque a veces requiere mayor número de evaluaciones de la función F, es muy útil en cierto tipo de problemas y, en concreto, en la obtención de la amplitud de paso multivariante (ver Sección 4), básicamente consiste en la siguiente operativa (ver e.g. /1/):

(1) $\gamma: 0$ si $\alpha_M \geq 1$; en caso contrario $\gamma = \lceil \ln \alpha_M / \ln \beta \rceil$; donde $\lceil a \rceil = a$ si a es entero y, en caso contrario, $\lceil a \rceil$ recoge el valor entero inmediatamente superior.

(2) $\alpha := \beta^\gamma$ para $0 < \beta < 1$

(3) Comprobar la condición (2.4c) tal que se incrementa γ en una unidad y se ejecuta de nuevo el paso (2) si la condición (2.4c) no se ha cumplido; en caso contrario, el valor de α es la amplitud de paso en la iteración correspondiente.

Es preciso notar que en la operativa anterior el escalador γ es el primer valor $0, 1, 2, \dots$ que preserva la factibilidad de la solución y satisface la condición (2.4c). En /1/ se demuestra que existe una combinación de β y μ , tal que hay un α que reduce suficientemente la función objetivo para una dirección de búsqueda que satisfaga la condición (2.1) y caso no degenerado (i.e., el menor valor de α que satisface la condición (4.a) es suficientemente pequeño). Valores típicos, $\beta = 0.5$ y $\mu = 0.1$.

La convergencia local de cualquier algoritmo encuadrado en la metodología arriba brevemente descrita, no depende del valor concreto de la amplitud de paso que satisfaga e.g. las condiciones GPW; el ritmo de convergencia depende de la dirección de búsqueda, d tal que la secuencia $\{X\}$ converge a X^* con un ritmo r si

$$0 < \limsup \frac{\|X - X^*\|}{\|X - X^*\|} \equiv \Omega < \infty$$

El ritmo de convergencia es lineal para $r = 1$

(siendo $0 < \Omega < 1$) y cuadrático para $r = 2$. La convergencia será Q-superlineal para $\Omega = 0$ y $r = 1$; el valor de r para el que resulte que $0 < \Omega < \infty$ (si existe) define el ritmo de convergencia.

Recientemente se han diseñado algoritmos -- (ver la operativa correspondiente en e.g. /3/, /14/) en los que se sustituye la utilización de la amplitud de paso por la condición de que la dirección de búsqueda, d se restrinja a una denominada región de confianza. En concreto, el problema a resolver en una determinada iteración consiste en

$$\min\{F(\bar{X} + d) : Ad = 0, U - \bar{X} > d \geq L - \bar{X}\} \quad (2.7a)$$

sujeta a

$$\|d\| \leq \theta \quad (2.7b)$$

tal que el nuevo punto factible será $X: \bar{X} + d$ en lugar de X (2.2), siendo θ un escalador que el algoritmo actualiza dinámicamente.

Aunque todavía no existe suficiente experiencia computacional, parece que la utilización de la amplitud de paso produce mejores resultados que la región de confianza, al menos para problemas de gran escala (cientos de condiciones y miles de variables); por tanto, el nuevo punto de obtiene con la expresión (2.2). Ahora bien, salvo en circunstancias excepcionales, sólo una variable (en el mejor de los casos) que pertenezca al conjunto $\bar{B}\bar{S}$ puede tomar en cada iteración el valor de uno de sus límites y, por tanto, se incluirá en el conjunto \bar{N} ; éste sería el caso en el que la variable, sea $X_j = \bar{X}_j + \alpha d_j$ es tal que el límite α_M de α en (2.3) tiene la expresión

$$\alpha_M = \alpha_j, \quad \alpha_j = \{\min_{k \in \bar{B}\bar{S}} \alpha_k\} \quad (2.8a)$$

donde

$$\alpha_k = |(a_k - \bar{X}_k) / d_k| \quad (2.8b)$$

siendo a_k el límite activo en la dirección del signo de d_k donde

$$a_k = U_k |d_k > 0, \quad a_k = L_k |d_k < 0 \quad (2.8c)$$

Por tanto, si e.g. 500 variables que en la solución factible inicial \bar{X} pertenecen al conjunto $\bar{B}\bar{S}$, deben pertenecer al conjunto \bar{N} en la solución óptima X^* , entonces resulta que serían necesarias, al menos, 500 itera--

ciones para lograr obtener el óptimo con independencia del grado de sofisticación del método utilizado para obtener la dirección de búsqueda. Mejores resultados se podrían obtener permitiendo que se active más de una variable en cada iteración, en este caso, la amplitud de paso no es un escalar, sino que está representada por un vector, sea $[\alpha]^{\#}$ donde el elemento k-ésimo recoge la amplitud a utilizar en d_k para $k \in \bar{B}\bar{S}$ en la iteración dada, tal que

$$X_k := \bar{X}_k + [\alpha]_k^{\#} d_k \quad \forall k \in \bar{B}\bar{S} \quad (2.9)$$

En este trabajo se describe la metodología a utilizar para obtener la amplitud de paso multivariante $[\alpha]^{\#}$, tal que el nuevo punto X (2.9) sea factible y reduzca suficientemente la función objetivo; el ritmo de convergencia no resulta modificado y depende de la metodología a utilizar para la obtención de la dirección de búsqueda, d; ver Sección 6.

Es preciso notar que las variables deactivadas en una iteración pueden convertirse en superbásicas en alguna iteración posterior, tanto si se utiliza una estrategia multivariante como si se utiliza una estrategia univariante. No es posible predecir "a priori" en qué estrategia se desactivan mas variables. Intuitivamente, parece que es más interesante activar más de una variable en cada iteración, si ello implica una mayor reducción en la función objetivo. No hemos encontrado ningún caso correspondiente a situaciones reales con estructura reticular (ver ---/11/) en la matriz A, en el que el tiempo de computo del punto X^* sea mayor cuando se utiliza la estrategia multivariante que cuando se utiliza la estrategia univariante.

3. CARACTERIZACION DE LA MATRIZ Z EN PROBLEMAS DE GRAN ESCALA.

Siguiendo un criterio clásico en problemas de gran escala /19/, la partición de la matriz A tal que $A=(BS N)$ se sustituye por la partición $A=(B S N)$, donde B es la matriz m.m básica, S es la matriz m.s superbásica y N es la matriz m.n no-básica. La dirección de búsqueda, d será factible si se cumple la condición

$$Ad \equiv (B S N) \begin{pmatrix} d_B \\ d_S \\ d_N \end{pmatrix} = 0 \quad (3.1)$$

donde la dirección no-básica d_N es $d_N=0$ por definición, la dirección básica d_B se utiliza para satisfacer la condición de factibilidad $Ad=0$, y la dirección superbásica d_S debe ser tal que el nuevo punto $X=\bar{X}+\alpha d$ reduzca su ficiientemente el valor de la función objetivo F. Sean \bar{B} , \bar{S} y \bar{N} los conjuntos básico, superbásico y no-básico de variables que corresponden a las matrices B, S y N, respectivamente, donde $m=|\bar{B}|$, $s=|\bar{S}|$ y $\bar{n}=|\bar{N}|$. (Es preciso notar que $L_j < \bar{X}_j < U_j \quad \forall j \in \bar{S}$, y $L_j = \bar{X}_j$ o $U_j = \bar{X}_j$ para $\forall j \in \bar{N}$; la partición $A=(B S N)$ se corresponde con la partición $A=(B N)$ en programación lineal, ya que en ésta no hay más de m variables que en el óptimo tomen un valor intermedio entre sus límites.

Dado que $d_N=0$ y d_S es libre en (3.1), resulta que

$$d_B = -B^{-1} S d_S \equiv -\rho d_S \quad (3.2)$$

tal que la caracterización, denominada reducción de variables, de la matriz \bar{Z} tiene la expresión

$$\bar{Z} = \begin{pmatrix} -\rho \\ I \\ 0 \end{pmatrix} \quad (3.3)$$

donde

$$\rho = B^{-1} S$$

es la matriz m.s de proyección básica-superbásica, I es la matriz s.s unitaria, y 0 es la matriz n.s nula.

El problema (reducido) de minimización de F en función del conjunto $\bar{B}\bar{S}\Delta\bar{B}\cup\bar{S}$, tiene la siguiente aproximación cuadrática

$$\min \left\{ \bar{h}^t d + 1/2 d^t \bar{H} d \right\} \quad (3.5)$$

tal que (1.5) y (1.9) recogen las expresiones del gradiente reducido \bar{h} y de la matriz Hessiana reducida \bar{H} . Denotando $\bar{g}^t = (\bar{g}_B^t, \bar{g}_S^t, \bar{g}_N^t)$ resulta que para \bar{Z} expresada por (3.3), \bar{h} tiene las siguientes expresiones alternativas

$$\bar{h} = \bar{g}_S^t - \rho^t \bar{g}_B^t \quad (3.6)$$

$$\bar{h} = \bar{g}_S^t - S^t \nu_B \quad (3.7a)$$

donde μ_B resuelve el sistema lineal de ecuaciones

$$\bar{g}_B = B^t \mu_B \quad (3.7b)$$

La obtención de la dirección superbásica, d_S , suficientemente descendente depende de la estructura de la matriz A y, por tanto, de la estructura de la matriz ρ . En cualquier caso, el vector \bar{h} precisa ser calculado utilizando alternativamente,

(a) La fórmula (3.6), tal que

$$\bar{h}_k = \bar{g}_k - \sum_{k' \in \beta_k} \rho_{k'k} \bar{g}_{k'} \quad \forall k \in \bar{S} \quad (3.8)$$

donde $\rho_{k'k}$ recoge el elemento relativo a la fila k' y columna k de la matriz ρ , y β_k es el conjunto, denominado conjunto equivalente básico, que recoge las variables básicas con elementos no-nulos en el vector columna ρ_k de la matriz ρ . La fórmula (3.8) es útil cuando la matriz ρ es poco densa; ver en /11/ /12/, /13/ un caso típico en el que los elementos no-nulos de ρ son ± 1 con una estructura reticular.

(b) La fórmula (3.7a) y la resolución del sistema (3.7b), tal que se utiliza la formulación LU de la matriz B; ver /7/. Se recomienda esta formulación en casos con estructura general en la matriz A; ver /7/, /8/, /9/ /10/, /19/.

La dirección d_S denominada dirección gradiente, tal que $d_S = -\bar{h}$

$$d_S = -\bar{h} \quad (3.9)$$

sólo precisa información de primer orden y, por tanto, su convergencia es cuasi-lineal. La dirección d_S denominada dirección Newton, tal que

$$d_S = -\bar{H}^{-1} \bar{h} \quad (3.10)$$

utiliza información de segundo orden. Es preciso destacar que, aún en el caso de que las matrices \bar{G} y \bar{Z} sean de poca densidad, la matriz \bar{H} (1.9) es normalmente muy densa; salvo en casos especiales, o en problemas de dimensión pequeña, su uso es prohibitivo. En un punto intermedio entre las direcciones gradiente y Newton existe una amplia gama de tipologías para obtener la dirección superbásica; --

ver /5/, /7/, /8/, /9/, /10/, /11/, /12/, /13/, /14/, /15/, /16/, /19/, entre otros.

La formulación alternativa de la dirección básica, d_B será de acuerdo con (3.2),

(a) Caso con estructura especial en la matriz A

$$d_{k'} = - \sum_{k \in \bar{S} | k' \in \beta_k} \rho_{k'k} d_k \quad \forall k' \in \bar{B} \quad (3.11)$$

(b) Caso general. Ver /8/.

$$(1) \underline{d} = S d_S \quad (3.12a)$$

(2) Obtener d_B como solución del sistema de ecuaciones

$$LU d_B = -\underline{d} \quad (3.12b)$$

donde LU es la correspondiente factorización de la matriz B.

En cualquier caso, una vez obtenido un nuevo punto $X = \bar{X} + \alpha d$, es posible que una variable básica o superbásica tome el valor de uno de sus límites. Si la variable, sea k es superbásica, entonces se convierte en no-básica, tal que $\bar{N} \Delta \bar{N} u \{k\}$ y $\bar{S} \Delta \bar{S} \setminus \{k\}$. Si la variable, sea k' es básica, entonces se intercambia -- con una variable superbásica, sea k tal que $\bar{B} \Delta \bar{B} u \{k\} \setminus \{k'\}$, $\bar{S} \Delta \bar{S} \setminus \{k\}$ y $\bar{N} \Delta \bar{N} u \{k'\}$ y, como resultado del pivote (k', k), se actualizan -- las matrices B y S, o alternativamente ρ , y la matriz \bar{H} o su correspondiente aproximación; la metodología en la actualización matricial dependerá de la estructura de la matriz A.

4. AMPLITUD DE PASO MULTIVARIANTE.

La amplitud de paso multivariante (ver 2.2), sólo permite, para el caso en que $\alpha = \alpha_M$ (ver 2.8), que sólo una variable del conjunto $\bar{B} \bar{S}$ -- pueda convertirse en variable activa, salvo -- en el caso excepcional de que se cumpla la -- condición (2.8a) para más de una variable; el esfuerzo computacional requerido (ver Sección 2) para obtener el punto óptimo depende fuertemente de la composición de los conjuntos $\bar{B} \bar{S}$ y \bar{N} en la solución inicial factible. A continuación se describe un algoritmo que permite la activación de más de una variable básica o superbásica en la misma iteración (en /1/, /2/, se proponen también algoritmos de tipo -- multivariante), si se cumple la condición (4.1)

tal que

$$\alpha_s < d_B \quad (4.1)$$

donde

$$\alpha_B = \min\{\alpha_k, \forall k' \in \bar{B}\} \quad (4.2)$$

$$\alpha_s = \min\{\alpha_k, \forall k \in \bar{S}\} \quad (4.3)$$

tal que $\alpha_k, \forall k \in \bar{S}$ es dado por la expresión (2.8b), donde a_k , el límite activo en la dirección del signo de d_k tiene la expresión - (2.8c); en caso contrario (i.e., $\alpha_B \leq \alpha_s$), sugerimos utilizar el método tradicional de línea de búsqueda aproximada (ver Sección 2).

Mediante una pequeña modificación en la expresión (2.8c), sea la siguiente formulación de a_k el límite activo en la dirección del signo contrario de h_k .

$$a_k = U_k |h_k < 0, \quad a_k = L_k |h_k > 0 \quad \forall k \in \bar{S} \quad (4.4)$$

tal que el conjunto, sea \bar{Q} de variables --- cuasi-activas en el conjunto \bar{S} tiene la expresión

$$\bar{Q} \Delta \{k \in \bar{S} \mid |a_k - \bar{x}_k| \leq \delta_q\} \quad (4.5)$$

donde la tolerancia δ_q será

$$\delta_q = \min\{\delta, \|\bar{x}_S - [\bar{x}_S - \bar{h}]^\# \|_2\} \quad (4.6)$$

tal que para $\forall k \in \bar{S}$

$$[\bar{x}_S - \bar{h}]^\#_k = \begin{cases} \min\{(\bar{x}_S - \bar{h})_k, a_k\} & \text{si } h_k < -\epsilon \\ (\bar{x}_S)_k & \text{si } |h_k| \leq \epsilon \\ \max\{(\bar{x}_S - \bar{h})_k, a_k\} & \text{si } h_k > \epsilon \end{cases} \quad (4.7)$$

para $\delta > 0$ (valor típico, 0.01) y $\epsilon = \epsilon_M^{1/2}$, - donde ϵ_M es la precisión utilizada en - cálculo con coma flotante (normalmente, --- 10E-15 aprox.). Por tanto, una variable superbásica será cuasi-activa si su valor obtenido en la iteración anterior está en la 'proximidad' de su límite activo en la dirección del signo contrario de su elemento en el gradiente reducido evaluado en dicho punto anterior; en este caso, la dirección $d_{\bar{Q}}$ del conjunto \bar{Q} , obtenida como parte integrante de la dirección superbásica d_S según la metodología referenciada en la Sección 3, se sustituye por la expresión

$$d_k = a_k - \bar{x}_k \quad \forall k \in \bar{Q} \quad (4.8)$$

donde a_k tiene la expresión (4.4).

Se puede observar que $d_{\bar{Q}}$ es una dirección descendente (i.e., $\bar{h}_{\bar{Q}}^t d_{\bar{Q}} < 0$), dado que d_k tiene el signo contrario de $h_k, \forall k \in \bar{Q}$. La influencia del conjunto \bar{Q} en la dirección de - búsqueda d será muy pequeña; probablemente, su influencia también hubiera sido pequeña sin la modificación (4.8); la ventaja de esta modificación consiste en que la obtención tentativa de la dirección básica d_B (ver más abajo) sólo utilizará el conjunto \bar{Q}_n (donde $\bar{Q}_n \Delta \bar{S} \setminus \bar{Q}$) y, por tanto, se reduce el esfuerzo computacional.

El nuevo punto factible $x^t = (x_B^t, x_S^t, x_N^t)$ puede expresarse tal que

$$x_N = \bar{x}_N \quad (4.9a)$$

$$x_S = \bar{x}_S + \langle [\alpha]^\# \rangle d_S \quad (4.9b)$$

$$x_B = x_B + d_B \quad (4.9c)$$

donde $\langle ab \rangle$ es un vector tal que $\langle ab \rangle_k = a_k b_k$; donde ahora, d_B no recoge la dirección básica (3.11) o (3.12), sino que es el desplazamiento básico (i.e., el desplazamiento del punto \bar{x}_B al punto x_B); y $[\alpha]^\#$ es el vector que recoge la amplitud de paso multivariante para el conjunto superbásico \bar{S} , tal que

$$[\alpha]^\#_k = \begin{cases} 1 & \forall k \in \bar{Q} \\ \min\{\alpha, \alpha_k, \forall k \in \bar{Q}_n\} & \end{cases} \quad (4.10)$$

donde α_k es el límite máximo (2.8b) de la amplitud de paso de la variable superbásica k , tal que su límite activo es dado por (2.8c). El desplazamiento básico tiene la expresión

$$d_B = -\rho \langle [\alpha]^\# \rangle d_S \quad (4.11)$$

o alternativamente,

$$d_B = -B^{-1} S \langle [\alpha]^\# \rangle d_S \quad (4.12)$$

El desplazamiento d_B será factible en el sentido de que $U_k \geq x_k \geq L_k, \forall k' \in \bar{B}$, si se cumple que

$$\alpha_{\bar{B}} \geq 1 \quad (4.13a)$$

para

$$\alpha_{\bar{B}} = \min\{\alpha_k, \forall k' \in \bar{B} \mid |d_k| > \epsilon\} \quad (4.13b)$$

tal que se utiliza la expresión (2.8b) para α_k , donde el límite activo es dado por --- (2.8c) pero d_k , es dado, alternativamente, por (4.11) y (4.12).

Sea α la amplitud de paso tentativa, tal -- que

$$\alpha = \beta^\gamma \quad (4.14)$$

donde $0 < \beta < 1$, y el parámetro γ toma el primer valor no-negativo $0, 1, 2, \dots$ que, mediante la utilización de las expresiones (4.9) a --- (4.12), produce un punto X_B (4.9c) factible (i.e., satisface las condiciones $AX=b$ y - (4.13)) y, además, cumple la condición -- (4.15) para $0 < \mu \leq 0.5$; valores típicos, ----- $\beta = 0.5$ y $\mu = 0.1$

$$F(X) \leq F(\bar{X}) + \mu \bar{h}_Q^t d_Q + \mu \bar{h}_{Q_n}^t < [\alpha]^\# d_{Q_n} \quad (4.15)$$

Es preciso notar que $\gamma = 0$ en (4.14) implica que se asume que la dirección superbásica d_{Q_n} es suficientemente 'correcta' como para utilizar $\alpha = 1$ en la primera amplitud de paso tentativa; ahora bien, si $\bar{\alpha}_S < 1$ para

$$\bar{\alpha}_S = \max\{\alpha_k \forall k \in \bar{Q}_n \mid |d_k| > \epsilon\} \quad (4.16)$$

resulta que el primer valor de γ será

$$\gamma = \lceil \ln \bar{\alpha}_S / \ln \beta \rceil \quad (4.17)$$

En la Figura 2 se recogen los tres casos po sibles de la amplitud de paso multivariante para $\bar{S} = 2$.

Si el punto X_B no satisface la condición (4.13) (i.e., $\alpha = \beta^\gamma$ no produce una solución factible), se actualiza el parámetro γ tal que

$$\gamma = \lceil \ln(\alpha_B \alpha) / \ln \beta \rceil \quad (4.18)$$

y se comprueban de nuevo las condiciones - (4.13) y (4.15) para el nuevo valor de α da do por (4.14) con γ dado por (4.18).

5. EXTRAPOLACION DE LA AMPLITUD DE PASO.

Si después de un número dado, sea t_1 de ten tativas en la amplitud de paso α , todavía no se satisfacen las condiciones (4.13) y --

(4.15), se sugiere extrapolar el valor de γ tal que

$$\gamma = \lceil \ln \alpha_S / \ln \beta \rceil \quad (4.19)$$

si β^γ para γ dado por (4.19) es inferior - al último valor analizado de α . Es preciso notar que para $\alpha \leq \alpha_S$, sólo es necesario com probar la condición (4.15), dado que la con dición (4.13) se satisface automáticamente, ya que por definición $\alpha_S < \alpha_B$ y $[\alpha]^\# = \alpha \forall k \in \bar{Q}_n$.

En cualquier caso, se asume que el procedi miento para obtener la amplitud de paso no ha dado los resultados apetecidos si $\alpha \leq \epsilon$ (el punto actual \bar{X} es cuasi-óptimo) o, si des-- pués de un número dado, sea t_2 de tentativas en la amplitud de paso todavía no se satisfi cen las condiciones (4.13) y (4.15). Valores típicos, $t_1 = 2$ ó 3 y $t_2 = 6$ ó 7.

El orden en que se analizan las condiciones (4.13) y (4.15) depende del esfuerzo compu tacional requerido para la evaluación de las expresiones (4.11) ó (4.12) por un lado, y la evaluación de la función objetivo F y la última parte de la condición (4.15) por otro. En problemas con matriz A muy densa y función F muy poco densa y no-linealidad muy suave, es preferible analizar en primer lu gar la condición (4.15). Normalmente, y so bre todo si se utiliza la modalidad definida por la matriz ρ , es preferible analizar la condición (4.13) en primer lugar.

La utilización, alternativa, de las expresio nes (4.11) y (4.12) para obtener d_B depende del criterio en que se basa la utilización, alternativa, de las expresiones (3.6) y (3.7) para obtener \bar{h} . Es preciso indicar que, sal vo en casos particulares, la utilización de la expresión (4.12) requiere un esfuerzo com putacional excesivo, dado que en cada tenta tiva de la amplitud de paso es preciso

(a) Computar $S < [\alpha]^\# d_S >$, sea \underline{d}

(b) Resolver el sistema $LUd_B = -\underline{d}$

El algoritmo ChLBM, Chacón Línea de Búsque da Multivariante (ver el apéndice), en el - cual hemos basado nuestra experiencia compu tacional, asume la disponibilidad de la ma triz ρ ; en caso contrario, la modificación

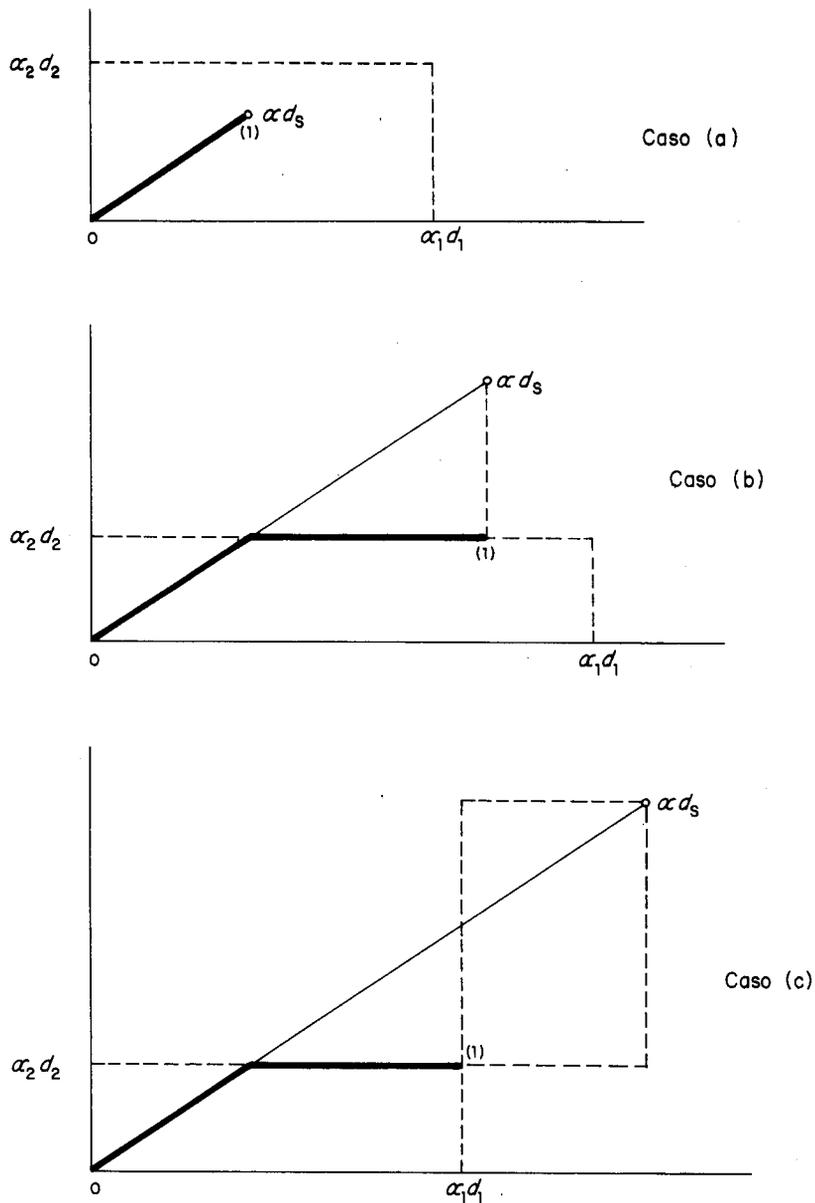


FIGURA 2.- AMPLITUD DE PASO MULTIVARIANTE

(1) : PRIMERA TENTATIVA

Caso (a): $\alpha < \alpha_k \rightarrow [\alpha]_k^* = \alpha$ para $k=1,2$

Caso (b): $\alpha_1 > \alpha > \alpha_2 \rightarrow [\alpha]_1^* = \alpha ; [\alpha]_2^* = \alpha_2$

Caso (c): $\alpha > \alpha_k \rightarrow [\alpha]_k^* = \alpha_k$ para $k=1,2$

es obvia (ver el párrafo anterior), aunque el esfuerzo computacional puede ser excesivo.

6. DISCUSION SOBRE LA CONVERGENCIA.

En primer lugar, es preciso notar que el número de manifolds es finito. Una manifold está definida por los valores de las variables activas (i.e., conjunto \bar{N}). Por tanto, si el algoritmo converge globalmente en la manifold, converge en el problema. El algoritmo converge globalmente una manifold si existe una iteración a partir de la cual se cumple la condición (2.1) para la dirección de búsqueda y la condición de aceptabilidad para la amplitud de paso. No es necesario que estas iteraciones sean consecutivas; si no son consecutivas, sólo se precisa que el valor de la función objetivo del que parte cada iteración no sea superior al valor obtenido en la iteración anterior; éste es el caso en que una manifold se repite después de un número de iteraciones debido a una interrupción prematura en su optimización y, por construcción del algoritmo, se cumple la condición anterior.

Es preciso indicar que no hay amplitud de paso multivariante en la 'cercanía' del valor óptimo de una manifold y, por tanto, el algoritmo no altera el tipo de convergencia local. El razonamiento es el siguiente: Si existe una iteración a partir de la cual $\alpha = 1$ cumple la condición de aceptabilidad de la amplitud de paso, resulta que (ver /20/, /21/ entre otros) que el algoritmo tiene el tipo de convergencia local que tendría el procedimiento utilizado para obtener la dirección de búsqueda si éste se aplicase a la optimización de una función cuadrática sin restricciones. Dado que $\alpha = 1$ implica que $\alpha_S \geq 1$, resulta que el algoritmo propuesto comprueba directamente la condición (4.15) y,

no necesita de la amplitud de paso multivariante. Si el algoritmo tiene convergencia global, siempre existe una iteración a partir de la cual $\alpha = 1$ es aceptable como amplitud de paso.

El algoritmo propuesto tiene convergencia global en la optimización de la manifold si d_S satisface (2.1) y el problema no es degenerado (i.e., el menor valor de α que satisface la condición (2.4a) es suficientemente inferior a α_S). El algoritmo analiza un conjunto de desplazamientos, sea $\{[\alpha]_k, d_k\} \forall k \in \bar{S}$

donde $\{d_k\} = d_S$, tal que $[\alpha]_k \geq \alpha_S > 0$, y, por tanto, el conjunto es descendente; por otra parte, $\exists [\alpha]_k$ igual a α_S (ver Figura 2). En /1/ se demuestra que, independientemente del algoritmo utilizado para la obtención del nuevo punto (ver (4.9), (4.11)-(4.13)), existe una combinación de β y μ tal que el primer valor de la secuencia $\gamma = 0, 1, 2, \dots$ que satisfaga (4.15) produce un desplazamiento que reduce suficientemente la función objetivo. La convergencia global se obtiene pues con independencia del procedimiento utilizado para la obtención del desplazamiento básico.

7. RESULTADOS COMPUTACIONALES.

En /11/ se presenta un algoritmo para resolver un problema del tipo (1.1)-(1.2) en el que la matriz de condiciones A tiene estructura reticular. El prototipo, denominado NLRNET, utilizado para obtener los resultados recogidos en /11/, /12/ tiene como componentes básicos: (a) El algoritmo propuesto para la obtención de la amplitud de paso multivariante; (b) El método de Newton Truncado preconditionado para obtener la dirección superbásica; (c) La identificación y explotación de conjuntos superbásicos independientes; y (d) La utilización de estructuras de datos para obtener la matriz virtual ρ que son clásicas en programación con función objetivo lineal. En /13/ se presentan los resultados de un análisis computacional que analiza la estrategia de conjuntos superbásicos independientes. En esta sección, y al objeto de aislar la influencia de la amplitud de paso multivariante, se presenta el análisis computacional de la utilización del prototipo NLRNET en la resolución de los casos descritos en /11/, pero no se utiliza la estrategia anterior. La Tabla 1 recoge los resultados de nuestro análisis computacional en el que se compara el procedimiento para obtener la amplitud de paso univariante descrito en la Sección 2, y el algoritmo propuesto para la obtención de la amplitud de paso multivariante.

En /1/ se indica que el procedimiento utilizado para obtener la amplitud de paso multivariante identifica rápidamente el conjunto de variables (y valores) activos. Si por ello se entiende que el número de manifolds

TABLA I

RESULTADOS COMPUTACIONALES SOBRE LA AMPLITUD DE PASO

Caso	m	n	Manil	Multivariante		vd	Ratio t
				manix	t (secs.)		
PI	28	64	5	5	2	1.07	1.00
PII	156	312	11	10	25	0.96	1.02
PIII	182	416	6	6	48	0.95	0.98
PIV	208	416	7	6	72	0.92	0.81
PV	208	416	4	5	103	0.97	0.77
PVI	312	416	4	4	85	0.94	0.84
PVII	312	624	6	5	293	0.98	0.75
PVIII	312	572	5	5	84	1.01	0.88
PIX	312	416	5	5	85	0.97	0.73
PX	468	676	11	9	142	0.98	0.42
PXI	520	884	8	7	225	0.91	0.74
PXII	572	728	5	4	201	0.93	0.93
PXIII	780	988	6	6	250	0.94	0.75
PXIV	832	1664	10	8	603	0.96	0.56
PXV	1196	1872	12	10	512	0.95	0.87

m : # condiciones.

n : # variables.

manil: # manifolds en la estrategia con amplitud de paso univariante.

manix: # manifolds en la estrategia con amplitud de paso multivariante.

t: Tiempo de CPU (sec.). PL/I OPT(2), VM/CMS, IBM 4381.

vd: ratio del # variables que se desactivan en la estrategia multivariante en comparación con la estrategia univariante.

Ratio t: ratio del tiempo de CPU requerido con la estrategia multivariante en comparación con la estrategia univariante.

a optimizar y el número de variables a desactivar se reduce fuertemente, nuestra experiencia computacional no evidencia dichos resultados. Esta sólo está basada en el tipo de problemas descritos en /11/.

Ahora bien, el número de variables activadas en las iteraciones en las que hay activación y no hay pivoteo (i.e., sólo se activan variables superbásicas) puede ser muy alto. Cuando el punto está en la 'cercaña' del óptimo de la manifold, no hay activación. Hay iteraciones en las que las variables superbásicas tienen un valor cercano a sus límites y todo el conjunto superbásico se activa en la misma iteración. Hay otras situaciones en las que $|\bar{S}_1|$ y $|\bar{S}_2|$ están en el rango 60-80 y 45-55, respectivamente, donde $|\bar{S}_1|$ y $|\bar{S}_2|$ recogen el número de variables superbásicas de dos iteraciones consecutivas en la misma manifold y a continuación no hay más activación. Ello provoca que se reduzca fuertemente el número de iteraciones y el esfuerzo computacional en cada una de ellas.

Por otra parte, el número de tentativas que requiere cada iteración para obtener una amplitud de paso multivariante que satisfaga la condición (4.15), normalmente, no es superior a 3; el 80% (aprox.) de las iteraciones requiere 2 tentativas y sólo hay 10 iteraciones que hayan requerido 4 tentativas sobre el total de 3687 iteraciones que han sido necesarias para la resolución del conjunto de problemas analizados. Nota: 6 de las 10 iteraciones corresponden a los problemas PXIII-PXV.

8. CONCLUSION.

En este trabajo se presenta un algoritmo para obtener la amplitud de paso multivariante en programación no-lineal con condiciones; la metodología en la que se encuadra el algoritmo particiona el conjunto de variables en básicas, superbásicas y no-básicas; la característica multivariante permite activar varias variables en la misma iteración. Aunque esta línea de investigación es muy prometedora, su campo de aplicabilidad debe ceñirse sólo al caso de condiciones lineales. En el caso no-lineal, la función a utilizar, denominada función de mérito, está integrada por

la función objetivo del problema y por componentes frecuentemente no-diferenciables basados en las condiciones; la utilización del algoritmo propuesto es dificultosa en el mejor de los casos.

Por otra parte, mayor rendimiento puede proporcionar el nuevo concepto de amplitud de paso si la partición de variables en la solución factible inicial se diferencia fuertemente de la partición en la solución óptima; el número de iteraciones ahorrado puede ser muy alto.

La estructura de la matriz de condiciones -- tiene una gran influencia en el esfuerzo computacional exigido; a igualdad en el número de iteraciones, el esfuerzo computacional -- puede variar fuertemente comparando los casos en los que sea y no sea preciso resolver el sistema de ecuaciones incluido en (4.12) en cada tentativa de amplitud de paso. Nuestra experiencia computacional con matrices de condiciones sin estructura especial (y, por tanto, utilizando la expresión (4.12)) no es muy concluyente pero, al menos, indica que hay que ser muy cauto en la utilización del algoritmo propuesto.

En cambio, el algoritmo tiene su máxima operatividad en los casos en los que, dada la estructura especial de la matriz de condiciones, se puede utilizar la matriz ρ (3.4) bien con representación explícita, bien a base de estructuras de datos sin excesivo esfuerzo computacional y utilización de espacio; por tanto, se utiliza la expresión (4.11). Hemos efectuado una extensa experimentación computacional (ver resultados en Sección 7) en el caso en que la matriz de condiciones representa un grafo, tal que en cada columna no hay más de dos elementos no-nulos, siendo éstos ± 1 ; los elementos no-nulos de la matriz ρ son también ± 1 ; se obtiene fácilmente la matriz ρ mediante estructuras de datos reticulares. Aparte de las ventajas que tiene la manipulación de estructuras de datos reticulares (e.g., la posibilidad de descomponer el conjunto básico-superbásico en subconjuntos independientes y, por tanto, con optimizaciones paralelas), la dificultad principal en este tipo de problemas consiste en el alto número de iteraciones con pivoteo (i.e., iteraciones en las que el conjunto básico es activo) y, por tanto, sin posibilidad de mul

ti-activar el conjunto superbásico.

En cualquier caso, menor esfuerzo computacional requerido al utilizar las expresiones (4.11) y (4.12), mayor operatividad de la amplitud de paso multivariante.

9. REFERENCIAS.

- /1/ D.P. BERTSEKAS : "Projected methods for optimization with simple constraints", SIAM Journal of Control and Optimization 20 (1982) 221-226.
- /2/ R.S. DEMBO : "A bending backtracking linesearch for linearly constrained optimization", Yale School of Organization and Management, working paper B74, Universidad de Yale, Illinois, 1984.
- /3/ J.E. DENNIS y R.S. SCHNABEL : "Numerical methods for unconstrained optimization and nonlinear equations", (Prentice-Hall, Englewood Cliffs, New Jersey, 1983).
- /4/ L.C.W. DIXON : "Introduction to numerical optimization" en: L.C.W. DIXON et.al. (eds.), Numerical optimization (Birkhauser, Londres, 1980) 1-29.
- /5/ L.F. ESCUDERO : "On Hessian matrices in unconstrained optimization", IBM Scientific Center report G320-3416, Palo Alto (California), 1980.
- /6/ L.F. ESCUDERO : "Lagrange multipliers estimates for constrained optimization", QUESTIO 5 (1981) 173-186.
- /7/ L.F. ESCUDERO : "An algorithm for large-scale quadratic programming and its extensions to the linearly constrained case", IBM Scientific Center report SCR-01.81, Madrid, 1981.
- /8/ L.F. ESCUDERO : "On diagonally preconditioning the Truncated Newton method for super-scale linearly constrained nonlinear programming", European Journal of Operational Research 17 (1984) 401-414.
- /9/ L.F. ESCUDERO : "On diagonally preconditioning the 2-step BFGS method with accumulated step for linearly constrained nonlinear programming", European Journal of Operational Research 18 (1984) 259-274.
- /10/ L.F. ESCUDERO : "Revisión crítica de la utilización del método del gradiente -- conjugado y extensiones en la minimización de funciones no-lineales con condiciones", QUESTIO 8 (1984) 165-177.
- /11/ F.L. ESCUDERO : "On nonlinear replicated networks", QUESTIO 9 (1985) 55-74.
- /12/ L.F. ESCUDERO : "A motivation for using the Truncated Newton approach in a very large-scale nonlinear network problem", Mathematical Programming Studies 26 (1986) 240-244.
- /13/ L.F. ESCUDERO : "Performance evaluation of independent superbasic sets on nonlinear replicated networks", European Journal of Operational Research 23 (1986) 343-355.
- /14/ R. FLETCHER : "Practical methods of optimization" (J. Wiley, Londres, 1980).
- /15/ P.E. GILL y W. MURRAY : "Safeguard steplength algorithms for optimization using descent methods", National Physical Laboratory report NAC 37, Teddington (UK), 1974.
- /16/ P.E. GILL, W. MURRAY y M.H. WRIGHT : "Practical Optimization" (J. Wiley, Londres, 1981).
- /17/ P.E. GILL, W. MURRAY y M.H. WRIGHT : "A note on a sufficient decrease for a non-derivate steplength", Mathematical Programming 23 (1982) 249-252.
- /18/ A.A. GOLDSTEIN y J.F. PRICE : "An effective algorithm for minimization", Numerische Mathematik 10 (1967) 184-189.
- /19/ B. MURTAGH y M. SAUNDERS : "Large-scale linearly constrained optimization", Mathematical Programming 14 (1978) 41-72.

/20/ M.J.D. POWELL : "The convergence of variable metric methods for nonlinear --- constrained optimization calculations", en : O.L. Mangasarian, R.R. Meyer y S.M. Robinson (eds.), Nonlinear Programming 3 (Academic Press, NY, 1978) 27-63.

/21/ M.J.D. POWELL : "Variable metric methods for constrained optimization", en : A. Bachem, M. Grotchel y B. Korte (eds.), Mathematical Programming : state-of-the-art (Springer-Verlag, Berlin, 1983) 288-311.

/22/ P. WOLFE : "Convergence conditions for ascent methods", SIAM Review 11 (1969) 226-234.

Q1. PENDICE.

ALG. ALITMO ChLBM (Línea de búsqueda multivariante).

Datos: Parámetros: $t_1, t_2, \mu, \beta, \epsilon, \delta, F(X)$

Vectores: $U, L, \bar{X}, \bar{h}, d_B, d_S$

Conjuntos: $\bar{B}, \bar{S}, \bar{N}$ y $\beta_k \quad \forall k \in \bar{S}$

Matrices: B, S, N

Matriz o estructura de datos ρ

/* Obtener vectores $\{a_k\}$ y $\{\alpha_k\} \quad \forall k \in \bar{B} \cup \bar{S} \quad */$

For $\forall k \in \bar{B} \cup \bar{S}$: If $d_k > 0$ then $a_k := U_k$; else $a_k := L_k$

$$\alpha_k := |(a_k - \bar{X}_k) / d_k|$$

$\alpha_S := \min\{\alpha_k \quad \forall k \in \bar{S}\}$; $\alpha_B := \min\{\alpha_k, \quad \forall k \in \bar{B}\}$

If $\alpha_B \leq \alpha_S$ then: utilizar la amplitud de paso univariante, stop */

/* Obtener el conjunto \bar{Q}^* */

For $\forall k \in \bar{S}$: If $\bar{h}_k < -\epsilon$ then $a'_k = U_k$; $[\bar{X}_S - \bar{h}]_k := \min\{(\bar{X}_k - \bar{h}_k), a'_k\}$

If $|h_k| \leq \epsilon$ then $a'_k = \infty$;

$$[\bar{X}_S - \bar{h}]_k := \bar{X}_k$$

If $\bar{h}_k > \epsilon$ then $a'_k = L_k$; $[X_S - h]_k := \max\{(\bar{X}_k - \bar{h}_k), a'_k\}$

$$\delta_q := \min\{\delta, \|X_S - [\bar{X}_S - h]\|_2\}$$

$$\bar{Q}^* := \{k \in \bar{S} \mid |a'_k - \bar{X}_k| \leq \delta_q\}; \quad d_k := a'_k - \bar{X}_k \quad \forall k \in \bar{Q}^*$$

/* Obtener magnitudes fijas */

$$f := F(\bar{X}) + \bar{h}_Q^t d_Q$$

$$d_k := - \sum_{k \in \bar{Q} \mid k' \in \beta_k} \rho_{k', k} d_{k'} \quad \forall k' \in \bar{B} \mid (\exists k \in \bar{Q} \mid k' \in \beta_k)$$

$$X_N := \bar{X}_N$$

$$X_Q := \bar{X}_Q + d_Q$$

$$\bar{Q}_n \Delta \bar{S} \setminus \bar{Q}; \quad \bar{\alpha}_S := \max\{\alpha_k \quad \forall k \in \bar{Q}_n \mid |d_k| > \epsilon\}$$

/* Obtener tentativas amplitud de paso */

If $\bar{\alpha}_S \geq 1$ then $\gamma := 0$; else $\gamma := \lceil \ln \bar{\alpha}_S / \ln \beta \rceil$

t := 0

b1: If $t = t_1, \forall \alpha \leq \epsilon$ then stop: no se puede obtener la amplitud de paso.

If $t = t_2$ then : $\gamma 1 := \lceil \ln \alpha_S / \ln \beta \rceil$

If $\beta^{\gamma 1} < \alpha$ then $\gamma := \gamma 1$

$$\alpha := \beta^\gamma$$

If $\alpha \leq \epsilon$ then go to b1

For $\forall k \in \bar{Q}_n$: $[\alpha]_k^\# := \min\{\alpha, \alpha_k\}$

$$d_k := [\alpha]_k^\# d_k$$

$$\alpha_B := \infty$$

For $k' \in \bar{B} \mid (\exists k \in \bar{Q}_n \mid k' \in \beta_k)$:

$$d_{k'} := d_{k'} - \sum_{k \in \bar{Q}_n \mid k' \in \beta_k} \rho_{k', k} d_k$$

If $|d_{k'}| \leq \epsilon \forall \alpha \leq \alpha_S$ then go to e1

If $d_{k'} > 0$ then $a_{k'} := U_{k'}$; else $a_{k'} := L_{k'}$

$$\alpha_{k'} := |(a_{k'} - \bar{X}_{k'}) / d_{k'}|$$

If $\alpha_{k'} < \alpha_B$ then $\alpha_B := \alpha_{k'}$

e1: end

If $\alpha_B < 1$ then : $t := t+1, \gamma := \lceil \ln(\alpha_B \alpha) / \ln \beta \rceil$; go to b1

$$X_{Q_n} := \bar{X}_{Q_n} + d_{Q_n}$$

$$X_B := \bar{X}_B + d_B$$

Evaluar $F(X)$

If $F(X) \leq f + \mu \bar{h}_Q^t d_{Q_n}$ then stop: X es el nuevo punto

t := t+1, $\gamma := \gamma+1$, go to b1

end.