Estudios de Economía Aplicada
N.º 9, 1998. Págs. 35-60

La dependencia lineal del conjunto de trabajo en el método de conjunto activo que controla la inercia

MANUEL ALBERTO GÓMEZ SUÁREZ
LUIS PEDRO PEDREIRA ANDRADE
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

Una hipótesis básica de los métodos de conjunto activo es la independencia lineal del conjunto de trabajo. Sin embargo, si el conjunto de trabajo contiene restricciones insatisfechas introduce el riesgo de que la restricción alcanzada en una iteración sea dependiente del conjunto de trabajo, a diferencia de lo que ocurre en los métodos primal-factibles. En este trabajo, se consideran el problema de la detección y tratamiento de la dependencia del conjunto de trabajo, la detección de la infactibilidad del programa y la actualización de la dirección de búsqueda ante la adición de una restricción al conjunto de trabajo en presencia de restricciones insatisfechas en éste

Palabras clave: Programación cuadrática, métodos de conjunto activo, métodos que controlan la inercia, métodos de una única fase.

Clasificación AMS (MOS): 90C20, 65K05

ABSTRACT

Linear independence of the working set is a basic assumption in active set methods. However, the working set can include violated restrictions in a single phase method. This fact can cause a restriction added to the working set being linearly dependent of the working set. In this work, it is considered the detection and resolution of linear dependence in the working set, the infeasibility of the quadratic programming problem detection and the search direction updating after a restriction adition in presence of violated restrictions in the working set.

Key words: General quadratic programming, active-set methods, inertia-controlling methods, single phase methods.

Artículo recibido en diciembre de 1997. Revisado en mayo de 1998.

1. Introducción

En este trabajo consideraremos el problema de la dependencia lineal del conjunto de trabajo en el método de conjunto activo que controla la inercia para la solución del programa cuadrático con restricciones de desigualdad (PCD):

$$\begin{array}{ll} \underset{x\in R^n}{minimizar} & q(x) = c^T x + \frac{1}{2} x^T Bx \\ \text{sujeto a :} & Ax \geq b \end{array} \qquad (1)$$

siendo B una matriz simétrica de orden n x n y A una matriz de orden m x n.

Los algoritmos de conjunto activo para la solución de PCD son métodos iterativos basados en la equivalencia de este problema con el programa cuadrático con restricciones de igualdad (PCI) resultante de considerar únicamente las restricciones activas en el óptimo de PCD, que forman el denominado *conjunto activo*. Básicamente, en cada iteración del algoritmo se realiza una predicción del conjunto activo, denominada *conjunto de trabajo*, formada por un conjunto de restricciones linealmente independientes que se tratarán como igualdades en la iteración actual. Si la predicción es correcta, la solución del PCI cuyas restricciones son las incluidas en el conjunto de trabajo es el óptimo de PCD. En caso de que el conjunto de trabajo no sea el conjunto activo, se determina un nuevo punto, un nuevo conjunto de trabajo y se inicia una nueva iteración.

La estrategia de controlar la inercia consiste en determinar el conjunto de trabajo en cada iteración de forma que la hessiana reducida tenga a lo sumo un autovalor no positivo. Para ello es preciso que la eliminación de restricciones se realice únicamente cuando la hessiana reducida es definida positiva.

El primer método que controla la inercia (MCI) fue propuesto por Fletcher (1971). Posteriormente, se desarrollaron otros MCIs, como los métodos de Gill y Murray (1978), Gill, Murray, Saunders y Wright (1991), Gould (1989) y Hoyle (1986).

El método QPSFA propuesto por Hoyle (1986) es un MCI primal de una única fase, en el que la sucesión de iterantes generada no es necesariamente factible. En esta situación, el conjunto de trabajo puede contener restricciones insatisfechas. Como consecuencia, existe la posibilidad de que la restricción alcanzada en una iteración sea linealmente dependiente del conjunto de trabajo. Hoyle (1986) obtuvo una serie de resultados para la detección y solución de este problema, desarrollando la que denominó regla de la singularidad. Una vez alcanzado un punto factible, QPSFA se convierte en un MCI primal-factible, semejante al método de Gill y Murray (1978).

En este trabajo, se consideran las particularidades introducidas en el MCI por el hecho de que la solución de iterantes no sea necesariamente factible, mientras no se alcanza la factibilidad. Especialmente, se analiza la posibilidad de que la restricción alcanzada en una iteración sea dependiente del conjunto de trabajo, y el problema de la detección de la infactibilidad de PCD.

La estructura de este trabajo es la siguiente. En §2 se presenta el método de Hoyle. La detección y tratamiento de la dependencia lineal de la restricción alcanzada del conjunto de trabajo, la actualización de la dirección de búsqueda en esta situación y la detección de la infactibilidad de PCD son objeto de estudio en §3. Finalmente, en §4 se presentan las conclusiones de este trabajo y futuras líneas de investigación.

2. Exposición general del método QPSFA

Sean x el iterante actual, A una matriz de orden t x n y rango t, cuyas filas son las normales de las restricciones incluidas en el conjunto de trabajo, r=Ax-b el vector de residuos en x, y g=Bx+c el gradiente de q en x.

El algoritmo QPSFA propuesto por Hoyle (1986) genera una sucesión de puntos, no necesariamente factibles, donde el sistema empleado para la determinación de la dirección es el mismo, sea o no factible el iterante actual. Por ello, su método es un método de una única fase. Esta denominación se emplea en contraposición a los métodos de conjunto activo primal-factibles que emplean dos fases: una primera de factibilidad, y una segunda de optimalidad. Básicamente, los pasos de QPSFA son los siguientes:

Paso 0: Determine qué restricciones (linealmente independientes) insatisfechas o satisfechas en el punto inicial están en el conjunto de trabajo inicial. Así, las restricciones en el conjunto de trabajo determinan la matriz A de orden t x n ($t \leq n$), y los vectores b y r. Se denotarán con el subíndice "$_I$" a vectores o matrices relativos a restricciones no incluidas en el conjunto de trabajo.

Paso 1: Test de convergencia.

Paso 2: Determine la dirección de búsqueda y el vector de multiplicadores de Lagrange, resolviendo el sistema de Karush-Kuhn-Tucker (KKT):

$$K\begin{pmatrix} -d \\ \lambda \end{pmatrix} = \begin{pmatrix} B & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} d \\ -\lambda \end{pmatrix} = -\begin{pmatrix} g \\ r \end{pmatrix} \qquad (2)$$

Paso 3: O se elimina una restricción o se calcula la longitud de paso:

a) Si se elimina una restricción, actualice K y vuelva al Paso 1.
b) Si se calcula la longitud de paso, vaya al Paso 4.

Paso 4: Actualización: $x \leftarrow x + \alpha d$, $r \leftarrow (1-\alpha)r$, $r_l \leftarrow r_l + \alpha A_l d$. Si $\alpha < 1$, se añade una restricción al conjunto de trabajo y se actualiza K. Vuelva al Paso 1.

Para la solución del sistema KKT (2) en el paso 2, Hoyle (1986) propone el empleo de la factorización TQ de la matriz A:

$$AQ = A\,(\,Z \mid Y\,) = (\,AZ \mid AY\,) = (\,0 \mid T\,) \qquad (3)$$

donde Q es una matriz ortogonal de orden n x n y $T=(t_{ij})$ es una matriz triangular inferior *reflejada* de orden t x t; esto es, $t_{ij}=0$, $\forall i,j \in \{1,\ldots,t\} \,/\, i+j \leq t$, de modo que es la imagen especular de una matriz triangular inferior. Las columnas de la matriz Z de orden n x (n-t) forman una base de $Ker(A)=\{u \in R^n \,/\, Au=0\}$, ya que $AZ=0$ y $rg(A^T \mid Z)=n$. En este caso se dice que Z es una matriz de espacio nulo de A. La matriz Z^TBZ se denomina *hessiana reducida* respecto de A, o símplemente, hessiana reducida.

En caso de que Z^TBZ no sea definida positiva, la hessiana B en el sistema (2) se sustituye por una hessiana modificada $\overline{B}$, definida según:

$$\overline{B} = \begin{cases} B & \text{si } Z^T BZ \text{ es definida positiva} \\ B + \sigma z z^T & en \text{ otro caso} \end{cases} \qquad (4)$$

El escalar $\sigma > 0$ en (4) se escoge de modo que $Z^T \overline{B}\, Z$ sea definida positiva, y z es la última columna de la matriz Z. De este modo, la determinación de la dirección de búsqueda se realiza en cada iteración con una hessiana reducida (quizá modificada) definida positiva, resolviendo el sistema de ecuaciones de espacio nulo equivalente a (2):

$$\begin{aligned} Td_Y &= -r \\ Z^T \overline{B} Z d_Z &= -Z^T g - Z^T BY d_Y \\ d &= Yd_Y + Zd_Z \\ T^T \lambda &= Y^T (g + Bd) \end{aligned} \qquad (5)$$

para cuya solución se emplea la factorización de Cholesky de $Z^T \overline{B}\, Z$.

La eliminación de una restricción en el paso 3 depende de si el iterante actual es factible o infactible respecto de las restricciones incluidas en el conjunto de trabajo. En cualquier caso, para poder realizarse la hessiana reducida ha de ser definida positiva. En un iterante factible es posible eliminar una restricción del conjunto de trabajo y determinar una dirección de búsqueda de descenso y factible respecto de la restricción eliminada siempre que el multiplicador correspondiente sea negativo. En un iterante no factible, sólo es posible asegurar que la nueva dirección es factible respecto de la dirección eliminada si B es definida positiva, de modo que sólo en este caso se permite la eliminación de restricciones en iterantes no factibles respecto del conjunto de trabajo.

La adición de restricciones se realiza cuando una restricción limita el paso en la dirección de búsqueda, y de modo que no se aumente la infactibilidad de ninguna restricción insatisfecha. Así, el máximo tamaño de paso factible en la dirección d se define como:

$$\alpha_F = \min_{i \notin J} \alpha_i \tag{6.a}$$

donde:

$$\alpha_i = \begin{cases} \dfrac{-r_i}{a_i^T d} & \text{si } a_i^T d<0,\ i\notin V(x) \\ 0 & \text{si } a_i^T d\le 0,\ i\in V(x), \qquad i\notin J \\ +\infty & \text{en otro caso} \end{cases} \tag{6.b}$$

y V(x) denota el conjunto de restricciones insatisfechas en el iterante actual. El tamaño de paso se toma como:

$$\alpha = \begin{cases} \min(1,\alpha_F) \text{si } Z^T BZ \text{ es definida positiva o si } r\neq 0 \\ \alpha_F \quad \text{en otro caso} \end{cases} \tag{6.c}$$

De este modo un paso unitario en la dirección de búsqueda si $r\neq 0$, hace activas las restricciones incluidas en el conjunto de trabajo.

2.1. El tratamiento de la dependencia lineal del conjunto de trabajo en QPSFA

En QPSFA se diferencia entre restricciones generales y cotas sobre las variables. Si el número de cotas en el conjunto de trabajo A es n_F, y el número de variables libres es $n_L=t-n_F$, la matriz A_L de orden t x n_L denota la parte de la matriz A correspondiente a las variables libres. Para determinar qué restricción del conjunto de trabajo (si existe) intercambiar por la restricción alcanzada, Hoyle (1986) propone resolver el siguiente sistema de ecuaciones para β:

$$\begin{pmatrix} A_L & h \\ k^T & m \end{pmatrix}\begin{pmatrix} \bar{d} \\ \beta \end{pmatrix} = -\begin{pmatrix} r \\ q \end{pmatrix} \tag{7}$$

donde $\bar{d}$ denota la nueva dirección, excepto la componente correspondiente a la restricción a eliminar, que se denota por β. Los casos son los siguientes:

i) Si se alcanza una cota e_i, entonces $k=e_i$, y $q=0$. De este modo, fuerza que la i-ésima componente de $\bar{d}$ sea igual a cero.

ii) Si se alcanza una restricción general a_i, entonces k consiste en los elementos libres de a_i y $q=r_i$ es el residuo correspondiente. De este modo, el punto $x+\bar{d}$ satisface la restricción i-ésima.

iii) Si se elimina la restricción general a_j del conjunto de trabajo, entonces $h=-e_j$.

iv) Si se elimina la j-ésima cota del conjunto de trabajo, h consistirá en los elementos de la j-ésima columna de A, correspondiente a la variable que se hace libre.

v) Si se alcanza la restricción general a_i, y se elimina la j-ésima cota, entonces m es el elemento de la i-ésima fila y j-ésima columna de A. En otro caso, m es cero.

En lugar de resolver para β un sistema de la forma (7) para cada restricción en el conjunto de trabajo, Hoyle (1986) muestra cómo simplificar su determinación empleando la factorización TQ de la matriz A_L:

$$A_L Q_L = A_L (Z_L | Y_L) = (A_L Z_L \,|\, A_L Y_L) = (0 \,|\, T_L)$$

Para ello se determina la solución u del sistema:

$$k^T Y_L = u^T T_L \tag{8}$$

y a continuación, se determina un valor de β para cada restricción en el conjunto de trabajo según:

$$\beta = \frac{q - u^T r}{u^T h - m} \tag{9}$$

El vector u sólo necesita ser calculado una vez para todos los β. Si el denominador es no nulo, β existe, de modo que el nuevo conjunto de trabajo sería independiente. Si β es positivo (suponiendo que la restricción eliminada j-ésima está en su cota inferior), la nueva dirección de búsqueda sería factible con respecto a la restricción eliminada. En consecuencia, se ha determinado una restricción a eliminar si β tiene el signo correcto.

Sin embargo, tras la resolución del sistema (8) y la determinación de un valor de b para cada restricción, aún restaría determinar la nueva dirección de búsqueda.

Además, la dirección de búsqueda se ha de determinar en cada iteración resolviendo el sistema (2) con una hessiana reducida (quizá modificada) $Z^T\bar{B}Z$ definida positiva, para que se pueda aplicar el teorema 2 sobre la detección de un problema infactible.

En el siguiente apartado, desarrollaremos los resultados precisos para la detección de la dependencia lineal del conjunto de trabajo, simplificando la determinación de una restricción conveniente del conjunto de trabajo para intercambiar por la restricción alcanzada y la detección de la infactibilidad de PCD. Además, la única suposición sobre la determinación de la dirección de búsqueda es que ésta sature las restricciones incluidas en el conjunto de trabajo.

3. La dependencia lineal del conjunto de trabajo

Una particularidad introducida por el hecho de que la sucesión de iterantes no sea necesariamente factible consiste en que la restricción alcanzada en una iteración pueda ser linealmente dependiente del conjunto de trabajo. En los MCIs primal-factibles que generan direcciones de espacio nulo, los sucesivos conjuntos de trabajo son linealmente independientes. Esto es así porque la dirección de búsqueda verifica que $Ad=0$, siendo A el conjunto de trabajo actual. Si se añade una restricción a_q al conjunto de trabajo, ello es porque $a_q^T d<0$. Si a_q fuese dependiente del conjunto de trabajo actual, tendríamos que $a_q=A^Ty$. Multiplicando ambos términos por d^T, resultaría una contradicción, pues:

$$d^T a_q = d^T A^T y = 0$$

Si el conjunto de trabajo actual contiene restricciones insatisfechas en el iterante actual, de modo que $r\neq 0$, la restricción alcanzada podría ser dependiente del conjunto de trabajo. En este caso, la dirección de búsqueda verifica que $Ad=-r$, de modo que:

$$d^T a_q = d^T A^T y = -d^T r$$

con lo que no se incurre, necesariamente, en la contradicción señalada con anterioridad.

Una hipótesis básica del método de conjunto activo es la independencia lineal del conjunto de trabajo. Si la adición de una restricción hace que el nuevo conjunto de trabajo sea dependiente, se debe determinar una restricción para eliminar de modo que el nuevo conjunto de trabajo sea independiente y la nueva dirección de búsqueda sea factible respecto de la restricción eliminada.

3.1. *Notación y resultados previos*

Supongamos que A es una matriz de orden t x n, t<n, tal que rg(A)=t, y que la matriz A* se diferencia de A únicamente en la última fila, que sería la normal a_N^T, de modo que:

$$A\ * = \begin{pmatrix} A \\ a_N^T \end{pmatrix} \qquad (9)$$

Supondremos que Z es una matriz de orden n x (n-t) de espacio nulo de A, de modo que AZ=0 y rg(A^T|Z)=n. La matriz Z* de orden n x (n-t-1) es una matriz de espacio nulo de A*; esto es, A*Z*=0 y rg(A^{*T}|Z*)=n. Por la estructura de A*, podemos tomar Z*=(Z|z). Las matrices KKT correspondientes a los conjuntos de trabajo A y A* serán, respectivamente:

$$K = \begin{pmatrix} B & A^T \\ A & 0 \end{pmatrix} \qquad (10)$$

$$K^* = \begin{pmatrix} B & A^{*T} \\ A^* & 0 \end{pmatrix} = \begin{pmatrix} B & A^T & a_N \\ A & 0 & 0 \\ a_N^T & 0 & 0 \end{pmatrix} \qquad (11)$$

Lema 1.(Haynsworth (1968)) Sea la matriz simétrica M de orden m x m, particionada como:

$$M = \begin{pmatrix} P & Q^T \\ Q & R \end{pmatrix}$$

donde la matriz cuadrada P es regular. Entonces, se verifica que:

$$In(M) = In(P) + In(M/P)$$

donde $M/P = R - QP^{-1}Q^T$.

A la matriz $M/P = R - QP^{-1}Q^T$ se le denomina *complemento de Schur* de P en M.

El siguiente resultado muestra la relación entre la inercia de K y la inercia de la hessiana proyectada $Z^T BZ$.

Lema 2. (Gould (1985, Lema 3.4)) Si A tiene rango pleno por filas t, se verifica que:

$$In(K) = In(Z^T BZ) + (t, t, 0)$$

3.2. *La detección de la singularidad en la matriz KKT*

La relación entre las inercias de dos matrices KKT sucesivas tras la adición de una restricción al conjunto de trabajo viene dada por el siguiente resultado.

Lema 3. Si K es regular, entonces:

$$In(K^*) = In(K) + In(-z^T a_N)$$

siendo z la parte apropiada de la solución del sistema

$$K\begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} B & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} a_N \\ 0 \end{pmatrix} \qquad (12)$$

Demostración. Por (12) y la regularidad de K, se verifica que:

$$\begin{pmatrix} z \\ y \end{pmatrix} = K^{-1}\begin{pmatrix} a_N \\ 0 \end{pmatrix}$$

El resultado se sigue de aplicar el *lema 1*:

$$In(K^*) = In(K) + In(K^*/K)$$

teniendo en cuenta que:

$$K^*/K = (0) - \begin{pmatrix} a_N^T & 0 \end{pmatrix} K^{-1}\begin{pmatrix} a_N \\ 0 \end{pmatrix} = -\begin{pmatrix} a_N^T & 0 \end{pmatrix}\begin{pmatrix} z \\ y \end{pmatrix} = -z^T a_N$$

A continuación, consideraremos las posibles fuentes de singularidad de la matriz KKT y cómo distinguirlas, tras la adición de una restricción al conjunto de trabajo.

Lema 4. Sea z la parte apropiada de la solución del sistema (12). Si K es regular, entonces K* es singular si y sólo si $a_N^T z = 0$. Además, si K* es singular:

a) z=0 si y sólo si $a_N \in rg(A^T)$,
b) $z \neq 0$ si y sólo si $Z^{*T}BZ^*$ es singular.

Demostración. La primera parte se sigue directamente del *lema 3*.

a) Si z=0, entonces de (12) resulta que $A^T y = a_N$(donde $y \neq 0$), de modo que $a_N \in rg(A^T)$. Por otra parte, si $a_N \in rg(A^T)$ entonces $\exists y \neq 0$ tal que $A^T y = a_N$. Por ser K regular, el sistema (12) admite una única solución, que será (0,y), como se comprueba por sustitución. En consecuencia, z=0.

b) Si $z \neq 0$, por a), A* tiene rango pleno. Por el *lema 2*,

$$In(K^*) = In(Z^{*T} BZ^*) + (t+1, t+1, 0)$$

En consecuencia, si K* es singular, también $Z^{*T}BZ^*$ lo es.

El *lema 4* muestra que, si K es regular, tras la adición de una restricción la nueva matriz K* puede ser singular por dos causas: la singularidad de la hessiana proyectada $Z^{*T}BZ^*$, o la dependencia lineal del conjunto de trabajo A*. Además, proporciona un método para distinguir entre ambos casos.

3.3. *El tratamiento de la dependencia lineal del conjunto de trabajo*

Una hipótesis básica de los métodos de conjunto activo es la independencia lineal del conjunto de trabajo. Si se alcanza una restricción linealmente dependiente del conjunto de trabajo actual, no puede ser añadida inmediatamente al conjunto de trabajo, sino que antes habrá que eliminar una restricción de éste. La restricción a eliminar se elegirá de modo que el nuevo conjunto de trabajo sea independiente, y la nueva dirección de búsqueda sea factible respecto de la restricción eliminada.

Supongamos que se alcanza una restricción a_N en la dirección d, y el tamaño de paso es $\alpha = \alpha_F < 1$. El siguiente teorema considera la situación en la que la restricción alcanzada a_N es dependiente del conjunto de trabajo actual. Si

$$A = P\begin{pmatrix} A_S \\ a_E^T \end{pmatrix}$$

es el conjunto de trabajo actual, donde P una matriz de permutación apropiada, denominaremos

$$\bar{A} = \begin{pmatrix} A_S \\ a_N^T \end{pmatrix}$$

al nuevo conjunto de trabajo, resultado de eliminar una restricción a_E del antiguo, y añadir la restricción alcanzada a_N. El vector de residuos correspondiente será:

$$\bar{r} = \begin{pmatrix} \bar{r}_S \\ \bar{r}_N \end{pmatrix}$$

donde $\bar{r}_N = r_N + \alpha a_N^T d$ y $\bar{r}_S = (1-\alpha)\, r_S$.

Teorema 1.

Supongamos que el conjunto de trabajo A es linealmente independiente, la dirección de búsqueda d verifica que

$$Ad = -r \Leftrightarrow \begin{pmatrix} A_S \\ a_E^T \end{pmatrix} d = -\begin{pmatrix} r_S \\ r_E \end{pmatrix} \qquad (13)$$

y se alcanza la restricción a_N, con un paso $\alpha < 1$, definido por (6). Supongamos que a_N es linealmente dependiente del conjunto de trabajo A; esto es, existe un vector $y \in R^t$ tal que:

$$a_N = A^T y = A_S^T y_S + a_E y_E \qquad (14)$$

Si $y_E > 0$, el conjunto de trabajo $\bar{A}$ resultante de eliminar a_E y añadir a_N al conjunto de trabajo, es linealmente independiente. Además, si la nueva dirección de búsqueda se escoge de modo que

$$\bar{A}\bar{d} = -\bar{r} \Leftrightarrow \begin{pmatrix} A_S \\ a_N^T \end{pmatrix} \bar{d} = -\begin{pmatrix} \bar{r}_S \\ \bar{r}_N \end{pmatrix} \qquad (15)$$

entonces, el punto $\bar{x}+\bar{d}$ satisface estrictamente la restricción a_E; esto es,

$$a_E^T \bar{d} > -\bar{r}_E$$

Demostración. Si a_N es linealmente dependiente del conjunto de trabajo A, por la independencia lineal de A existe un vector $y \neq 0$ único tal que se verifica (14). Puesto que $y_E \neq 0$, es posible intercambiar a_N por a_E, y el conjunto resultante sigue siendo independiente, de modo que el conjunto de trabajo $\bar{A}$ es linealmente independiente.

Denominaremos $\bar{x} = x + \alpha d$, y

$$\tilde{r} = \begin{pmatrix} \bar{r}_S \\ \bar{r}_E \end{pmatrix} = A(x+\alpha d) - b = r - \alpha r = (1-\alpha) r = \begin{pmatrix} (1-\alpha) r_S \\ (1-\alpha) r_E \end{pmatrix} \qquad (16)$$

O bien $\alpha>0$, y entonces $r_N=0$ (se alcanza la restricción a_N antes inactiva), o bien $\alpha=0$, en cuyo caso $\bar{r}_N=0$ si la restricción a_N era activa y no estaba en el conjunto de trabajo, o $r_N<0$ si a_N es una restricción insatisfecha. En cualquier caso, $r_N\leq 0$.

Multiplicando (14) por $\bar{d}^T$ y empleando (15), tenemos que, por una parte

$$\bar{d}^T A^T y = \bar{d}^T a_N = -\bar{r}_N$$

y por otra, empleando (14),

$$\bar{d}^T A^T y = \bar{d}^T A_S^T y_S + \bar{d}^T a_E y_E = -\bar{r}_S^T y_S + \bar{d}^T a_E y_E$$

Igualando ambas expresiones y puesto que, por hipótesis, $y_E\neq 0$, tenemos que:

$$d^T a_E = \frac{\bar{r}_S^T y_S - \bar{r}_N}{y_E} \tag{17}$$

Empleando (16), tenemos que:

$$\bar{r}^T y = (1-\alpha) r^T y = \bar{r}_S^T y_S + \bar{r}_E y_E$$

Empleando (13) y (14), por ser $\alpha<1$, resulta que

$$\bar{r}^T y=(1-\alpha)r^T y=-(1-\alpha)d^T A^T y=\begin{cases}(1-\alpha)\overbrace{d^T a_N}^{>0} > 0, \text{ si } a_N \text{ es satisfecha}\\ (1-\alpha)\overbrace{d^T a_N}^{\geq 0} \geq 0, \text{ si } a_N \text{ es insatisfecha}\end{cases}$$

de modo que

$$\bar{r}_S^T y_S > -\bar{r}_E y_E \geq 0, \text{si } a_N \text{ es satisfecha} \tag{18.a}$$

$$\bar{r}_S^T y_S \geq -\bar{r}_E y_E \geq 0, \quad \text{si } a_N \text{ es insatisfecha} \tag{18.b}$$

pues $-\bar{r}_E y_E = -(1-\alpha) r_E y_E \geq 0$. Empleando (17) y (18), tenemos que:

$$a_E^T d = \frac{\bar{r}_S^T y_S - \bar{r}_N}{y_E} \geq \frac{\bar{r}_S^T y_S}{y_E} > -\bar{r}_E \quad \text{si } a_N \text{ es satisfecha}$$

$$a_E^T \bar{d} = \frac{\bar{r}_S^T y_S - \bar{r}_N}{y_E} > \frac{\bar{r}_S^T y_S}{y_E} \geq -\bar{r}_E \quad \text{si } a_N \text{ es violada}$$

En cualquier caso, resulta que:

$$a_E^T \bar{d} > -r_E$$

de modo que con la dirección así determinada, el punto $\bar{x}+\bar{d}$ satisface estrictamente la restricción eliminada a_E.

Supongamos que se alcanza una restricción dependiente, y que se determina una restricción del conjunto de trabajo para intercambiarla por la restricción alcanzada, empleando el teorema 1. El siguiente lema muestra que si las columnas de la matriz Z forman una base de Ker(A), también forman una base de Ker($\bar{A}$). Por lo tanto, el intercambio de restricciones no afecta a la inercia de la hessiana reducida, de modo que la estrategia de controlar la inercia no se ve afectada.

Lema 5. Supongamos que las columnas de la matriz Z forman una base de Ker(A). En las hipótesis del *teorema 1*, si se ha determinado una restricción conveniente para eliminar a_E, las columnas de la matriz Z también forman una base de Ker($\bar{A}$).

Demostración. La matriz Z verifica que AZ=0. Puesto que $A^Ty=a_N$, tenemos que

$$A_S Z = 0, \quad a_N^T Z = y^T AZ = 0 \quad \Rightarrow \quad \bar{A}Z = 0$$

de modo que las columnas de la matriz Z también forman una base de Ker($\bar{A}$).

En consecuencia, tras el intercambio de restricciones, la matriz Z no varía, de modo que la hessiana reducida tampoco lo hace: $Z^T BZ = Z^T BZ$.

Si se realiza un tratamiento diferenciado de las restricciones generales y las cotas, denominaremos C al conjunto de trabajo actual, particionado según:

$$C = \begin{pmatrix} A_L & A_F \\ 0 & I_F \end{pmatrix} \qquad (19)$$

Las filas de la matriz $A=(A_L|A_F)$ son las normales de las restricciones generales incluidas en el conjunto de trabajo actual, particionada en la parte A_L correspondiente a las variables *libres* y la parte A_F correspondiente a las variables *fijas* en el valor de la cota correspondiente incluida en el conjunto de trabajo.

En este caso, particionamos el vector a_N según:

$$a_N = \begin{pmatrix} a_L \\ a_F \end{pmatrix}$$

Entonces, calcularemos el vector "y" de (12), según:

$$\begin{pmatrix} 0 & A_L^T \\ I_F & A_F^T \end{pmatrix}\begin{pmatrix} y_C \\ y_G \end{pmatrix}=\begin{pmatrix} a_L \\ a_F \end{pmatrix} \Leftrightarrow \begin{aligned} A_L^T y_G &= a_L \\ y_C &= a_F - A_F^T y_G \end{aligned} \tag{20}$$

donde y_G es la parte del vector "y" correspondiente a restricciones generales en el conjunto de trabajo, e y_C es la parte del vector "y" correspondiente a las cotas incluidas en el conjunto de trabajo.

El siguiente resultado muestra que la hessiana reducida correspondiente a las nuevas variables libres, coincide con la antigua. Si Z_L es una matriz de espacio nulo de A_L, se verifica que

$$Z=\begin{pmatrix} Z_L \\ 0 \end{pmatrix}$$

es una matriz de espacio nulo de C, pues:

$$CZ=\begin{pmatrix} A_L & A_F \\ 0 & I_F \end{pmatrix}\begin{pmatrix} Z_L \\ 0 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{21}$$

Lema 6. En las hipótesis del teorema 1, supongamos que C es el conjunto de trabajo actual, se intercambia la restricción alcanzada dependiente por una restricción incluida en C, y que $\overline{C}$ es el nuevo conjunto de trabajo. Entonces, se verifica que $\overline{Z}_L^T \overline{B}_L \overline{Z}_L = Z_L^T B_L Z_L$, donde $\overline{B}_L$ es la parte de la hessiana B, correspondiente a las nuevas variables libres.

Demostración. Por el *lema 5*, las columnas de la matriz Z forman una base de Ker($\overline{C}$).

Si se alcanza la restricción general a_N, la particionaremos según

$$a_N^T = \begin{pmatrix} a_L^T & a_F^T \end{pmatrix}$$

Si se alcanza la cota e_N, la particionaremos según $e_N^T = \begin{pmatrix} e_L^T & 0 \end{pmatrix}$.

CASO 1:

Intercambio de dos restricciones generales. Puesto que $C^T y = a_N$, observando (21), tenemos que:

$$A_L Z_L = 0, \qquad a_N^T Z = a_L^T Z_L = y^T CZ = 0$$

de modo que $\overline{A}_L Z_L = 0$, y las columnas de la matriz Z_L también forman una base de Ker($\overline{A}_L$). Puesto que $\overline{B}_L = B_L$ (ya que las variables libres y fijas no han variado), resulta que:

$$\overline{Z}_L^T \overline{B}_L \overline{Z}_L = Z_L^T B_L Z_L$$

CASO 2:

Intercambio de dos cotas. Como $\overline{A}_L = A_L$, las columnas de Z_L también forman una base de Ker($\overline{A}_L$). Puesto que $C^T y = e_N$, tenemos que

$$A_L Z_L = 0, \qquad e_N^T Z = e_L^T Z_L = y^T C Z = 0$$

Además, se intercambian las filas t_E-ésima y t_N-ésima de B, de modo que:

$$\overline{B}_L = B_L + (b_L - b_E)e_L^T + e_L(b_L^T - b_E^T) - b_{dd} e_L e_L^T$$

donde b_L es la parte *libre* de la columna añadida (la t_N-ésima) y b_E denota la parte *libre* de la columna eliminada (la t_E-ésima) de B_L, y b_{dd} es el t_N-ésimo elemento de b_L-b_E .

Ahora,

$$\overline{Z}_L^T \overline{B}_L \overline{Z}_L = Z_L^T B_L Z_L + Z_L^T (b_L - b_E) e_L^T Z_L + Z_L^T e_L (b_L^T - b_E^T) Z_L - b_{dd} Z_L^T e_L e_L^T Z_L = Z_L^T B_L Z_L$$

CASO 3:

Se añade una restricción general y se elimina una cota. Puesto que se elimina una cota, una variable fija se hace libre. Supondremos, por simplicidad, que es la primera variable fija la que se hace libre (en otro caso, se reordenan las variables fijas). Ahora, (21) resulta:

$$CZ = \left(\begin{array}{c|c} A_L & A_F \\ \hline 0 & I_F \end{array}\right) \begin{pmatrix} Z_L \\ 0 \end{pmatrix} = \left(\begin{array}{cc|c} A_L & a & A_F \\ 0 & 0 & \overline{I}_F \\ \hline 0 & 1 & 0 \end{array}\right) \left(\begin{array}{c} Z_L \\ 0 \\ \hline 0 \end{array}\right) = 0$$

Para el nuevo conjunto de trabajo, tenemos que:

$$\overline{C}\overline{Z} = \begin{pmatrix} \overline{A}_L & \overline{A}_F \\ 0 & \overline{I}_F \end{pmatrix} \begin{pmatrix} \overline{Z}_L \\ 0 \end{pmatrix} = \overline{C}Z = \begin{pmatrix} a_L^T & \alpha & a_F^T \\ A_L & a & A_F \\ 0 & 0 & \overline{I}_F \end{pmatrix} \begin{pmatrix} Z_L \\ 0 \\ 0 \end{pmatrix} = 0$$

de modo que las columnas de $\overline{Z}_L = \begin{pmatrix} Z_L \\ 0 \end{pmatrix}$ forman una base de Ker($\overline{A}_L$).

La nueva parte libre de la hessiana se forma orlando por la fila y la columna correspondiente a la variable que se ha hecho libre:

$$B_L = \begin{pmatrix} B_L & b \\ b^T & \beta \end{pmatrix}$$

y la nueva hessiana reducida *libre* sería:

$$\overline{Z}_L^T \overline{B}_L Z_L = \begin{pmatrix} Z_L^T & 0 \end{pmatrix} \begin{pmatrix} B_L & b \\ b^T & \beta \end{pmatrix} \begin{pmatrix} Z_L \\ 0 \end{pmatrix} = Z_L^T B_L Z_L$$

CASO 4:

Se añade una cota y se elimina una restricción general. Puesto que se añade una cota, una variable antes libre ahora se hace fija. Supondremos, por simplicidad, que es la última variable libre la que se hace fija (en otro caso, se reordenarían las variables libres). El sistema (21) sería ahora:

$$CZ = \begin{pmatrix} A_L & A_F \\ 0 & I_F \end{pmatrix} \begin{pmatrix} Z_L \\ 0 \end{pmatrix} = \left(\begin{array}{cc|c} a_L^T & \alpha & a_F^T \\ A_L & a & A_F \\ \hline 0 & 0 & I_F \end{array} \right) \begin{pmatrix} Z_L \\ 0 \end{pmatrix} = 0$$

Para el nuevo conjunto de trabajo, tenemos que:

$$\overline{C}\overline{Z} = \begin{pmatrix} A_L & A_F \\ 0 & I_F \end{pmatrix} \begin{pmatrix} Z_L \\ 0 \end{pmatrix} = \begin{pmatrix} A_L & a & A_F \\ 0 & 0 & I_F \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} Z_L \\ 0 \\ 0 \end{pmatrix} = \overline{C}Z = 0$$

de modo que

$$Z_L = \begin{pmatrix} \overline{Z}_L \\ 0 \end{pmatrix}$$

Puesto que una variable antes libre, ahora se hace fija, la nueva parte libre de la hessiana, sería:

$$B_L = \begin{pmatrix} B_L & b \\ b^T & \beta \end{pmatrix}$$

y la nueva hessiana reducida *libre*:

$$Z_L^T B_L Z_L = \begin{pmatrix} \overline{Z}_L^T & 0 \end{pmatrix} \begin{pmatrix} B_L & b \\ b^T & \beta \end{pmatrix} \begin{pmatrix} \overline{Z}_L \\ 0 \end{pmatrix} = \overline{Z}_L^T \overline{B}_L \overline{Z}_L$$

En consecuencia, en todos los casos, se verifica que $\overline{Z}_L^T \overline{B}_L \overline{Z}_L = Z_L^T B_L Z_L$.

Los lemas 5 y 6 indican que si se mantiene una factorización de la hessiana reducida y se produce un intercambio de restricciones en el conjunto de trabajo, fruto de la aplicación de la regla de la singularidad, no es necesario actualizar la factorización de Cholesky (quizá modificada) de la hessiana reducida, pues ésta no varía. Además, muestran que la estrategia de controlar la inercia no se ve afectada por el intercambio de restricciones, pues la inercia de la hessiana reducida no varía.

3.4. *La detección de la infactibilidad de PCD*

El siguiente teorema demuestra que, en las condiciones del teorema 1, si $y \le 0$, entonces el problema PCD es infactible.

Teorema 2.

En las hipótesis del teorema 1, si $y \le 0$, entonces el problema PCD es infactible.

Demostración. Uno de los teoremas de la alternativa (véase Rockafellar (1970, p. 201)), indica que de las dos alternativas siguientes sólo se cumple una:

a) $\exists x$ tal que $Gx \ge h$,

b) $\exists w \le 0$ tal que $G^T w = 0$, $w^T h < 0$.

Tenemos que $A^T y = a_N$, de modo que $\exists w = \begin{pmatrix} y \\ -1 \end{pmatrix} \le 0$, tal que $\begin{pmatrix} A^T & a_N \end{pmatrix} \begin{pmatrix} y \\ -1 \end{pmatrix} = 0$,

donde $w \le 0$, puesto que $y \le 0$. Además, empleando (18), tenemos que:

$$\tilde{r}^T y > 0, \quad \text{si } a_N \text{ es satisfecha}$$

$$\tilde{r}^T y \ge 0, \quad \text{si } a_N \text{ es insatisfecha}$$

de modo que, en cualquier caso:

$$\tilde{r}^T y - \bar{r}_N > 0 \quad \Leftrightarrow \quad \begin{pmatrix} y^T & -1 \end{pmatrix} \begin{pmatrix} -\tilde{r} \\ \\ -\bar{r}_N \end{pmatrix} = -y^T \tilde{r} + \bar{r}_N < 0$$

Identificando en las alternativas anteriores, $G = \begin{pmatrix} A \\ \\ a_N^T \end{pmatrix}$, $h = \begin{pmatrix} -\tilde{r} \\ \\ -\bar{r}_N \end{pmatrix}$, resulta que no existe un vector d tal que

$$Ad \geq -\tilde{r}$$
$$a_N^T d \geq -\bar{r}_N$$

de modo que el problema PCD es infactible.

De los teoremas 1 y 2, se sigue que es posible intercambiar la restricción dependiente alcanzada por una restricción a_E del conjunto de trabajo de modo que la nueva dirección de búsqueda sea factible respecto de la restricción eliminada, si y sólo si la correspondiente componente del vector y es positiva; esto es, $y_E > 0$. Si no es posible efectuar este intercambio, porque $y \leq 0$, el problema es infactible.

Hoyle (1986, teorema 2) demuestra un resultado similar al teorema 2 anterior, derivado a partir de su regla de la singularidad. Sin embargo, en el resultado allí expuesto se supone que en el sistema (2) para la determinación de la dirección de búsqueda, la matriz hessiana B se sustituye por la matriz definida en (4), de forma que $Z^T \bar{B} Z$ sea definida positiva. El teorema 1 no exige que la determinación de búsqueda se realice mediante el sistema (2). Sólo se supone que la dirección de búsqueda en cada iteración satura las restricciones incluidas en el conjunto de trabajo; esto es, que Ad=-r. En consecuencia, existe más libertad para la determinación de la dirección de búsqueda que la empleada por Hoyle (1986).

3.5. *La detección de la dependencia del conjunto de trabajo*

Si se emplea la factorización RQ del conjunto de trabajo, la dependencia de la restricción alcanzada del conjunto de trabajo puede ser detectada al realizar su actualización. Si se distingue entre restricciones generales y cotas, el conjunto de trabajo sería C definido en (21). Si tenemos la factorización RQ de la matriz A_L:

$$A_L Q_L = A_L (Z_L \mid Y_L) = (A_L Z_L \mid A_L Y_L) = (0 \mid R_L)$$

donde R_L es una matriz triangular superior de orden $n_L x n_L$, se obtiene la factorización RQ del conjunto de trabajo según:

$$CQ=\begin{pmatrix} A_L & A_F \\ 0 & I_F \end{pmatrix}\begin{pmatrix} Z_L & Y_L & 0 \\ 0 & 0 & I_F \end{pmatrix}=\begin{pmatrix} A_L Z_L & A_L Y_L & A_F \\ 0 & 0 & I_F \end{pmatrix}=\begin{pmatrix} 0 & R_L & A_F \\ 0 & 0 & I_F \end{pmatrix}=\begin{pmatrix} 0 & R \end{pmatrix}$$

donde R es una matriz triangular superior de orden txt.

Si se añade una restricción general de normal $a_N^T=\begin{pmatrix} a_L^T & a_F^T \end{pmatrix}$al conjunto de trabajo, se coloca en la primera posición, de modo que:

$$\overline{A}_L Q_L=\begin{pmatrix} a_L^T \\ A_L \end{pmatrix}\begin{pmatrix} Z_L & Y_L \end{pmatrix}=\begin{pmatrix} a_L^T Z_L & a_L^T Y_L \\ A_L Z_L & A_L Y_L \end{pmatrix}=\begin{pmatrix} w_Z^T & w_Y^T \\ 0 & R_L \end{pmatrix} \quad (22)$$

Para reconstruir la factorización RQ de la matriz $\overline{A}_L$ ha de reducirse el vector w_Z a un múltiplo no nulo de e_{n-t}:

$$w_Z^T \Omega_Z=\begin{pmatrix} 0 & \gamma \end{pmatrix}=\gamma e_{n-t}^T$$

Entonces, tendríamos la factorización RQ de $\overline{A}_L$ de:

$$\overline{A}_L \overline{Q}_L=\overline{A}_L Q_L \Gamma_L=\begin{pmatrix} w_Z^T & w_Y^T \\ 0 & R_L \end{pmatrix}\begin{pmatrix} \Omega_Z & 0 \\ 0 & I_L \end{pmatrix}=\begin{pmatrix} 0 & \gamma & w_Y^T \\ 0 & 0 & R_L \end{pmatrix}=\begin{pmatrix} 0 & \overline{R}_L \end{pmatrix}$$

de modo que:

$$\overline{A}_L \overline{Q}_L=\overline{A}_L Q_L \Gamma_L=\begin{pmatrix} w_Z^T & w_Y^T \\ 0 & R_L \end{pmatrix}\begin{pmatrix} \Omega_Z & 0 \\ 0 & I_L \end{pmatrix}=\begin{pmatrix} 0 & \gamma & w_Y^T \\ 0 & 0 & R_L \end{pmatrix}=\begin{pmatrix} 0 & \overline{R}_L \end{pmatrix} \quad (23)$$

Sin embargo, si w_Z es nulo, no es posible realizar el proceso anterior, de modo que la restricción añadida es dependiente de las anteriores (la nueva R_L sería singular).

Para determinar la restricción a eliminar del conjunto de trabajo (si hay alguna restricción adecuada), se resuelve el sistema (20). Multiplicando la primera ecuación de (20) por Y_L^T, resulta el sistema equivalente:

$$\begin{aligned} R_L^T y_G &= w_Y \\ y_C &= a_F - A_F^T y_G \end{aligned} \quad (24)$$

pues $Y_L^T a_L = w_Y$ (véase (22)).

Si se añade una cota al conjunto de trabajo, la t_N-ésima, el nuevo conjunto de trabajo sería $\bar{C}=\begin{pmatrix} C \\ e_N^T \end{pmatrix}$, de modo que:

$$\bar{C}Q=\begin{pmatrix} A_L & A_F \\ 0 & I_F \\ e_L^T & 0 \end{pmatrix}\begin{pmatrix} Z_L & Y_L & 0 \\ 0 & 0 & I_F \end{pmatrix}=\begin{pmatrix} A_LZ_L & A_LY_L & A_F \\ 0 & 0 & I_F \\ e_L^TZ_L & e_L^TY_L & 0 \end{pmatrix}=\begin{pmatrix} 0 & R_L & A_F \\ 0 & 0 & I_F \\ q_Z^T & q_Y^T & 0 \end{pmatrix}$$

donde el vector definido por

$$q_L=\begin{pmatrix} q_Z \\ q_Y \end{pmatrix}$$

es la t_N-ésima fila de la matriz Q_L, correspondiente a la variable hecha fija; esto es, q_L^T es la fila t_N-ésima de Z_L, y q_L^T es la fila t_N-ésima de Y_L. En consecuencia, si q_z es nula, la cota añadida es combinación lineal de las restricciones en el conjunto de trabajo.

En este caso, (20) es equivalente al sistema:

$$\begin{aligned} R_L^T y_G &= q_Y \\ y_C &= a_F - A_F^T y_G \end{aligned} \qquad (25)$$

Por lo tanto, en cualquier caso, la dependencia lineal del conjunto de trabajo puede ser detectada antes de actualizar las matrices Q_L y R_L.

3.6. *Actualización de la dirección de búsqueda*

En este apartado, por simplicidad, trataremos todas las restricciones como generales. Ante la presencia de restricciones insatisfechas en el conjunto de trabajo, ya se ha señalado que los teoremas 1 y 2 sólo exigen que la dirección de búsqueda haga activas las restricciones incluidas en el conjunto de trabajo. Supongamos que se divide la dirección de búsqueda d en dos componentes:

$$d = d^Z + d^Y$$

La componente de espacio nulo d^Z es tal que $Ad^Z=0$. La componente de factibilidad d^Y verifica que $Ad^Y=-r$, sistema que es equivalente a:

$$\begin{aligned} \bar{R}d_Y &= -\bar{r} \\ d^Y &= Y\bar{d}_Y \end{aligned} \qquad (26)$$

En consecuencia, la dirección de búsqueda d^Y ya está en las hipótesis de los teoremas 1 y 2, y podría escogerse la componente d^Z de forma arbitraria. Por ello se introduce mayor libertad de elección en la dirección de búsqueda que en QPSFA, mientras alguna restricción en el conjunto de trabajo sea insatisfecha.

Supongamos que se realiza un movimiento, de tamaño $\alpha<1$, en la dirección de búsqueda $d=d^Z+d^Y$, y que se añade la restricción a_N al conjunto de trabajo. Los siguientes resultados muestran cómo llevar a cabo la actualización del vector d^Y, de modo que verifique que:

$$\begin{aligned} \bar{R}\bar{d}_Y &= -r \\ \bar{d}^Y &= \bar{Y}\bar{d}_Y \end{aligned} \tag{27}$$

donde r es el vector de residuos de las restricciones incluidas en el nuevo conjunto de trabajo A.

Lema 7. (Adición de restricciones: caso independiente) Supongamos que las filas de A y la nueva restricción a_N son linealmente independientes, y que esta restricción se añade al conjunto de trabajo con un tamaño de paso $\alpha<1$, para dar el nuevo conjunto de trabajo

$$\bar{A} = \begin{pmatrix} a_N^T \\ A \end{pmatrix}$$

La nueva matriz $\bar{Y}$ sería $\bar{Y} = \begin{pmatrix} y & Y \end{pmatrix}$ (véase (23)), y el nuevo factor triangular del conjunto de trabajo es:

$$\bar{R} = \begin{pmatrix} \tau & t^T \\ 0 & R \end{pmatrix}$$

Si d^Y verifica (26), y se define el escalar:

$$\gamma = \frac{-\bar{r}_N - (1-\alpha)t^T d_Y}{\tau} \tag{28}$$

entonces, el vector

$$\bar{d}^Y = \gamma y + (1-\alpha)d^Y \tag{29}$$

satisface (27).

Demostración. Podemos expresar (27) como:

$$\bar{R}\bar{d}_Y = \begin{pmatrix} \tau & t^T \\ 0 & R \end{pmatrix}\begin{pmatrix} \delta_Y \\ \Delta_Y \end{pmatrix} = -\begin{pmatrix} \bar{r}_N \\ (1-\alpha)r \end{pmatrix} \tag{30}$$

de modo que:

$$R\Delta_Y = -(1-\alpha)r \tag{31.a}$$

$$\delta_Y = \frac{-\bar{r}_N - t^T\Delta_Y}{\tau} \tag{31.b}$$

Empleando (26), obtenemos que

$$\Delta_Y = (1-\alpha)d_Y \tag{32}$$

con lo que, de (30) y (31), y empleando la definición (28), tenemos que:

$$\bar{d}_Y = \begin{pmatrix} \dfrac{-\bar{r}_N - (1-\alpha)t^T d_Y}{\tau} \\ (1-\alpha)d_Y \end{pmatrix} = \begin{pmatrix} \gamma \\ (1-\alpha)d_Y \end{pmatrix} \tag{33}$$

Multiplicado (33) por $\bar{Y}$, resulta (29), ya que:

$$\bar{d}^Y = \bar{Y}\bar{d}_Y = \begin{pmatrix} y & Y \end{pmatrix}\begin{pmatrix} \gamma \\ (1-\alpha)d_Y \end{pmatrix} = \gamma y + (1-\alpha)Yd_Y = \gamma y + (1-\alpha)d^Y$$

Si la restricción alcanzada es dependiente del conjunto de trabajo, es necesario intercambiarla por una restricción incluida en el conjunto de trabajo. En este caso, tenemos el siguiente resultado.

Lema 8. (Adición de restricciones: caso dependiente) Supongamos que las filas de A y la restricción alcanzada a_N son linealmente dependientes, de modo que existe un vector "y" tal que:

$$A^T y = \begin{pmatrix} A_S^T & a_E \end{pmatrix}\begin{pmatrix} y_S \\ y_E \end{pmatrix} = a_N \tag{34}$$

y que, aplicando el *teorema 1*, se intercambia a_N por la restricción a_E.

Sean $\bar{A} = \begin{pmatrix} a_N^T \\ A_S \end{pmatrix}$ el nuevo conjunto de trabajo, y su factorización RQ:

$$\bar{A}\bar{Q} = (\bar{A}\bar{Z} \quad \bar{A}\bar{Y}) = (0 \quad \bar{R})$$

Si d^Y verifica (25), y se define el escalar:

$$\eta = -\bar{r}_N + (1-\alpha)r^T y \tag{35}$$

entonces, el vector

$$\bar{d}^Y = (1-\alpha)d^Y + \frac{\eta}{\tau}\bar{Y}^1 \tag{36}$$

verifica (27), donde τ denota el primer elemento diagonal de $\bar{R}$, e $\bar{Y}^1$ denota la primera columna de $\bar{Y}$.

Demostración. La solución d^Y de (26), verifica que:

$$Ad^Y = -r \Leftrightarrow \begin{pmatrix} A_S d^Y \\ a_E^T d^Y \end{pmatrix} = -\begin{pmatrix} r_S \\ r_E \end{pmatrix} \tag{37}$$

La nueva $\bar{d}^Y$, solución de (27), ha de verificar que:

$$\bar{A} \quad \bar{d}^Y = -\bar{r} \Leftrightarrow \begin{pmatrix} a_N^T \quad \bar{d}^Y \\ A_S \bar{d}^Y \end{pmatrix} = -\begin{pmatrix} r_N \\ \bar{r}_S \end{pmatrix} = -\begin{pmatrix} \bar{r}_N \\ (1-\alpha)r_S \end{pmatrix} \tag{38}$$

Supongamos que

$$\bar{d}^Y = (1-\alpha)d^Y + \delta \tag{39}$$

Sustituyendo (39) en (38), resulta:

$$\bar{A} \quad \bar{d}^Y = \begin{pmatrix} (1-\alpha)a_N^T \quad d^Y + a_N^T \quad \delta \\ (1-\alpha)A_S d^Y + A_S\delta \end{pmatrix} = \begin{pmatrix} (1-\alpha)a_N^T d^Y + a_N^T \quad \delta \\ -(1-\alpha)r_S + A_S\delta \end{pmatrix} = -\begin{pmatrix} \bar{r}_N \\ (1-\alpha)r_S \end{pmatrix}$$

de modo que

$$\begin{aligned} a_N^T \quad \delta &= -\bar{r}_N - (1-\alpha)a_N^T \quad d^Y \\ A_S\delta &= 0 \end{aligned} \tag{40}$$

Empleando (34) y (37), tenemos que:

$$a_N^T \quad d^Y = y^T A d^Y = -y^T r$$

de modo que, para calcular d, se resuelve el sistema:

$$\begin{matrix} a_N^T \quad \delta = -\bar{r}_N + (1-\alpha) r^T y \\ A_S \delta = 0 \end{matrix} \Leftrightarrow \bar{A}\,\delta = \begin{pmatrix} a_N^T \\ A_S \end{pmatrix} \delta = \begin{pmatrix} \eta \\ 0 \end{pmatrix} \qquad (41)$$

donde η viene definido por (35). Empleando la factorización RQ de $\bar{A}$, una solución a (41) puede obtenerse como:

$$\begin{aligned} \bar{R}\delta_Y &= \eta e_1 \\ \delta &= \bar{Y}\delta_Y \end{aligned} \qquad (42.a,b)$$

Puesto que $\bar{R}$ es triangular superior, el cálculo de d_Y se reduce a una única división:

$$\delta_Y = \frac{\eta}{\tau} e_1 \qquad (43)$$

donde τ es el primer elemento diagonal de $\bar{R}$.
Sustituyendo en (42.b) y después en (39), tenemos que:

$$\bar{d}^Y = (1-\alpha) d^Y + \bar{Y}\delta_Y = (1-\alpha) d^Y + \frac{\eta}{\tau} \bar{Y}^1$$

donde $\bar{Y}^1$ denota la primera columna de $\bar{Y}$.
Para comprobar que, efectivamente, $\bar{d}^Y$ es solución de (27); o sea, que es combinación lineal de las columnas de Y, mostraremos que $\bar{Z}^T \bar{d}^Y = 0$. Puesto que, $Z^T \bar{d}^Y = 0$ si y sólo si $\bar{d}^Y \in rg(A^T)$, y como las columnas de $\bar{Y}$ forman una base ortonormal de $rg(\bar{A}^T)$, ya que $\bar{Z}^T \bar{Y} = 0$, tendremos también que $\bar{d}^Y \in rg(\bar{Y})$.
Por el *lema 5*, podemos tomar $\bar{Z}$=Z, de modo que, empleando (36):

$$\bar{Z}^T \bar{d}^Y = Z^T \bar{d}^Y = (1-\alpha) Z^T d^Y + \frac{\eta}{\tau} Z^T \bar{Y}^1 = 0$$

ya que $Z^T d^Y = Z^T Y d_Y = 0$, y $Z^T \bar{Y} = 0$.

4. Conclusiones

En este trabajo se ha estudiado el problema de la dependencia lineal del conjunto de trabajo en el MCI de una única fase. Los resultados obtenidos permiten una mayor libertad en la determinación de la dirección de búsqueda, mientras no se alcance la factibilidad respecto del conjunto de trabajo, que la desarrollada en el método QPSFA: únicamente es preciso que la dirección de búsqueda sature las restricciones incluidas en el conjunto de trabajo. Una futura línea de investigación será cómo aprovechar de la forma más eficiente posible esta mayor libertad de elección de la dirección de búsqueda.

También se han presentado una *regla de la singularidad* distinta a la propuesta por Hoyle (1986), y fórmulas para la actualización eficiente de la que hemos denominado componente de factibilidad de la dirección de búsqueda. Además del interés puramente teórico de este trabajo, resta efectuar una implementación práctica del MCI de una única fase, con las modificaciones aquí expuestas, que compruebe si este método es competitivo en la práctica con los métodos ya existentes para programación cuadrática.

5. Bibliografía

FLETCHER, R. (1971): "A general quadratic programming algorithm", Journal of the Institute of Mathematics and its Applications 7, pp. 76-91.

GILL, P. E. Y MURRAY, W. (1978): "Numerically stable methods for quadratic programming", Mathematical Programming 14, pp. 349-372.

GILL, P. E.; MURRAY, W.; SAUNDERS, M. A. Y WRIGHT, M. H. (1984): "Procedures for optimization problems with a mixture of bounds and general linear constraints", ACM Transactions on Mathematical Software 10, pp. 282-298.

GILL, P. E.; MURRAY, W.; SAUNDERS, M. A. Y WRIGHT, M. H. (1991): "Inertia-controlling methods for general quadratic programming", SIAM Review 33, pp. 1-36.

GILL, P. E.; MURRAY, W. Y WRIGHT, M. H. (1981): Practical Optimization, Academic Press, London.

GÓMEZ, M. A. (1997): "Métodos de conjunto activo que controlan la inercia para la solución de programas cuadráticos generales", Tesis Doctoral, Departamento de Economía Aplicada II, Universidad de A Coruña.

GOULD, N. I. M. (1985): "On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem", Mathematical Programming 32, pp. 90-99.

GOULD, N. I. M. (1989): "An algorithm for large-scale quadratic programming", Report CSS 219, AERE, Atomic Energy Research Establishment, Harwell, UK.

HAYNSWORTH, E. V. (1968): "Determination of the inertia of a partitioned Hermitian matrix", Linear Algebra and its Applications 1, pp. 63-78.

HOYLE, S. C. (1986): "A single-phase method for quadratic programming", Report SOL 86-9, Department of Operations Research, Stanford University, Stanford, California.

MURRAY, W. (1971): "An algorithm for finding a local minimum of an indefinite quadratic program", Report NAC 1, National Physical Laboratory, United Kingdom.

ROCKAFELLAR, R. T. (1972): Convex Analysis, Princenton University Press, Princenton, New Jersey.