

MAT

Serie 

Conferencias, Seminarios
y Trabajos de Matemática.

ISSN: 1515-4904

19

VII Italian

Latin American

Conference on

Industrial and

Applied Mathematics

First Part

Domingo A. Tarzia (Ed.)

Departamento
de Matemática,
Rosario,
Argentina
Octubre 2014

UNIVERSIDAD AUSTRAL

FACULTAD DE CIENCIAS EMPRESARIALES



MAT

SERIE A: CONFERENCIAS, SEMINARIOS Y TRABAJOS DE MATEMÁTICA

No. 19

VII ITALIAN - LATIN AMERICAN CONFERENCE ON INDUSTRIAL AND APPLIED MATHEMATICS

First Part

Domingo A. Tarzia (Ed.)

INDICE

Graciela M. Croceri - Graciela N. Sottosanto, “Comparación de dos algoritmos de lagrangiano aumentado para el problema de optimización con restricciones de igualdad”, pp. 1-7.

Martha Hilda Timoteo Sánchez - Yolanda Santiago Ayala, “Sobre la existencia y unicidad de la solución de un modelo de propagación del sonido en un fluido compresible”, pp. 9-16.

Carlos Andrés Trujillo-Salazar - Hernán Darío Toro-Zapata, “Análisis teórico de la transmisión y el control del VIH/SIDA en un centro de reclusión”, pp. 17-26.

Santiago C. Rojas Romero, “Sobre el espectro de Fucik para un sistema acoplado”, pp. 27-34.

Manuel Maurette, “A topological approach to the repulsive central motion problem”, pp. 35-42.

Rosario, Octubre 2014

COMPARACIÓN DE DOS ALGORITMOS DE LAGRANGIANO AUMENTADO PARA EL PROBLEMA DE OPTIMIZACIÓN CON RESTRICCIONES DE IGUALDAD

G. M. Croceri[†] y G.N. Sottosanto[†]

[†]*Departamento de Matemática, Universidad Nacional del Comahue, Santa Fe 1400, 8300 Neuquén, Argentina,
graciela.croceri@faea.uncoma.edu.ar, graciela.sottosanto@faea.uncoma.edu.ar*

Resumen: En este trabajo se presentan dos algoritmos basados en la minimización secuencial del Lagrangiano Aumentado para resolver el problema de optimización con restricciones de igualdad. Uno de los algoritmos combina una técnica de gradiente conjugado y región de confianza. El segundo utiliza como estrategia de resolución un método de gradiente proyectado, y la convergencia desde cualquier punto inicial, se obtiene mediante un esquema tipo Armijo. Un estimado del multiplicador de Lagrange se actualiza en cada iteración y el parámetro de penalización se modifica para obtener suficiente reducción en la norma de las restricciones. La información de segundo orden, en ambos algoritmos, se actualiza mediante una técnica secante. Los algoritmos fueron implementados en Scilab y se reportan resultados numéricos sobre un conjunto de problemas test de la colección CUTer.

Palabras clave: *Lagrangiano Aumentado, gradiente conjugado, gradiente proyectado, región de confianza, regla de Armijo.*

2000 AMS Subject Classification: 21A54, 55P54

1. INTRODUCCIÓN

El objetivo de este trabajo es hallar la solución $x^* \in R^n$ del problema general de optimización con restricciones de igualdad y cotas sobre las variables

$$\begin{aligned} \underset{s.a.}{\text{mín}} \quad & f(x) & (1) \\ & c_j(x) = 0, \quad j = 1, \dots, p, \\ & l_i \leq x_i \leq u_i, \quad i = 1, \dots, n, \end{aligned}$$

donde $f : R^n \rightarrow R$ y $c : R^n \rightarrow R$ son funciones dos veces continuamente diferenciables y donde algunas de las componentes del vector l y u de cotas sobre las variables pueden fijarse en $-\infty$ y $+\infty$ respectivamente, significando que no existen cotas sobre esas componentes.

Los dos algoritmos que presentamos para resolver (1) están basados en la minimización secuencial del Lagrangiano Aumentado. El método, desarrollado originalmente para resolver el problema de minimización con restricciones de igualdad [11], fue aplicado después a la resolución de problemas de cuadrados mínimos con restricciones de igualdad [6] donde se incorporó una aproximación secante del tipo BFGS [8],[4].

2. LOS ALGORITMOS

El Lagrangiano Aumentado [3] asociado a este problema es ,

$$L_\mu(x, \lambda) = f(x) + \lambda^T c(x) + \frac{1}{2\mu} \|c(x)\|_2^2, \quad (2)$$

donde λ es el vector de multiplicadores de Lagrange y $\mu > 0$ el parámetro de penalización.

El método es iterativo y para la resolución del problema (1) utiliza dos tipos de iteraciones: externas e internas.

En cada iteración externa se encuentra una solución aproximada x_{k+1} de

$$\underset{x}{\text{mín}} L_\mu(x_k, \lambda_k), \quad l_i \leq x_i \leq u_i, \quad i = 1, \dots, n, \quad (3)$$

donde λ_k es el estimado actual del vector de multiplicadores de Lagrange y $\mu > 0$ es el parámetro de penalización. Ambos parámetros se actualizan al final de cada iteración externa.

En cuanto a la iteración interna, en la que se resuelve el subproblema de minimizar la función Lagrangiano Aumentado, se usaron dos algoritmos que utilizan distintas estrategias, a fin de comparar su performance. El que llamaremos, en lo sucesivo, Algoritmo 1 combina una técnica de gradiente conjugado [14] y región de confianza [5]. El Algoritmo 2 utiliza un método de gradiente proyectado con un esquema del tipo Armijo con el objetivo de lograr convergencia desde cualquier punto inicial.

En el caso del Algoritmo 1, si el problema original posee cotas sobre las variables, se transforma en un problema con restricciones de igualdad agregando variables de holgura. Es decir, para cada i con $i = 1, \dots, n$ se reemplaza $l_i \leq x_i \leq u_i$ por $x_i - u_i + s_i^2 = 0$ y $x_i - l_i - w_i^2 = 0$.

2.1. ALGORITMO 1

El subproblema irrestricto se resuelve usando el método de gradiente conjugado que tiene incorporado una estrategia de región de confianza. El paso se obtiene como solución del problema

$$\min_{\|s\| \leq \delta_k} Q_k(s, \lambda_k, \mu_k), \quad (4)$$

donde Q_k es el modelo cuadrático del Lagrangiano Aumentado alrededor de x_k y $\delta_k > 0$ es el radio de la región de confianza. Si la reducción de la función $L_{\mu}(x_k, \lambda_k)$ en el minimizador del modelo cuadrático es suficiente, se acepta el correspondiente paso de prueba como nuevo iterado, caso contrario, se reduce el radio de la región de confianza de manera estándar [5].

Cuando $\|\nabla L_{\mu_k}(x_{k+1}, \lambda_k)\| \leq \epsilon_k$ para alguna sucesión dada $\{\epsilon_k\}$, $\epsilon \rightarrow 0$, la iteración interna termina, con $x_{k+1} = (x_*)_{k+1}$ y el método regresa a la iteración externa.

Con el método de gradiente conjugado [14] se puede hallar la solución aproximada del subproblema cuadrático a través una técnica *dogleg* de tal manera que la curva solución $s(\delta)$ resulta aproximada por una poligonal. La información de segundo orden se aproxima utilizando estrategias secantes en una forma estructurada.

Aún cuando se involucren matrices arbitrarias, el método resulta bien definido y en caso que la aproximación de la matriz Hessiana sea definida positiva se obtiene convergencia local superlineal en el minimizador local del subproblema irrestricto.

Descripción del Algoritmo 1

Supongamos que están dados $x_k \in R^n$, $\lambda_k \in R^p$, $\|\lambda_k\|_2 \leq M$, $M > 0$, constantes y $0 < \mu_{min} \leq \mu_k$. Dadas, además las tolerancias ϵ_1 y $\epsilon_2 > 0$, los siguientes pasos permiten hallar x_* y su multiplicador asociado λ_* .

Algorithm 2.1 (Algoritmo básico de la iteración externa.)

Paso I. Hallar $(x_*)_{k+1}$, solución aproximada del subproblema de minimización

$$\min_x L_{\mu_k}(x, \lambda_k)$$

usando el algoritmo de gradiente conjugado/región de confianza [14].

Paso II. (Test de convergencia)

- Si $\|\nabla l(x_*)_{k+1}, \lambda_k\|_2 \leq \epsilon_1$ y $\|c((x_*)_{k+1})\|_2 \leq \epsilon_2$ entonces TERMINAR, $(x_*)_{k+1} = x_*$ es la solución aproximada del problema de minimización sin restricciones,
- de lo contrario, actualizar el parámetro de penalización tal que $0 < \mu_{k+1} < \mu_k$.
 - Si $\mu_{k+1} \geq \mu_{min}$ actualizar el estimado del multiplicador e ir al paso I,
 - de lo contrario terminar con la solución aproximada $(x_*)_{k+1} \cong x_*$.

2.2. ALGORITMO 2

Este algoritmo combina un método de gradiente proyectado con una regla tipo Armijo [1], [2].

El algoritmo del gradiente proyectado es una extensión natural del algoritmo de dirección de máximo descenso para el caso de problemas con restricciones de cotas.

En este trabajo seguimos una implementación diseñada por [10] que usa una estrategia BFGS recursiva para actualizar la inversa generalizada de una parte de la matriz Hessiana.

El conjunto factible de este problema es el conjunto

$$\Omega = \{x \in R^n : l_i \leq x_i \leq u_i, i = 1, \dots, n\}.$$

Decimos que la i -ésima restricción es activa en x si $(x)_i = l_i$ ó $(x)_i = u_i$. En cada $x \in \Omega$ los conjuntos de índices $\mathcal{A}(x)$ e $\mathcal{I}(x)$ son respectivamente los conjuntos de índices activos e inactivos.

Recordemos que las condiciones de Karush Kuhn y Tucker para el problema (3) establecen que un punto $x_* \in \Omega$ es estacionario para el subproblema si y sólo si

$$x_* = \mathcal{P}(x_* - \alpha \nabla L(x_*))$$

para todo $\alpha \geq 0$, donde \mathcal{P} es la proyección sobre Ω

$$\mathcal{P}(x)_i = \begin{cases} l_i & \text{si } (x)_i \leq l_i \\ (x)_i & \text{si } l_i < (x)_i < u_i \\ u_i & \text{si } (x)_i \geq u_i. \end{cases}$$

El método del gradiente proyectado tiene una propiedad destacada, que consiste en que bajo determinadas condiciones, el conjunto de índices de restricciones activas $\mathcal{A}(x_k)$, a partir de un número determinado k de iteraciones, es el mismo que el que se tiene en la solución $\mathcal{A}(x_*)$. Para asegurar el cumplimiento de esta propiedad, el método tiene que identificar, en cada iteración, el conjunto de restricciones activas.

Dado un iterado actual x_k el nuevo iterado se obtiene como

$$x_+ = \mathcal{P}(x_k - \alpha \nabla L_\mu(x_k)),$$

donde α es un parámetro de longitud de paso correspondiente a una regla tipo Armijo, de manera tal que asegure suficiente decrecimiento en la función objetivo. Para $\alpha > 0$ se define

$$x(\alpha) = \mathcal{P}(x - \alpha \nabla L_\mu(x)).$$

Para el caso de restricciones de cotas sobre las variables, la condición de suficiente decrecimiento se expresa como

$$L_\mu(x(\alpha)) - L_\mu(x) \leq \frac{-\beta}{\alpha} \|x - x(\alpha)\|^2,$$

donde β es un parámetro cuyo valor típico es 10^{-4} .

Como criterio de terminación para nuestro algoritmo elegimos $\|x_k - \mathcal{P}(x_k - \nabla L_\mu(x_k))\| < \epsilon$ para cierta tolerancia ϵ .

En la implementación que usamos en este trabajo, a efectos de mejorar la convergencia del método, en lugar de elegir como dirección de descenso, la opuesta a la del gradiente, se usa un vector d que se actualiza recursivamente premultiplicando por una aproximación secante estructurada de la inversa generalizada de la matriz Hessiana.

Descripción del Algoritmo 2

Supongamos que están dados $x_k \in R^n$, $\lambda_k \in R^p$, $\|\lambda_k\|_2 \leq M$, $M > 0$, constantes y $0 < \mu_{min} \leq \mu_k$. Dadas, además las tolerancias ϵ_1 y $\epsilon_2 > 0$, los siguientes pasos permiten hallar x_* y su multiplicador asociado λ_* .

Algorithm 2.2 (Algoritmo básico de la iteración externa)**Paso I.** Hallar $(x_*)_{k+1}$, solución aproximada del subproblema de minimización

$$\begin{aligned} \min_x \quad & L_\mu(x, \lambda) \\ \text{s.a} \quad & l_i \leq x_i \leq u_i, \quad i = 1, \dots, n. \end{aligned}$$

usando el algoritmo de gradiente proyectado con búsqueda lineal tipo Armijo y actualización BFGS de la inversa de la matriz Hessiana [10].

Paso II. (Test de convergencia)

- Si $\|(x_*)_{k+1} - \mathcal{P}((x_*)_{k+1} - \nabla l((x_*)_{k+1}, \lambda_k))\|_2 \leq \epsilon_1$ and $\|c((x_*)_{k+1})\|_2 \leq \epsilon_2$ entonces TERMINAR,
 $(x_*)_{k+1} = x_*$ es la solución aproximada del problema,
- de lo contrario, actualizar el parámetro de penalización tal que $0 < \mu_{k+1} < \mu_k$.
 - Si $\mu_{k+1} \geq \mu_{min}$ actualizar el estimado del multiplicador e ir al paso I,
 - de lo contrario terminar con la solución aproximada $(x_*)_{k+1} \cong x_{x_*}$.

2.3. ACTUALIZACIÓN DEL PARÁMETRO DE PENALIZACIÓN

Un punto determinante del método de Lagrangiano Aumentado es la elección del parámetro de penalización μ_k . Para asegurar convergencia global se debe generar una sucesión no decreciente $\{\mu_k\}$. Aunque los métodos pueden ser formulados con p diferentes parámetros de penalización, uno para cada componente de $c(x)$, en la práctica, esto no presenta ventajas numéricas.

Para ello, empleamos, en el Algoritmo 1, el siguiente esquema: Si $\|c(x_{k+1})\|_2$ no es suficientemente menor que $\|c(x_k)\|_2$ entonces el parámetro de penalización es disminuido por la regla $\mu_{k+1} = \varphi \mu_k$ donde $\varphi \in (0, 1)$. Aquí consideramos que $\|c(x_{k+1})\|_2$ debe ser menor que $0,05\|c(x_k)\|_2$. El comportamiento del parámetro de penalización es independiente del vector de multiplicadores. Nuestro criterio para actualizar el vector de multiplicadores es mantenerlo constante a lo largo de la minimización cuasi Newton y actualizarlo al final del mismo, por medio de alguna fórmula. En este trabajo, elegimos una fórmula de primer orden para actualizar el vector de multiplicadores de Lagrange,

$$\lambda_{k+1} = \lambda_k + \frac{1}{\mu_k} c_k.$$

Cuando el problema se resuelve aproximadamente utilizando el algoritmo 2, se mide la condición $\|c(x_k)\|_2 < \epsilon$. Si esta se cumple, el parámetro de penalización no se cambia para la siguiente iteración porque su valor actual está produciendo un nivel aceptable de violación de las restricciones y el multiplicador se actualiza con una fórmula de primer orden. Caso contrario, se aumenta el parámetro de penalización para asegurar que la siguiente iteración pondrá más énfasis en la disminución de las restricciones y el multiplicador no se actualiza [13].

3. RESULTADOS NUMÉRICOS

Los algoritmos fueron codificados en Scilab en entorno LINUX. A fin de evaluar su performance, se resolvieron 100 problemas test con restricciones de igualdad y cotas sobre las variables de la colección CUTer [9] y se realizaron dos tipos de comparaciones. En todas ellas los parámetros utilizados en nuestra implementación numérica son los siguientes: $\epsilon_1 = 10^{-5}$ y $\epsilon_2 = 10^{-5}$ y $\mu_{min} = 10^{-10}$. El multiplicador inicial, λ_0 , y el punto inicial, x_0 de cada problema era el proporcionado por CUTer.

Por un lado se evaluó el desempeño de los dos algoritmos planteados y por otro, ambos se compararon con un nuevo algoritmo basado en la estrategia de región de confianza [12]. En este último, en cada iteración, el paso de prueba se obtiene minimizando una aproximación cuadrática de la función Lagrangiano Aumentado en la región de confianza. La función Lagrangiano Aumentado se utiliza también como función

de mérito para decidir si el paso de prueba es aceptado. Un estimado del multiplicador de Lagrange se actualiza en cada iteración, y el parámetro de penalización se modifica para conseguir una reducción suficiente en la norma de la violación de las restricciones. A este algoritmo lo llamaremos Algoritmo 3.

Utilizamos la técnica de *performance profile* propuesta por Dolan y Moré [7] para mostrar el comportamiento de cada algoritmo en el conjunto de problemas test considerados.

En la figura 1 se muestra el desempeño observado basado en el número de evaluaciones de funciones en las iteraciones externas de los Algoritmos 1 y 2. El Algoritmo 1 en el 77 % de los casos resulta más eficiente que el Algoritmo 2 y la superioridad del primero es clara.

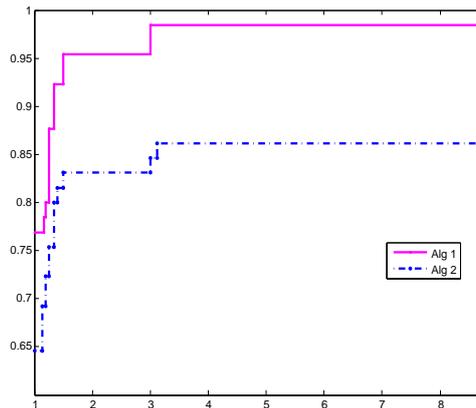


Figura 1: Número de evaluaciones de funciones en las iteraciones externas

En las figuras 2 (a) y (b) se muestran los resultados obtenidos respecto del número de evaluaciones de funciones en las iteraciones internas de los Algoritmos 1 y 2 para diferentes rangos. En el caso (a) se muestra el desempeño en el intervalo $[0, 50]$ y en el (b), $[0, 25]$. Aquí el Algoritmo 2 se impone sobre el 1 en aquellos problemas que requieren menor tiempo de ejecución y ocurre lo contrario con los que requieren mayor tiempo.

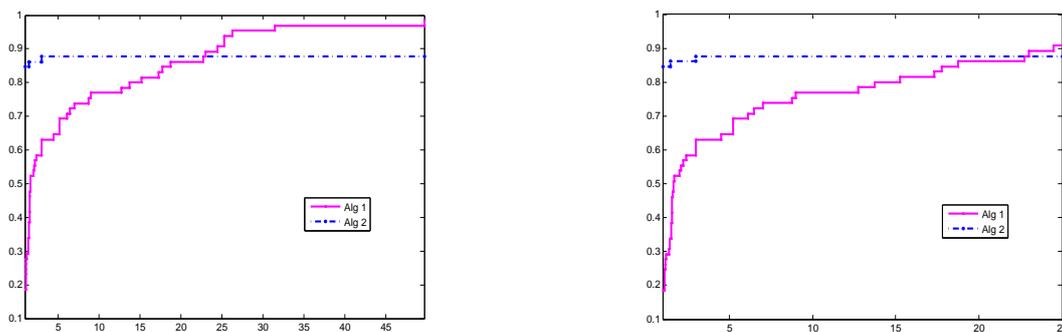


Figura 2: Número de evaluaciones de funciones en las iteraciones internas (a) $[0, 50]$, (b) $[0, 25]$.

En la figura (3) se muestran los resultados relacionados con el número de evaluaciones externas de los Algoritmos 1 y 3. Se ve que el Algoritmo 1 resulta ganador en el 98 % de los casos estudiados y tienen comportamientos similares en $[32, 45]$.

En la figura 4 (a) y (b) se observa la performance para los Algoritmos 1 y 3 respecto del número de iteraciones externas de funciones para distintos rangos, $[0, 100]$ y $[0, 25]$ respectivamente. Es claro que el comportamiento del Algoritmo 2 es, en general, superior al 3.

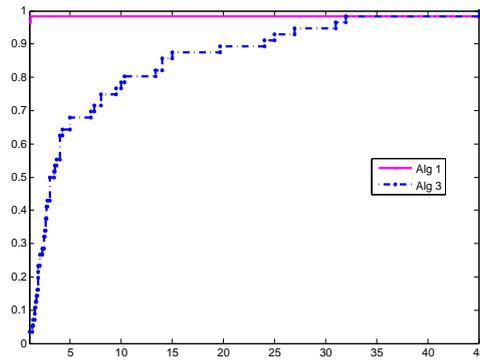


Figura 3: Número de evaluaciones de funciones en las iteraciones externas

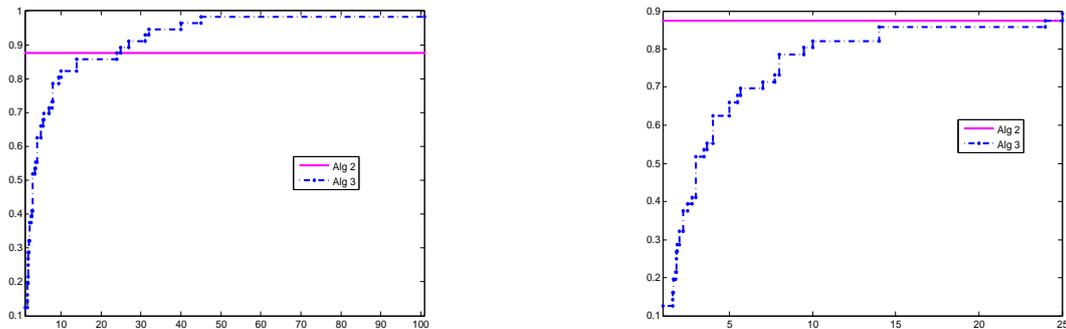


Figura 4: Número de evaluaciones de funciones en las iteraciones externas (a) $[0, 100]$, (b) $[0, 25]$.

4. CONCLUSIONES

Hemos propuesto dos algoritmos de Lagrangiano Aumentado, para resolver el problema general de optimización con restricciones de igualdad y cotas sobre las variables, que combinan diferentes técnicas y estrategias globalizadoras.

Los algoritmos fueron comparados entre sí y con un tercer algoritmo de región de confianza.

Las pruebas mostraron que entre los dos algoritmos de Lagrangiano Aumentado, el que usa una estrategia de región de confianza resultó más eficiente para problemas que requieren más tiempo de ejecución, y en todos los casos, el número de evaluaciones de funciones en las iteraciones externas fue inferior.

Cuando los dos algoritmos de Lagrangiano Aumentado se compararon con el tercer algoritmo de región de confianza, los test mostraron que, en general, resultaron más eficientes para problemas que requieren menos tiempo de ejecución, y frente al algoritmo propuesto por Niu y Yuan [12], este último requiere mayor cantidad de evaluaciones de funciones en las iteraciones externas.

REFERENCIAS

- [1] D.P. BERTSEKAS, *On the Goldstein-Levitin-Polyak Gradient Projection Method*, IEEE Transactions on Automatic Control, 21 (1976), pp.174-184.
- [2] D.P. BERTSEKAS, *Projected Newton methods for optimization problems with simple constraints*, SIAM J. Control Optim, 20 (1982), pp.221-246.

- [3] D.P. BERTSEKAS, *Constrained Optimization and Lagrange Multiplier Methods*, Athenas Scientific, 1996.
- [4] J.E. DENNIS Y R.B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, 1996.
- [5] A.R. CONN, N.I.M. GOULD Y PH.L. TOINT, *Trust-Region Methods*, SIAM, 2000.
- [6] G.M. CROCERI, M.C. MACIEL Y G.N. SOTTOSANTO, *Augmented penalty algorithms based on BFGS secant approximations and trust regions*, *Applied Numerical Mathematics*, 57 (2007), pp.320-334.
- [7] E. DOLAN Y J. MORÉ, *Bechmarking optimization software with performance profiles*, *Mathematical Programming*, 91(2002), pp.201-213.
- [8] R. FLETCHER, *Practical Methods of Optimization*, John Wiley, 1987.
- [9] N.I.M. GOULD, D. ORBAN Y PH.L. TOINT, *CUTEr (and SifDec), a constrained and unconstrained testing environment, revisited*, Technical Report TR/PA/01/04 (2001), CERFACS.
- [10] C.T. KELLEY, *Iterative Methods for Optimization*, SIAM, 1999.
- [11] M.C. MACIEL Y G.N. SOTTOSANTO, *An augmented penalization algorithm for the equality constrained minimization problem*, *Seleta do XXIV CNMAC*, 3(2002), pp.171-180.
- [12] L. NIU Y Y. YUAN, *A new trust-region algorithm for nonlinear constrained optimization*, *Journal of Computational Mathematics*, Vol.28, No.1 (2010), pp. 72-86.
- [13] C.J. NOCEDAL Y S. WRIGHT, *Numerical Optimization*, Springer, 2006.
- [14] T. STEihaug, *The conjugate gradient method and trust regions in large scale optimization*, *SIAM Journal on Numerical Analysis*, 20(1983), pp.626-637.