Análisis de convergencia global de un método de restauración inexacta sin el uso de derivadas que utiliza la técnica de filtro inclinado

María Mercedes Olea[1 2], Raúl P. Vignau[1], and María Laura Schuverdt[1 2]

[1] Departamento de Matemática, Facultad de Ciencias Exactas,
Universidad Nacional de La Plata, Argentina,
[2] CONICET, Argentina,
molea@mate.unlp.edu.ar

Resumen En este trabajo se presenta el estudio de convergencia de un método de restauración inexacta sin derivadas para resolver problemas de optimización no lineal con restricciones de igualdad que utiliza la técnica de filtro inclinado. Este método trata a la función objetivo y a la restricción como dos objetivos independientes. Cada iteración del algoritmo está compuesta de dos fases: la de restauración, en la cual se reduce la infactibilidad de las restricciones, y una fase de minimización, en la cual se reduce la función objetivo. En la fase de restauración se emplea un algoritmo Quasi-Newton que utiliza una búsqueda lineal no monótona sin derivadas y en la de minimización se emplea un algoritmo de región de confianza sin derivadas. Los algoritmos de filtros definen una región prohibida memorizando pares obtenidos por iteraciones previas y luego evitan pares que estén dominados por los pares memorizados.

Keywords: Método de filtros, Método de Restauración Inexacta, Filtro inclinado, Optimización sin derivadas

1. Introducción

En este trabajo consideraremos el problema de programación no lineal

$$\begin{cases} min & f(x) \\ \text{s.a} & c(x) = 0 \end{cases}$$

donde las funciones $f : \mathbb{R}^n \to \mathbb{R}, c : \mathbb{R}^n \to \mathbb{R}^m$ son continuamente diferenciables pero sus derivadas no se encuentran disponibles. Denotaremos por $J_c(.)$ a la matriz jacobiana de c y consideraremos la función h que mide la infactibilidad de las restricciones en cada punto $x \in \mathbb{R}^n$ siendo $h(x) = \|c(x)\|$, donde $\|.\|$ denota una norma arbitraria.

En [5] los autores definen un método de filtros globalmente convergente para problemas de programación no lineal considerando que las derivadas de la función objetivo y de las restricciones están disponibles. Tal algoritmo de filtros pertenece a la clase de métodos que tratan a f y h como dos objetivos independientes.

Cada iteración del método que presentamos consta de dos fases: la fase de restauración o de factibilidad en la cual debe reducirse la infactibilidad sin hacer uso de la función objetivo, y la fase de optimización o minimización en la cual se mejoran los valores de la función objetivo sobre una aproximación tangente de las restricciones. Como es conocido, los métodos de filtro definen una región prohibida memorizando pares $(f(x^k), h(x^k))$ de las iteraciones anteriores. En [5] para definir esa región prohibida se utiliza la siguiente regla de dominación:

$$x \text{ es dominado por } y \text{ si y sólo si } \quad f(x) \geq f(y) - \alpha h(y) \quad \text{y} \quad h(x) \geq (1-\alpha)h(y)$$

donde $\alpha \in (0,1)$ es una constante fija. Evitando los puntos dominados por la regla anterior se generan regiones prohibidas que llamaremos filtro recto.

En este trabajo, usamos otra regla de dominación, que inicialmente fue propuesta por Chin y Fletcher en [1]:

$$x \text{ es dominado por } y \text{ si y sólo si } \quad f(x) + \alpha h(x) \geq f(y) \quad \text{y} \quad h(x) \geq (1-\alpha)h(y)$$

donde $\alpha \in (0,1)$ es una constante fija. Evitando los puntos dominados por esta última regla se generan regiones prohibidas que llamaremos filtro inclinado.

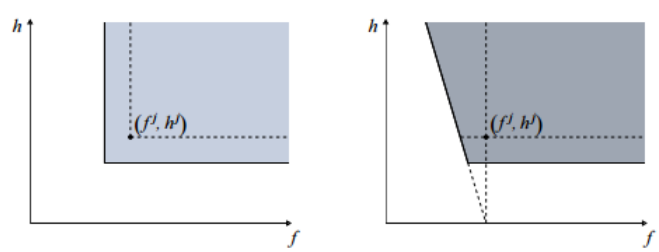

Figura 1. Regiones prohibidas considerando filtro recto y filtro inclinado

En [4] se prueba que el algoritmo de filtros sin derivadas, que utiliza la regla de dominación del filtro recto, genera una sucesión $\{x^k\}_{k \in \mathbb{R}^n}$ que posee un punto límite factible $\bar{x} \in \mathbb{R}^n$, $\quad \bar{x} = \lim\limits_{x \in K} x^k$ para algún conjunto infinito K de $\mathbb{N}$, que satisface:

$$\lim_{x \in K} \|d_c(x^k)\| = 0$$

donde $d_c(z) = P_{L(z)}(z - \nabla_s f(z)) - z$, $L(z) = \{x \in \mathbb{R}^n : A(z)(x - z) = 0\}$, $A(z)$ es una aproximación de la matriz Jacobiana $J_c(z)$ y $\nabla_s f$ indica el gradiente simplex de f ([2, Ch. 2]). Esos puntos factibles $\bar{x}$ fueron llamados puntos *cuasi-estacionarios*.

En este trabajo se analiza la convergencia global del algoritmo de restauración inexacta sin el uso de derivadas definido en [4], cuando se utiliza la técnica del filtro inclinado en lugar de la estrategia del filtro recto.

2. Hipótesis para la convergencia global de los algoritmos de filtros sin derivadas

Se presentará un algoritmo que genera sucesiones $\{x^k\}$, $\{z^k\} \subset \mathbb{R}^n$ y para obtener la convergencia global del mismo, se asumen las siguientes hipótesis:

(H1) Los iterados x^k y z^k se mantienen dentro de un conjunto compacto convexo $X \subset \mathbb{R}^n$.

(H2) Las funciones f, c_i para $i = 1, ..., m$ son continuamente diferenciables en un conjunto abierto que contiene a X.

(H3) Las funciones $\nabla f, \nabla c_i$ para $i = 1, ..., m$ son Lipschitz continuas en un conjunto abierto que contiene a X con constantes $L_1, L_2 > 0$ respectivamente, es decir:

$$\|\nabla f(x) - \nabla f(y)\| \leq L_1 \|x - y\|$$

$$\|\nabla c_i(x) - \nabla c_i(y)\| \leq L_2 \|x - y\|$$

para $i = 1, ..., m$; para todo x, y en un conjunto abierto que contiene a X.

Introducimos algunos conceptos y resultados de modelos de interpolación polinomial multivariada de la función objetivo y de las restricciones que se utilizarán durante este trabajo.

Cada conjunto de interpolación $Y = \{y^0, y^1, ..., y^n\} \subset \mathbb{R}^n$ que está contenido en la bola $B(y^0, \Delta(Y))$ centrada en y^0 y con radio $\Delta(Y) = \max\limits_{1 \leq i \leq n} \left\| y^i - y^0 \right\|$ se dice *equilibrado* para interpolación lineal si la matriz de direcciones $S = [y^1 - y^0 \;\; y^2 - y^0 \; ... \; y^n - y^0]^T$ es no singular.

El *gradiente simplex* de f en y^0 está definido por $\nabla_s f(y^0) = S^{-1} \delta f(Y)$ donde $\delta f(Y) = (f(y^1) - f(y^0), f(y^2) - f(y^0), ..., f(y^n) - f(y^0))^T$.

Si consideramos $m_f(x) = f(y^0) + g_f^T(x - y^0)$ el modelo de interpolación lineal de $f(x)$ sobre el conjunto Y, entonces se tiene que $g_f = \nabla_s f(y^0)$. Por lo tanto el gradiente simplex está cercanamente relacionado con la interpolación lineal multivariada.

Las propiedades geométricas de Y determinan la calidad del correspondiente gradiente simplex g_f como una aproximación al gradiente verdadero de la función objetivo f. Nos interesa la calidad de $m_f(x)$ y g_f en la bola $B(y^0, \Delta(Y))$.

Para todo $x \in B(y^0, \Delta(Y))$, considerando la matriz escalada $\bar{S} = \frac{S}{\Delta(Y)}$, se obtiene que

$$|f(x) - m_f(x)| \leq k_{ef}\Delta^2(Y),$$
$$\|\nabla f(x) - \nabla m_f(x)\| \leq k_{eg}\Delta(Y),$$

donde $k_{eg} = L_1(1 + \frac{\sqrt{n}}{2}\|\bar{S}^{-1}\|)$ y $k_{ef} = k_{eg} + \frac{L_1}{2}$ están dados en los Teoremas 2.11 y 2.12 de [2].

Análogamente, bajo las mismas hipótesis, si consideramos para todo $j = 1, ..., m$, $m_{c_j}(x) = c_j(y^0) + g_{c_j}^T(x - y^0)$ como el modelo lineal de interpolación de $c_j(x)$ sobre Y, se tiene que $g_{c_j} = \nabla_s c_j(y^0)$ así como también las siguientes cotas de error:

$$|c_j(x) - m_{c_j}(x)| \leq k_{ec}\Delta^2(Y),$$
$$\|\nabla c_j(x) - \nabla m_{c_j}(x)\| \leq k_{eg_c}\Delta(Y),$$

donde $k_{eg_c} = L_2(1 + \frac{\sqrt{n}}{2}\|\bar{S}^{-1}\|)$ y $k_{ec} = k_{eg_c} + \frac{L_2}{2}$.

Si consideramos a $A(y)$ como la aproximación de la matriz $J_c(y)$, donde la fila j-ésima es la traspuesta de ∇m_{c_j} entonces tenemos que

$$\|J_c(y) - A(y)\| \leq k_{eJ_c}\Delta(Y),$$

donde $k_{eJ_c} = \sqrt{m}k_{eg_c}$.

Suponemos que es posible mantener las constantes k_{ef}, k_{eg} y k_{eJ_c} uniformemente acotadas a lo largo del proceso iterativo de nuestro algoritmo.

Dado un iterado z^k consideramos la siguiente hipótesis:

(H4) El gradiente simplex usado como aproximación del gradiente de la función objetivo satisface la siguiente cota del error:

$$\|\nabla f(z^k) - \nabla_s f(z^k)\| \leq k_{eg}\Delta_f^k$$

donde Δ_f^k es el radio de la bola que contiene los puntos de interpolación.

Las derivadas simplex usadas para aproximar a la verdadera matriz Jacobiana satisfacen la siguiente cota del error:

$$\|J_c(z^k) - A(z^k)\| \leq k_{eJ_c} \Delta_c^k$$

donde Δ_c^k es el radio de la bola que contiene los puntos de interpolación.

Se tiene el siguiente resultado:

Lemma 1. *(Lemma 1 de [4]) Dado $\epsilon > 0, z^k \in \mathbb{R}^n$, si $\|d_c(z^k)\| > \epsilon$ y $\|\nabla f(z^k) - \nabla_s f(z^k)\| < \frac{\epsilon}{4}$ entonces*

$$\|z^k - P_{L(z^k)}(z^k - \nabla f(z^k))\| > \frac{3}{4}\epsilon,$$

$$\nabla^T f(z^k) d_c(z^k) < -\frac{1}{4}\|d_c(z^k)\|^2.$$

Observación: En la demostración se emplea la siguiente desigualdad:

$$\nabla_s^T f(z^k) d_c(z^k) \leq -\frac{\|d_c(z^k)\|^2}{2} \tag{1}$$

que será utilizada en los resultados de convergencia global.

3. Algoritmo de restauración inexacta sin el uso de derivadas utilizando Filtro Inclinado (DFF-I)

Dados $x^0 \in \mathbb{R}^n, F_0 = \emptyset, \mathcal{F}_0 = \emptyset, \alpha \in (0,1), \beta > 0, \epsilon_f > 0, \epsilon_I > 0, \{\delta_k\}_{k \in \mathbb{N}}, \delta_k > 0, \delta_k \to 0$.
Considerar $k \leftarrow 0$.

Paso 1: Definir $(\tilde{f}, \tilde{h}) = (f(x^k) - \alpha h(x^k), (1-\alpha) h(x^k))$.
Construir el conjunto $\bar{F}_k = F_k \cup \{(\tilde{f}, \tilde{h})\}$.
Definir el conjunto $\bar{\mathcal{F}}_k = \mathcal{F}_k \cup \{x \in \mathbb{R} : f(x) + \alpha h(x) \geq f(x^k), h(x) \geq \tilde{h}\}$.

Paso 2: Fase de restauración
Si $h(x^k) = 0$ definir $z^k = x^k$.
Si no, calcular $z^k \notin \bar{\mathcal{F}}_k$ tal que $h(z^k) < (1-\alpha) h(x^k)$ y $\|z^k - x^k\| \leq \beta h(x^k)$.
Si esto no es posible, detener el proceso sin éxito. FIN.

Paso 3: Fase de optimalidad
3.1 Construir o actualizar $Y_c^k = \{z^k, y_c^1, ..., y_c^n\}$, un conjunto de puntos de interpolación centrado en z^k tal que el radio $\Delta_c^k = \max\limits_{i=1,\ldots,n} \{\|y_c^i - z^k\|\}$ verifique $\Delta_c^k \leq \beta \min\{\max\{h(x^k), H_k\}, \delta_k\}$, donde H_k será definida en (2).
Calcular $A_k = A(z^k)$ usando derivadas simplex interpolando sobre Y_c^k.
Definir $L(z^k) = \{x \in \mathbb{R}^n : A_k(x - z^k) = 0\}$.
Construir o actualizar $Y_f^k = \{z^k, y_f^1, ..., y_f^n\}$, un conjunto de puntos de interpolación centrado en z^k tal que el radio $\Delta_f^k = \max\limits_{i=1,\ldots,n} \{\|y_f^i - z^k\|\}$ verifique $\Delta_f^k \leq \delta_k$.

Calcular $\nabla_s f(z^k)$ interpolando sobre Y_f^k y $d_c(z^k) = P_{L(z^k)}(z^k - \nabla_s f(z^k)) - z^k$.
3.2 Si $h(x^k) = 0, \text{máx}\{\Delta_f^k, \Delta_c^k\} < \epsilon_I$ y $\|d_c(z^k)\| < \epsilon_f$ finalizar el proceso con convergencia finita. FIN.
3.3 Calcular, con un algoritmo sin derivadas, $x_T \notin \bar{\mathcal{F}}_k$ tal que $x_T \in L(z^k)$ y $f(x_T) \leq f(z^k)$.
Si $z^k = x^k$ y no existe x_T tal que $f(x_T) < f(z^k)$ considerar $\Delta_f^k = \alpha\Delta_f^k$, $\Delta_c^k = \alpha\Delta_c^k$ e ir al Paso 3.1.
Si no, definir $x^{k+1} = x_T$.

Paso 4: Actualización del filtro
Si $f(x^{k+1}) < f(x^k)$ considerar $F_{k+1} = F_k, \mathcal{F}_{k+1} = \mathcal{F}_k$ (f-iteración).
Si no, considerar $F_{k+1} = \bar{F}_k, \mathcal{F}_{k+1} = \bar{\mathcal{F}}_k$ (h-iteración).
$k \leftarrow k+1$ e ir al Paso 1.

Los siguientes resultados generales se desprenden directamente de la construcción del Algoritmo:

(R1) Dado $k \in \mathbb{N}, x^{k+p} \notin \mathcal{F}_{k+1}$ para todo $p \geq 1$.

(R2) Dado $k \in \mathbb{N}$, al menos una de las siguientes situaciones debe ocurrir:

1. $f(x^{k+1}) + \alpha h(x^{k+1}) < f(x^k)$
2. $h(x^{k+1}) < (1-\alpha)h(x^k)$

(R3) Dado $k \in \mathbb{N}, h_j > 0$ para todo $j \in \mathbb{N}$ tal que $(f_j, h_j) \in F_k$. Consecuentemente $H_k > 0$ para todo $k \in \mathbb{N}$, siendo H_k la *holgura del filtro*, que se define como:

$$H_k = \text{mín}\{1, \text{mín}\{h_j : (f_j, h_j) \in F_k, f_j \leq f(x^k)\}\} \tag{2}$$

El siguiente teorema garantiza que cualquier punto de acumulación de la sucesión generada por el Algoritmo DFF-I es factible. Este resultado es demostrado por Chin y Fletcher en [1] para un algoritmo general que utiliza la técnica de filtro inclinado.

Theorem 1. *Consideremos la sucesión $\{x^k\}$ generada por el Algoritmo DFF-I. Entonces $h(x^k) \to 0$ y consecuentemente, cualquier punto límite de la sucesión es factible.*

4. Algoritmos Internos

Los métodos de restauración inexacta ofrecen la posibilidad de utilizar diferentes métodos para resolver cada fase. Los algoritmos que se utilizan en cada fase, verifican las condiciones requeridas para obtener convergencia global del algoritmo DFF-I que enunciaremos en la próxima sección.

En la fase de restauración se emplea el algoritmo BCDF-QNB [3]. Dicho método encuentra una solución aproximada al problema $c(x) = 0$ con $x \in \Gamma$ siendo $\Gamma = \{x \in \mathbb{R}^n : l \leq x \leq u\}$. BCDF-QNB es un método del tipo Quasi-Newton para resolver sistemas indeterminados de ecuaciones no lineales con restricciones de cotas en las variables que utiliza una búsqueda lineal no monótona sin el uso de derivadas.

En la fase de minimización se debe encontrar $x^{k+1} \notin \overline{\mathcal{F}}_k$ tal que $x^{k+1} \in L(z^k)$ y $f(x^{k+1}) \leq f(z^k)$ utilizando algún método que no utilice las derivadas. Empleamos el método de región de confianza descripto en [4].

Dado z^k generado por la fase de factibilidad, el algoritmo de región de confianza utiliza el modelo lineal $m_k(x) = f(z^k) + \nabla_s^t f(z^k)(x - z^k)$. Consideramos un radio $\Delta > 0$ y resolvemos el problema

$$\begin{cases} min \quad m_k(x) \\ \text{s.a} \quad x \in L(z^k), \\ \qquad \|x - z^k\| \leq \Delta. \end{cases}$$

Como el modelo es lineal se sabe que la solución del problema es un punto $z^k + d(z^k, \Delta)$ tal que

$$d(z^k, \Delta) = \Delta \frac{d_c(z^k)}{\|d_c(z^k)\|}$$

si $d_c(z^k) \neq 0$, donde $d_c(z^k)$ es la dirección del gradiente proyectado definido por $P_{L(z^k)}(z^k - \nabla_s f(z^k)) - z^k$.

Definimos la *reducción estimada* por el modelo para el paso $d(z^k, \Delta)$ como

$$pred(z^k, \Delta) = m_k(z^k) - m_k(z^k + d(z^k, \Delta))$$

y la *reducción actual* como

$$ared(z^k, \Delta) = f(z^k) - f(z^k + d(z^k, \Delta)).$$

El paso $d(z^k, \Delta)$ es aceptado sólo si se satisface la siguiente condición de suficiente decrecimiento:

$$ared(z^k, \Delta) > \eta \; pred(z^k, \Delta)$$

siendo $\eta \in (0, 1)$.

Como $pred(z^k, \Delta) = -\nabla_s^T f(z^k) d(z^k, \Delta) = -\nabla_s^T f(z^k) \frac{d_c(z^k)}{\|d_c(z^k)\|} \Delta$, considerando (1), tenemos que

$$pred(z^k, \Delta) \geq \frac{\Delta}{2} \|d_c(z^k)\|. \tag{3}$$

5. Convergencia Global del Algoritmo DFF-I

Es conocido en la literatura, ver por ejemplo [7], que los métodos de filtros deben cumplir, además de las hipótesis usuales, una condición específica para que resulten globalmente convergentes. Una condición análoga para el contexto de optimización sin derivadas es la siguiente: dado un punto factible no cuasi-estacionario $\bar{x} \in X$, existe un entorno V de $\bar{x}$ tal que para cualquier iterado $x^k \in V$ se verifica que

$$f(x^k) - f(x^{k+1}) = \Omega(\sqrt{H_k}),$$

es decir, existe una constante positiva M tal que $f(x^k) - f(x^{k+1}) \geq M\ \sqrt{H_k}$.

Para garantizar el cumplimiento de esta condición alcanza con que los puntos obtenidos en la fase de restauración satisfagan:

(C1) Condición del paso de restauración: En todas las iteraciones $k \in \mathbb{N}$, el paso de restauración satisface:

$$\|z^k - x^k\| = O(h(x^k));$$

y que los puntos obtenidos en la fase de minimización satisfagan:

(C2) Condición del paso de minimización: Dado un punto factible no cuasi-estacionario $\bar{x} \in X$, existe un entorno V de $\bar{x}$ tal que para cualquier iterado $x^k \in V$ se verifica que

$$f(z^k) - f(x^{k+1}) = \Omega(\sqrt{H_k}).$$

La condición (C1) se encuentra garantizada por el Algoritmo BC-DFQNB utilizado en la fase de restauración.

Para lograr el cumplimiento de la condición (C2) necesitamos los siguientes resultados:

Lemma 2. *(Lemma 3 de [4]) Sea $\bar{x} \in X$ un punto de acumulación factible no cuasi-estacionario de una sucesión $\{x^k\}_{k\in\mathbb{N}}$ generada por el Algoritmo DFF-I. Entonces existen un entorno $\tilde{V}$ de $\bar{x} \in X, \tilde{\Delta} > 0$ y una constante $\tilde{c} > 0$ tales que para cualquier $z^k \in \tilde{V}$ y cualquier $\Delta \in (0, \tilde{\Delta})$*

$$ared(z^k, \Delta) > \eta pred(z^k, \Delta) \geq \eta \tilde{c} \Delta.$$

Lemma 3. *(Lemma 4 de [4]) Suponiendo que la matriz A_k, aproximación de $J_c(z^k)$, es calculada por derivadas simplex utilizando un radio Δ_c^k, entonces si $z^k + d \in L(z^k)$ se tiene que*

$$\left|h(z^k + d) - h(z^k)\right| \leq \sqrt{M} L_2 \left\|d\right\|^2 + \kappa_{eJ_c} \Delta_c^k \left\|d\right\|.$$

Demostraremos a continuación un lema que nos garantiza el cumplimiento de la condición (C2) cuando utilizamos el filtro inclinado:

Lemma 4. *Sea $\bar{x} \in X$ un punto de acumulación factible no cuasi-estacionario de una sucesión $\{x^k\}_{k\in\mathbb{N}}$ generada por el Algoritmo DFF-I. Asumiendo que se verifican la condición (C1) y las hipótesis de los lemas anteriores, entonces existe un entorno V de $\bar{x}$ tal que si $x^k \in V$ se tiene que*

$$f(z^k) - f(x^{k+1}) = \Omega(\sqrt{H_k}),$$
$$f(z^k) - f(x^{k+1}) = \Omega\left(\left\|x^{k+1} - z^k\right\|\right)$$

donde z^k es el punto restaurado y x^{k+1} es el iterado obtenido en la fase de minimización.

Demostración. Sea $\{x^k\}_{k\in K}$ una subsucesión tal que $\lim\limits_{k\in K} x^k = \bar{x}$.

Por (C1) y como $h(x^k)$ tiende a cero, se sigue que $\lim\limits_{k\in K} z^k = \bar{x}$.

Sea $\tilde{V} \subset X$ y $\tilde{\Delta} > 0$ el entorno de $\bar{x}$ y el radio dado por el Lema 2 tal que para cualquier $z^k \in \tilde{V}, k \in K$ y para cualquier $\Delta \in (0, \tilde{\Delta})$,

$$ared(z^k, \Delta) > \eta.pred(z^k, \Delta) \geq \eta\tilde{c}\Delta.$$

El algoritmo comienza con un radio $\Delta \geq \Delta_{\text{mín}}$ y calcula $d(z^k, \Delta_j)$, $\Delta_j = 2^{-j}\Delta$ para $j = 0, 1, ...$ hasta que $z^k + d(z^k, \Delta_j) \notin \bar{\mathcal{F}}_k$ y $ared(z^k, \Delta_j) > \eta.pred(z^k, \Delta_j)$. Entonces define $\Delta_k = \Delta_j$.

Definimos $\hat{\Delta}$ como el primer Δ_j tal que

$$ared(z^k, \Delta_j) > \eta.pred(z^k, \Delta_j) \text{ y} \tag{4}$$

$$z^k + d(z^k, \Delta_j) \notin \bar{\mathcal{F}}_k \text{ o } f(z^k + d(z^k, \Delta_j)) + \alpha h(z^k + d(z^k, \Delta_j)) \geq f(x^k) \tag{5}$$

donde $(\tilde{f}, \tilde{h}) = (f(x^k) - \alpha h(x^k), (1-\alpha)h(x^k))$ es la entrada temporaria del filtro.

Vamos a denotar $\hat{d} = d(z^k, \hat{\Delta})$ y $\hat{x} = z^k + \hat{d}$. Notemos que $\hat{\Delta} \geq \Delta_k$ y que $\hat{\Delta} > \Delta_k$ ocurre solo cuando $f(\hat{x}) + \alpha h(\hat{x}) \geq f(x^k)$.

Observar que por el Lema 3, para un Δ fijo tenemos que existe una constante $k_{eJ_c}\Delta_c^k > 0$ tal que:

$$\left|h(z^k + d(z^k, \Delta)) - h(z^k)\right| \leq k_{eJ_c}\Delta_c^k\|d(z^k, \Delta)\| + \sqrt{m}L_2\|d(z^k, \Delta)\|^2.$$

Por definición de holgura, se sabe que si $H_k \leq 1$ y x^k está en un entorno de un punto factible se verifica que $h(x^k) \leq H_k$.

Por lo tanto, considerando que $\|d(z^k, \Delta)\| = \Delta$ y que el radio Δ_c^k cumple que $\Delta_c^k \leq \beta \,\text{mín}\{\text{máx}\{h(x^k), H_k\}, \delta_k\}$, tenemos que:

$$\left|h(z^k + d) - h(z^k)\right| \leq k_{eJ_c} \beta H_k \Delta + \sqrt{m} L_2 \Delta^2. \tag{6}$$

Consideremos $\bar{\Delta}$ tal que $\bar{\Delta} \leq \frac{\alpha}{4\beta k_{eJ_c}}$ y $\bar{\Delta} < \frac{\tilde{\Delta}}{2} < \tilde{\Delta}$.

1. Supongamos que $\hat{\Delta} \geq \bar{\Delta}$. Por (3) tenemos que

$$pred(z^k, \hat{\Delta}) \geq \frac{\hat{\Delta}}{2} \|d_c(z^k)\| \geq \frac{\tilde{\epsilon}}{2} \hat{\Delta}.$$

Considerando $\tilde{c} = \frac{\tilde{\epsilon}}{2}$ tenemos que

$$pred(z^k, \hat{\Delta}) \geq \tilde{c}\hat{\Delta} \geq \tilde{c}\bar{\Delta}.$$

Por definición de $\hat{\Delta}$ se verifica (4) entonces:

$$f(z^k) - f(\hat{x}) > \eta \, pred(z^k, \hat{\Delta}) > \eta \tilde{c} \bar{\Delta} = \Omega(1).$$

Como $H_k \leq 1$ se deduce

$$f(z^k) - f(\hat{x}) = \Omega(\sqrt{H_k}).$$

2. Supongamos que $\hat{\Delta} < \bar{\Delta}$ entonces $2\hat{\Delta} < 2\bar{\Delta} < \tilde{\Delta}$ y $2\hat{\Delta}$ no verifica (5). Por el Lema 2,

$$ared(z^k, d(z^k, 2\hat{\Delta})) > \eta \, pred(z^k, d(z^k, 2\hat{\Delta}))$$

y por (5) se tiene que $z^k + d(z^k, 2\hat{\Delta}) \in \bar{\mathcal{F}}_k$ y $f(z^k + d(z^k, 2\hat{\Delta})) + \alpha h(z^k + d(z^k, 2\hat{\Delta})) < f(x^k)$.

Consecuentemente, por definición de H_k, se tiene que $h(z^k + d(z^k, 2\hat{\Delta})) \geq H_k$.

Por construcción $h(z^k) < (1 - \alpha) h(x^k) \leq (1 - \alpha) H_k$.

Asi, $h(z^k + d(z^k, 2\hat{\Delta})) - h(z^k) \geq \alpha H_k$.
Entonces, usando (6)

$$\alpha H_k \leq h(z^k + d(z^k, 2\hat{\Delta})) - h(z^k) \leq k_{eJ_c} \beta H_k 2\hat{\Delta} + 4\sqrt{m} L_2 \hat{\Delta}^2$$

obtenemos que

$$H_k \leq \frac{2\beta}{\alpha} k_{eJ_c} H_k \hat{\Delta} + \underbrace{\frac{4}{\alpha} \sqrt{m} L_2 \hat{\Delta}^2}_{O(\hat{\Delta}^2)}$$

pues $\hat{\Delta} < \bar{\Delta}$ y $\bar{\Delta} \leq \frac{\alpha}{4\beta k_{eJ_c}}$.

Por lo tanto,

$$\frac{1}{2}H_k = O(\hat{\Delta}^2) \text{ o } \hat{\Delta} = \Omega(\sqrt{H_k}).$$

Usando el Lema 2 con $\hat{\Delta} < \bar{\Delta} < \tilde{\Delta}$,

$$f(z^k) - f(\hat{x}) = ared(z^k, \hat{\Delta}) \geq \eta\tilde{c}\hat{\Delta} = \eta\tilde{c}\Omega(\sqrt{H_k}).$$

Así, para los dos casos tenemos que: $f(z^k) - f(\hat{x}) = \Omega(\sqrt{H_k})$. Entonces el paso $\hat{d}$ satisface las condiciones del lema.

Además, como

$$ared(z^k, \hat{\Delta}) = f(z^k) - f(\hat{x}) \geq \eta\, pred(z^k, \hat{\Delta}) = \eta\tilde{c}\hat{\Delta} = \eta\tilde{c}\|\hat{x} - z^k\|$$

llamando $M = \eta\tilde{c} > 0$ obtenemos que $f(z^k) - f(\hat{x}) \geq M\|\hat{x} - z^k\|$,
por lo tanto

$$f(z^k) - f(\hat{x}) = \Omega(\|\hat{x} - z^k\|).$$

Para finalizar la demostración debemos probar que para $k \in K$ suficientemente grande, $f(\hat{x}) + \alpha h(\hat{x}) < f(x^k)$ lo que implica que $\hat{x} \notin \bar{\mathcal{F}}_k$ y así $\hat{x} = x^{k+1}$.

De $f(z^k) - f(x^k) = \Omega(\sqrt{H_k})$ sabemos que existe una constante positiva M tal que

$$f(\hat{x}) \geq f(z^k) - M\sqrt{H_k}. \tag{7}$$

De la condición (C1) tenemos que existe una constante positiva N tal que

$$f(z^k) \leq f(x^k) + Nh(x^k). \tag{8}$$

Entonces, combinando (7) y (8) tenemos que

$$f(\hat{x}) + \alpha h(\hat{x}) \leq f(z^k) - M\sqrt{H_k} + \alpha h(\hat{x})$$

$$\leq f(x^k) + Nh(x^k) - M\sqrt{H_k} + \alpha h(\hat{x}) < f(x^k) + NH_k - M\sqrt{H_k} + \alpha H_k =$$

$$f(x^k) + N(\sqrt{H_k})^2 - M\sqrt{H_k} + \alpha(\sqrt{H_k})^2 = f(x^k) + \sqrt{H_k}(\sqrt{H_k}(N + \alpha) - M).$$

Como H_k tiende a cero, para k suficientemente grande se verifica que $\sqrt{H_k} < \frac{M}{N+\alpha}$, por lo que $\sqrt{H_k}(N+\alpha) - M < 0$. Por lo tanto $f(\hat{x}) + \alpha h(\hat{x}) < f(x^k)$ completando la demostración.

Puede probarse (Lemma 6 de [4]) que las condiciones (C1) y (C2) implican la siguiente condición:

(C3) Dado un punto factible no cuasi-estacionario $\bar{x} \in X$, existe un entorno V de $\bar{x}$ tal que para cualquier iteración $x^k \in V$, se verifica que

$$f(x^k) - f(x^{k+1}) = \Omega(\sqrt{H_k}).$$

También puede probarse que:

Dado un punto factible no cuasi-estacionario $\bar{x} \in X$, existe un entorno V de $\bar{x}$ tal que para cualquier $x^k \in V$

$$f(x^k) - f(x^{k+1}) = \Omega(\left\|x^k - x^{k+1}\right\|).$$

El resultado anterior puede demostrarse de manera análoga a la demostración del Lema 3.2 de [6] cuando se utilizan las derivadas y la definición de puntos estacionarios.

Los siguientes lemas nos permitirán demostrar el resultado principal de convergencia global del Algoritmo DFF-I.

Lemma 5. *(Lemma 3.3 de [6]) Existe una constante $M > 0$ tal que para cualquier iterado k*

$$f(x^{k+1}) \leq f(x^k) + Mh(x^k).$$

Lemma 6. *(Lemma 3.4 de [6]) Consideremos una sucesión finita de iteraciones $I = \{\bar{k}, \bar{k}+1, ..., K\}$ tal que para $k \in I$, $f(x^k) \geq f(x^{\bar{k}})$ y sea $M > 0$ la constante dada por el lema anterior. Entonces*

$$f(x^K) \leq f(x^{\bar{k}}) + \frac{M}{\alpha} h(x^{\bar{k}}).$$

Theorem 2. *(Theorem 3.5 de [6]) La sucesión $\{f(x^k)\}$ converge.*

El lema presentado a continuación se utiliza para demostrar la convergencia global del Algoritmo DFF-I. La demostración es análoga a la del Lema 3.6 de [6] adaptada para el caso sin el uso de derivadas.

Lemma 7. *Sea $\bar{x} \in X$ un punto factible no cuasi-estacionario. Entonces existe un entorno V de $\bar{x}$ y $\delta > 0$ tal que para $x^k \in V$, existe $l_k \in \mathbb{N}$ tal que*

$$f(x^k) - f(x^{k+l_k}) \geq \delta.$$

El siguiente es el resultado principal del trabajo y afirma que cualquier punto de acumulación de la sucesión generada por el Algoritmo DFF-I, es cuasi-estacionario.

Theorem 3. *Cualquier punto de acumulación de la sucesión $\{x^k\}$ generada por el Algoritmo DFF-I, es cuasi-estacionario.*

Demostración. Supongamos que existen $\bar{x} \in X$ no cuasi-estacionario y un conjunto infinito $K \subset \mathbb{N}$ tal que $\lim_{k \in K} x^k = \bar{x}$. Por el Teorema 1, $\bar{x}$ es factible.

Por el Lema 7, existe $\delta > 0$ tal que para $k \in K$ grande, existe $l_k \in \mathbb{N}$ tal que

$$f(x^k) - f(x^{k+l_k}) \geq \delta.$$

Esto significa que la sucesión $\{f(x^k)\}$ no es de Cauchy, por lo que $\{f(x^k)\}$ no converge, contradiciendo el Teorema 2.

Por lo tanto, todo punto límite de la sucesión generada por el Algoritmo DFF-I es cuasi-estacionario como queríamos demostrar.

6. Conclusiones

En este trabajo analizamos la convergencia global de un algoritmo de restauración inexacta sin el uso de derivadas que utiliza la técnica de filtro inclinado para la aceptación de los iterados.Empleando esta técnica logramos demostrar que se cumplen las condiciones que permiten probar los teoremas de convergencia global.

Además se extendieron los resultados de convergencia demostrados en [4] ya que en el presente trabajo se prueba que todo punto límite es cuasi-estacionario, mientras que en [4] solo se garantiza la existencia de un punto límite cuasi-estacionario.

Agradecimientos: A nuestra querida profesora Nelly Echebest, sus enseñanzas y cariño nos acompañan siempre.

Referencias

1. Chin, C. M., Fletcher, R.: On the global convergence of an SLP-filter algorithm that takes EQP steps. Mathematical Programming, 96 (1), 161–177 (2003)
2. Conn, A., Scheinberg, K., Vicente, L. N.: Introduction to derivative-free optimization SIAM Book Series on Optimization, Philadelphia (2009)
3. Echebest, N., Schuverdt, M. L., Vignau, R. P.: A derivative-free method for solving box-constrained underdetermined nonlinear systems of equations. Mathematics and Computation, 219 (6), 3198–3208 (2012)
4. Echebest, N., Schuverdt, M. L., Vignau, R. P.: An inexact restoration derivative-free filter method for nonlinear programming. Computational and Applied Mathematics, 36 (1), 693–718 (2017)
5. Gonzaga, C. C., Karas, E. W., Vanti, M. : A globally convergent filter method for nonlinear programming. SIAM Journal on Optimization 14 (3), 646–669 (2004)
6. Karas, E. W., Oening, A. P., Ribeiro, A. A. : Global convergence of slanting filter methods for nonlinear programming. Applied Mathematics and Computation 200, 486–500 (2008)
7. Ribeiro, A. A., Karas, E. W.: Otimização Contínua: Aspectos teóricos e computacionais. Cengage Learning, São Paulo (2013)