Optimización

Trabajo de Iniciación a la Investigación

Análisis de convergencia de un método quasi-Newton libre de derivadas para resolver sistemas indeterminados de ecuaciones no lineales

María Mercedes Olea

Director: Raúl P. Vignau

Profesora a cargo: María Laura Schuverdt

Facultad de Ciencias Exactas, Universidad Nacional de La Plata

16 de diciembre de 2012

Iniciación

Nº

30

A Raúl S. Olea:
mi padre, mi ángel

Agradecimientos:

A mamá, por ser la mejor de las mamás, porque trabajaste toda tu vida para que yo pudiera tener la mejor educación, por ser mi gran compañera, porque desde hace unos años también sos mi papá y por ser la mujer mas fuerte (pero sensible a la vez) del mundo, que supo sobrellevar grandes dolores en su vida, pero siempre entera, para tratar de que yo no sufriera tanto.

A papá, mi angelito de la guarda, porque sos la persona que más quería que llegara este momento, creo que hasta mas que yo. Te fuiste demasiado pronto, pero sé que desde algún lugar me seguís aconsejando y ayudando cómo lo hiciste siempre. Te extraño demasiado y hay días que siento que no puedo vivir sin tu compañía física, pero esos mismos días son los que más siento tu presencia junto a mi. Éste trabajo está dedicado a vos, porque recién hoy puedo cumplirte uno de los tres deseos que pediste en tu último cumpleaños y porque daría todo lo que tengo para volver a verte un segundo mas.

A mi abuela Tita, mi otro angelito, porque hubiera querido tenerte presente en este momento. Porque me malcriaste por ser tu única nieta, tu Muñeca (como me llamabas), pero también me criaste con todo el amor. Por ser tu vida y tus ojos.

A mis tíos, primos y primos segundos. Por preocuparse por mi en todo momento, por estar a mi lado en momentos difíciles. Un recuerdo muy especial a mi tío Cesar, quien intentó cumplir el rol de padre cuando faltó el mío y lo cumplió muy bien. También te fuiste hace poco, pero nunca del todo. Otro recuerdo para mi prima María Paula, porque fuiste como una hija para mis papás. Porque ahora te están cuidando papá y la abuela.

A Raúl, por aceptar ser mi director mucho tiempo antes de que llegara este momento. Porque fue re lindo trabajar con vos, porque me haces reir mucho y sobre todo por tu excelente predisposición para mis consultas. Porque cada ida a tu oficina implicaba un gran adelanto en la elaboración de ésta tesis. Porque siento que al elegirte, fue la mejor decisión que pude haber tomado.

A mi jurado: Nelly, por ser mi directora en la beca que no pudo ser, por ayudar a Raúl en mi formación de tesista. Porque tenés un corazón muy grande y mucho para dar. Marité, por preocuparse tanto por mí en la carrera y en todo, porque no olvido tu abrazo aquel 6 de Noviembre de 2008. Victoria, porque aun recuerdo cuando papá me llevaba a tu oficina y yo te mostraba el boletín de la primaria. A las tres, por aceptar ser mi jurado. Les agradezco sobre todo el tiempo que le dedicaron a este trabajo.

A Laura, también por acompañar a Raúl en este proceso. Porque muchas veces diste tu oido para escucharme. Por aconsejarme sobre mi futuro. Por prepararme el plan de trabajo. Por compartir tantas cosas lindas.

A Edu, mi novio, porque con vos conocí el amor. Porque cada día que pasa más te quiero. Porque vivo pensando en vos y vos en mí. Porque quiero pasar el resto de mi vida a tu lado. Porque este último año que pasé al lado tuyo, fueron muchas las veces que sonreí... y hacía tiempo que no lo hacia.

A mis suegros Susana y Miguel. A mis cuñados Andrés y Lía. Por haberme integrado en su familia como una hija / hermana más. Porque son para mí un gran apoyo. Porque

estoy muy contenta de que también sean parte de mi familia. Por los lindos paseos compartidos.

A mis docentes del Departamento de Matemática: profesores, JTP y ayudantes. Porque sé que no me resultó fácil pero ustedes siempre trataron de que se me haga lo mas fácil posible. Por haberme aprobado cuando tenia que aprobar y desaprobado cuando lo merecía. Por la gran paciencia que tuvieron conmigo, por la buena predisposición que tuvieron siempre.

A los trabajadores no-docentes de la Facultad de Ciencias Exactas. Por ayudarme en todo lo que necesite, desde la búsqueda de un libro hasta completar papeles. Por ser tan buenos compañeros de trabajo de mi papá.

A los equipos docentes con quienes trabajo/trabajé en algún momento. Por priorizar mis momentos de estudio y por incentivarme a que rinda. Por cubrirme los días en los que no podía ir a trabajar por tener que rendir. Por preguntarme y preocuparse por mi situación academica.

A todos mis compañeros de carrera: Noe, Lupe, Adri, Iván, Hernán, Gi, Yami, Nahuel, Laura, Pía, Gabi C, Eli, Juliana, Eva, Lucila, Mariana J, Noemí L, Gastón, Leo, Joaquin... A Paulita y Eugenio, por ser tan lindos amigos. Los "chicus": Nadia, Pablo, Herni, Hernancito, Gabi R, Matias, Noe, Juli, Mariana F, Mimí, Nico, Manu, Leandro y Blas... Porque además de mis compañeros, son mis amigos. Por tantas tardes de estudio compartidas y tantas otras fuera de la facultad. A Ceci y Emma, por ser tan buenas compañeras de estudio y sobre todo tan buenas personas y a amigas, por hacer que las quiera tanto. A Eugenia, por pasarse tantas horas a mi lado explicandome todo, por tanto. A Dani, por tomarme las exposiciones previas a los finales, por traducirme el portugués, porque estoy felíz que seas parte del grupo de Optimización. Porque me atrasé en las cursadas, pero gracias a eso, tuve el honor de cursar y conocer a tantas y excelentes personas.

A mis amigos Romi y Juan. Porque son grandes amigos mios. Porque son capaces de sacarse lo que tienen puesto para darmelo a mí y así evitar que nada me falte. Porque me cuidan y me ayudan cada día, en todo.

A mis amigos Erica y Maxi Riddick. Por tantos años de amistad, una gran amistad. Tantas charlas, mates y momentos compartidos. Porque hace poco Eri y Jorge me hicieron la mas orgullosa de las tías. A Larita, porque la quiero desde antes de saber que estaba en la panza.

A mi amigo Walter, porque la amistad que comenzamos en mi primer año de carrera, la mantenemos hasta el día de hoy. Por los lindos momentos de estudio de las materias de primer año y por quererme tanto. A mis compañeros del Observatorio de mi primer año en la carrera.

A mis alumnos de Matemática C de Ingeniería, de Algebra I de cibex y del 4to año Economía de la escuela Inmaculada de Ensenada. A mis exalumnos de Matemática A y C de Ingeniería, Análisis I de cibex, Matemática 1 y 2 de Informática y de los ingresos a Ingeniería y a Informática. Porque todos me ayudan/ayudaron a formarme como docente. Por preocuparse por mis finales. Por desearme buenos deseos.

A Javierito, Marita y Ramiro Funes, la familia que me regaló mi abuela. Por estar con mamá y conmigo en momentos difíciles.

A mis compañeros del proyecto de extensión "Matemática en acción" y del voluntariado universitario "Conectar con la igualdad", por que cada actividad que iniciamos y desarrollamos fue siempre un gran desafío para todos y una alegría compartida.

A Luz Monteverde, mi profesora de Matemática de la secundaria. Por hacerme aprender a amar a la Matemática. Porque en parte estudié esta carrera gracias a ella. A mis compañeras del colegio Misericordia, porque nunca perdimos el contacto.

A Flora, mi madrina de bautismo porque pese a la distancia, siempre se preocupó por mí académicamente y emocionalmente. Por compartir lindos momentos con mis papás cuando yo era chica. Porque siempre espero tus llamadas telefónicas en los cumpleaños.

A mi madrina de confirmación, mi tía Ana, porque a pesar de que hay veces que pasa mucho tiempo sin vernos ni hablarnos, cuando lo hacemos, siempre tenés palabras lindas y de cariño hacia mi.

Estoy infinitamente agradecida a la vida por tener esta lista tan larga... Por tener mucha gente que quiero y que me quiere.

Índice

1. Introducción **4**

2. Generalidades **5**
2.1. El método de Newton 6
2.2. El método de Newton discreto 8
2.3. Métodos quasi-Newton 9

3. Búsqueda lineal no monótona libre de derivadas **11**
3.1. Algoritmo de búsqueda lineal 11
3.2. Análisis para el caso M=1 17

4. Método de quasi-Newton libre de derivadas usando la fórmula de actualización de Broyden (DF-QNB) **18**
4.1. Algoritmo DF-QNB 19
4.2. Análisis de convergencia global 20
4.2.1. Análisis de convergencia global cuando M=1 27
4.3. Análisis de convergencia local 32

5. Conclusiones y trabajo futuro **34**

1. Introducción

Consideraremos el problema de determinar una solución $x^* \in \mathbb{R}^n$ que verifique el sistema de ecuaciones no lineales

$$F(x) = 0 \tag{1}$$

donde $F : \mathbb{R}^n \to \mathbb{R}^m$ es una función continuamente diferenciable, su matriz Jacobiana de F en x es Lipschitz continua y $m \leq n$. Llamaremos **sistemas indeterminados** a este tipo de sistemas de ecuaciones no lineales.

En este trabajo estudiamos la teoría de convergencia del método quasi-Newton inexacto propuesto en [7] para resolver sistemas de ecuaciones no lineales no cuadrados. Basados en esa metodología en este trabajo analizaremos las hipótesis necesarias para obtener un resultado de convergencia local para el método quasi-Newton propuesto en [3] para resolver la misma clase de problemas sin el uso de derivadas.

En [3] los autores presentaron dos métodos libres de derivadas para resolver sistemas indeterminados de ecuaciones no lineales. Uno de ellos, llamado DF-QNB, es un método quasi-Newton que utiliza la fórmula adaptada de Broyden de rango uno para la actualización de las matrices. Consideraron la función de merito continuamente diferenciable $f : \mathbb{R}^n \to \mathbb{R}$ que consiste en una medida residual de F en x, $f(x) = \frac{1}{2} \|F(x)\|^2$. El proceso iterativo se basa en una condición de descenso no monótona en la cual no se utiliza información sobre las derivadas. Los autores han demostrado resultados de convergencia global y han analizado el desempeño computacional del método en problemas test de la literatura de pequeño y mediano tamaño mostrando su efectividad.

Notaciones:

- $x \in \mathbb{R}^n$, $x = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$
- $F : \mathbb{R}^n \to \mathbb{R}^m$, $F(x) = \begin{pmatrix} F_1(x) \\ F_2(x) \\ \cdot \\ \cdot \\ \cdot \\ F_m(x) \end{pmatrix}$ donde $F_i : \mathbb{R}^n \to \mathbb{R}$ para cada $i = 1, \ldots, m$.

- $J(x)$ representa a la matriz Jacobiana de F en x, es decir: $J(x) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \cdots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}$

- $f(x) = \frac{1}{2} \|F(x)\|^2$.

- $g(x) = J(x)^t F(x) = \nabla \left(\frac{1}{2} \|F(x)\|^2 \right) = \nabla (f(x))$.

- $\| \cdot \|$ representa la norma 2.

- Si$\{x_k\}_{k \in \mathbb{N}}$ es una sucesión y $K = \{k_1, k_2, k_3, \ldots\}$ es un subconjunto infinito de $\mathbb{N}$ tal que $k_i < k_j$ si $i < j$, denotamos:

$$\lim_{k \in K} x_k = \lim_{j \to \infty} x_{k_j}.$$

- B es una matriz $\mathbb{R}^{m \times n}$ y se puede pensar como una aproximación de la matriz Jacobiana $J(x)$.

2. Generalidades

En esta sección daremos algunos conceptos básicos que serán utilizados durante el presente trabajo.

Definición 2.1. *Si $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ decimos que x^* es **minimizador global** de f en Ω si $f(x^*) \leq f(x)$ para todo $x \in \Omega$.*

Definición 2.2. *Si $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ decimos que x^* es **minimizador local** de f en Ω si $f(x^*) \leq f(x)$ para todo $x \in \Omega \cap B(x^*, r)$.*

Definición 2.3. *Decimos que x^* es **punto estacionario** del problema (1) si es solución de (1) o es un minimizador global de $\|F(x^*) + J(x^*)(x - x^*)\|$. Como esta función objetivo es convexa, es equivalente a decir que x^* es un minimizador local de $\|F(x^*) + J(x^*)(x - x^*)\|$.*

Definición 2.4. *Un Algoritmo se dice **globalmente convergente** si todo punto límite de la sucesión generada por el Algoritmo es un punto estacionario del problema.*

Definición 2.5. *Diremos que un método es **localmente convergente** en relación a un determinado tipo de solución del problema considerado si, dada una solución x^*, existe $\epsilon > 0$ tal que toda sucesión $\{x_k\}$ generada por el Algoritmo converge a x^*, siempre que $\|x_0 - x^*\| \leq \epsilon$.*

Definición 2.6. *Si la sucesión $\{x_k\}$ converge a x^* de manera que*

$$\lim_{k \to \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = C$$

donde $C \in (0,1)$ se dice que la sucesión ***converge linealmente*** *a x^*.*

Definición 2.7. *La convergencia se dice que es* ***superlineal*** *si*

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.$$

Ejemplo 2.8. La sucesión $\{x_k\} = \{1 + k^{-k}\}$ converge superlinealmente a $x^* = 1$.

Definición 2.9. *Si la sucesión $\{x_k\}$ converge a x^* y existen una constante $C > 0$ y $k_0 \geq 0$ tales que*

$$\|x_{k+1} - x^*\| \leq C \|x_k - x^*\|^2$$

para todo $k \geq k_0$, se dice que la sucesión ***converge cuadráticamente*** *a x^*.*

Definición 2.10. *Decimos que $J(x)$ es* ***Lipschitz continua*** *con constante $L > 0$ si*

$$\|J(x) - J(y)\| \leq L \|x - y\|$$

para todo $x, y \in \mathbb{R}^n$.

2.1. El método de Newton

En esta sección esbozaremos brevemente el Método de Newton para resolver sistemas cuadrados de ecuaciones no lineales, asumiendo que el sistema a resolver tiene solución.

Vamos a considerar el problema de determinar $x^* \in \mathbb{R}^n$ que satisfaga $F_c(x) = 0$ donde $F_c : \mathbb{R}^n \to \mathbb{R}^n$. La solución de este problema se irá aproximando por una sucesión de puntos $\{x_k\}$ generada por el método de Newton.

Dada cada aproximación x_k se construye un subproblema con una resolución mas simple. La aproximación x_{k+1} es la solución del subproblema. El subproblema cambia de una iteración a otra y su solución estará, bajo ciertas condiciones, cada vez más próxima a la solución del problema original.

En nuestro problema a considerar, el k-ésimo subproblema viene de considerar la aproximación de Taylor de orden 1 de $F_c(x)$ alrededor del punto x_k, es decir, la función $L_k(x)$ dada por:

$$F_c(x) \approx L_k(x) = F_c(x_k) + J(x_k)(x - x_k).$$

Siguiendo lo que describimos antes, el punto x_{k+1} es una solución de

$$L_k(x) = 0. \tag{2}$$

Si $J(x_k)$ es no singular, (2) tiene solución única. La iteración de Newton consiste en resolver el sistema lineal

$$\begin{aligned} J(x_k)s_k &= -F_c(x_k) \\ x_{k+1} &= x_k + s_k \end{aligned} \tag{3}$$

donde $J(x_k)$ es la matriz Jacobiana que presentamos antes, evaluada en x_k.

El método de Newton posee una propiedad única entre los Algoritmos para resolver sistemas: la invarianza por traslaciones en coordenadas, tanto en el dominio como en el codominio. En el codominio significa que las iteraciones de Newton aplicadas a $F_c(x) = 0$ son las mismas que las aplicadas al sistema $AF_c(x) = 0$ para cualquier matriz A no singular. La invarianza en el dominio consiste en que si $\{x_k\}$ es la sucesión newtoniana para $F_c(x) = 0$, entonces las iteraciones para el sistema $F_c(Ax + b) = 0$ con A no singular y con la aproximación inicial $Ax_0 + b$ son puntos de la forma $Ax_k + b$.

Ventajas y desventajas del método de Newton:

a) Ventajas:

- Es cuadráticamente convergente si el punto inicial x_0 es una buena aproximación de x^* y $J(x^*)$ es no singular.
- Solución exacta en una iteración para una F_c afín.

b) Desventajas:

- No es globalmente convergente para muchos problemas.
- Requiere la evaluación de $J(x_k)$ en cada iteración.
- Cada iteración requiere la solución de un sistema de ecuaciones lineales, que puede ser singular o mal condicionado.

Presentamos a continuación un teorema de convergencia local para el método de Newton:

Teorema 2.11. *Sea $F_c : \mathbb{R}^n \to \mathbb{R}^n$ continuamente diferenciable en un conjunto abierto convexo $D \subset \mathbb{R}^n$. Supongamos que existen $x^* \in \mathbb{R}^n$ y $r, \beta > 0$ tales que $\{x \in \mathbb{R}^n : \|x - x^*\| \leq r\} \subset D, F_c(x^*) = 0$, existe $(J(x^*))^{-1}$ con $\left\|(J(x^*))^{-1}\right\| \leq \beta$ y $J(x)$ es Lipschitz continua con constante γ en el conjunto $\{x \in \mathbb{R}^n : \|x - x^*\| \leq r\}$. Entonces, existe $\epsilon > 0$ tal que para todo $x_0 \in \{x \in \mathbb{R}^n : \|x - x^*\| \leq \epsilon\}$ la sucesión $x_1, x_2, \ldots$ generada por*

$$x_{k+1} = x_k - (J(x_k))^{-1} F_c(x_k) \text{ para todo } k = 0, 1, \ldots$$

está bien definida, converge a x^* y cumple que

$$\|x_{k+1} - x^*\| \leq \beta\gamma \|x_k - x^*\|^2 \text{ para todo } k = 0, 1, \ldots$$

Demostración. Se encuentra demostrado en el Teorema 5.2.1 de [2]. ■

2.2. El método de Newton discreto

Una variación de (3) con practicamente las mismas propiedades teóricas y prácticas que evita el cálculo de derivadas es el llamado **método de Newton discreto**. El esquema de ese

método es análogo a resolver el sistema (3) con la excepción de que las derivadas consideradas no son las analiticas, sino sus aproximaciones por diferencias finitas. Más precisamente, la columna j de $J(x_k)$ es sustituida por

$$\frac{F_c(x_k + he_j) - F_c(x_k)}{h}$$

donde h es un paso pequeño y $\{e_1, \ldots, e_n\}$ es la base canónica de $\mathbb{R}^n$.

La implementación de una iteración del método de Newton discreto, aunque no exige el cálculo de derivadas, si no la evaluación de la función F_c en $n+1$ puntos, puede ser bastante costosa computacionalmente, por eso siempre que sea posible, deben ser utilizadas las derivadas analíticas.

La resolución del sistema lineal (3) cuando la matriz Jacobiana es no singular, se puede obtener vía factorización LU con un costo de $O\left(\frac{n^3}{3}\right)$ operaciones.

Por lo tanto, el trabajo realizado en una iteración del método de Newton consiste en la evaluación de F_c en x_k y sus derivadas, más las $O\left(\frac{n^3}{3}\right)$ operaciones necesarias para resolver (3).

En el caso que $J(x_k)$ sea singular se debe adoptar alguna estrategia para no afectar el procedimiento del método.

Presentamos a continuación un teorema de convergencia local para el método de Newton discreto:

Teorema 2.12. *Sean $F_c : \mathbb{R}^n \to \mathbb{R}^n$ y x^* cumpliendo las hipótesis del Teorema 2.11, entonces existen $\epsilon, h > 0$ tal que si $\{h_k\}$ es una sucesión real con $0 < |h_k| \leq h$ y $x_0 \in \{x \in \mathbb{R}^n : \|x - x^*\| \leq \epsilon\}$, la sucesión $x_1, x_2, \ldots$ generada por*

$$x_{k+1} := x_k - (A_k)^{-1} F_c(x_k) \text{ para } k = 0, 1, \ldots$$

donde

$$(A_k)_j = \frac{F_c(x_k + h_k e_j) - F_c(x_k)}{h_k} \text{ para } j = 1, \ldots, n$$

está bien definida y converge linealmente a x^*. Si

$$\lim_{k \to 0} h_k = 0$$

la convergencia es superlineal.

Si existe alguna constante c_1 tal que

$$|h_k| \leq c_1 \|x_k - x^*\|,$$

o equivalentemente, una constante c_2 tal que

$$|h_k| \leq c_2 \|F_c(x_k)\|$$

entonces la convergencia es cuadrática.

Demostración. Se encuentra demostrado en el Teorema 5.4.1 de [2]. ■

2.3. Métodos quasi-Newton

Se definen como métodos quasi-Newton a aquellos Algoritmos para resolver sistemas cuadrados de ecuaciones cuyas iteraciones tienen el siguiente formato:

$$\begin{aligned} B_k s_k &= -F_c(x_k) \\ x_{k+1} &= x_k + s_k \end{aligned} \tag{4}$$

donde $B_k \in \mathbb{R}^{n \times n}$ es una aproximación de la matriz Jacobiana $J(x_k)$.

Los métodos de quasi-Newton prácticos son aquellos en que B_{k+1}^{-1} puede ser obtenida facilmente a partir de B_k^{-1}, esto es, con a lo sumo $O(n^2)$ operaciones. De esta manera, los cálculos efectuados en (4) podrían ser efectuados con un costo de $O(n^2)$ en términos de tiempo por operación. Algunas implementaciones de los métodos de quasi-Newton trabajan con factorizaciones de las matrices B_k y no con sus inversas. En esos casos, se busca que la factorización de B_{k+1} se pueda obtener a partir de la factorización de B_k en tiempo del orden $O(n^2)$.

De esta forma, vemos que el esfuerzo computacional $O\left(\frac{n^3}{3}\right)$ usado por Newton disminuye a $O(n^2)$ cuando se utilizan métodos quasi-Newton adecuados.

Los **métodos secantes** son una familia de métodos del tipo quasi-Newton. Como hicimos antes, pensemos que en la iteración k la función $F_c(x)$ es aproximada por

$$L_k(x) = F_c(x_k) + B_k(x - x_k).$$

Escribiendo el mismo tipo de aproximación para la iteración $k+1$ tenemos:

$$F_c(x) \approx L_{k+1}(x) = F_c(x_{k+1}) + B_{k+1}(x - x_{k+1}).$$

La idea «secante» consiste en imponer que la función lineal $L_{k+1}(x)$ interpole a la función verdadera en los puntos x_{k+1} y x_k, es decir:

$$\begin{aligned} L_{k+1}(x_{k+1}) &= F_c(x_{k+1}) \ \ y \\ L_{k+1}(x_k) &= F_c(x_k) \end{aligned}$$

La condición $L_{k+1}(x_{k+1}) = F_c(x_{k+1})$ se satisface automáticamente por la definición de L_{k+1}.

La condición $L_{k+1}(x_k) = F_c(x_k)$ es equivalente a

$$F_c(x_k) = F_c(x_{k+1}) + B_{k+1}(x_k - x_{k+1})$$

o a

$$B_{k+1}s_k = y_k \tag{5}$$

donde $y_k = F_c(x_{k+1}) - F_c(x_k)$ y $s_k = x_{k+1} - x_k$. La ecuación (5) es llamada la **ecuación secante**. Podemos pensar (5) como un sistema lineal cuya incógnita es la matriz B_{k+1}. Así interpretado, el sistema tiene n^2 variables (las entradas de B_{k+1}) y n ecuaciones.

Solamente en el caso $n = 1$ el sistema tiene solución única.

Si $n > 1$ y $s_k \neq 0$ hay infinitas matrices B que satisfacen $Bs_k = y_k$. Diferentes elecciones de esa matriz definen **diferentes métodos secantes**. Por ejemplo, si buscamos B_{k+1} de manera que la diferencia $\Delta B_k \equiv B_{k+1} - B_k$ sea una matriz de rango uno, tenemos por (5) que

$$\Delta B_k s_k = y_k - B_k s_k$$

y podemos tomar

$$\Delta B_k = \frac{(y_k - B_k s_k)\, w_k^T}{w_k^T s_k}$$

con $w_k \in \mathbb{R}^n$ arbitrario y no ortogonal a s_k.

La elección $w_k = s_k$ define el **primer método de Broyden**. Esta actualización minimiza $\|B - B_k\|$ sobre todas las matrices B que satisfacen $Bs_k = y_k$ (ver Lema 11.4 de [9]).

Si $w_k = y_k - B_k s_k$ es el método conocido como **corrección simétrica de rango uno**.

Presentamos a continuación un teorema de convergencia local para el método de quasi-Newton:

Teorema 2.13. *Supongamos que todas las hipótesis del Teorema 2.11 se cumplen y que existen constantes positivas ϵ, δ tal que*

$$\|x_0 - x^*\| \leq \epsilon \;\; y \;\; \|B_0 - J(x^*)\| \leq \delta.$$

Entonces la sucesión $\{x_k\}$ generada por el método de Broyden está bien definida y converge superlinealmente a x^*.

Demostración. Se encuentra demostrado en el Teorema 8.2.2 de [2]. ■

Los métodos de Newton y de quasi-Newton **no** son globalmente convergentes para resolver el problema $F_c(x) = 0$. Para garantizar la convergencia global es que se introduce una búsqueda lineal sobre la función de mérito f que va a permitir demostrar la convergencia global del método de quasi-Newton que estudiaremos en una sección siguiente.

3. Búsqueda lineal no monótona libre de derivadas

En esta sección nos ocuparemos de la búsqueda lineal, que en este trabajo será no monótona y libre de derivadas, que se utilizará en el método de quasi-Newton que usaremos en la sección siguiente.

La estrategia se basa en la búsqueda lineal propuesta en [4]. Dada la iteración actual x_k y una dirección de búsqueda d_k, el Algoritmo busca una longitud de paso α_k tal que

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq M-1} f(x_{k-j}) + \eta_k - \gamma \alpha_k^2 f(x_k) \tag{6}$$

donde M es un entero no negativo, $0 < \gamma < 1$, $\eta_k > 0$ y $\sum_{k=0}^{\infty} \eta_k = \eta < \infty$.

Este procedimiento combina la siguiente estrategia de búsqueda lineal no monótona de optimización sin restricciones:

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq M-1} f(x_{k-j}) + \gamma \alpha_k \nabla f(x_k)^t d_k \tag{7}$$

con el esquema de búsqueda lineal libre de derivadas de [5]:

$$\|F(x_k + \alpha_k d_k)\| \leq (1 + \eta_k) \|F(x_k)\| - \gamma \alpha_k^2 \|d_k\|^2 . \tag{8}$$

En este trabajo dada una dirección de búsqueda d_k, basándonos en (7) y (8) consideraremos la siguiente condición de aceptación en la búsqueda lineal:

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq M-1} f(x_{k-j}) + \eta_k - \gamma \alpha_k^2 \|d_k\|^2 ,$$

donde $\|d_k\|^2$ reemplaza a $f(x_k)$ en (6).

3.1. Algoritmo de búsqueda lineal

Dados x_k, $d \in \mathbb{R}^n$, $0 < \tau_{min} < \tau_{max} < 1$, $0 < \gamma < 1$, $M \in \mathbb{N}$, $M > 0$, $\{\eta_k\}$ tal que $\eta_k > 0$ y $\sum_{k=0}^{\infty} \eta_k = \eta < \infty$.

Paso 1: Calcular $\overline{f_k} = max\{f(x_k), \ldots, f(x_{max\{0,k-M+1\}})\}$
Fijar $\alpha_+ = \alpha_- = 1$.

Paso 2: Si $f(x_k + \alpha_+ d) \leq \overline{f_k} + \eta_k - \gamma \alpha_+^2 \|d\|^2$,

definir $d_k = d, \alpha_k = \alpha_+, x_{k+1} = x_k + \alpha_k d_k$
si no
si $f(x_k - \alpha_- d) \leq \overline{f_k} + \eta_k - \gamma\alpha_-^2 \|d\|^2$,
definir $d_k = -d, \alpha_k = \alpha_-, x_{k+1} = x_k + \alpha_k d_k$
si no
elegir $\alpha_{+new} \in [\tau_{min}\alpha_+, \tau_{max}\alpha_+], \alpha_{-new} \in [\tau_{min}\alpha_-, \tau_{max}\alpha_-]$,
definir $\alpha_+ = \alpha_{+new}, \alpha_- = \alpha_{-new}$ e iniciar nuevamente el Paso 2.

Proposición 3.1. *El Algoritmo de búsqueda lineal está bien definido.*

Demostración. Para $\eta_k > 0$ la condición

$$f(x_k + \alpha_+ d) \leq \bar{f}_k + \eta_k - \gamma\alpha_+^2 \|d\|^2$$

se cumple después de un número finito de reducciones de α_+.

Análogamente con

$$f(x_k - \alpha_- d) \leq \bar{f}_k + \eta_k - \gamma\alpha_-^2 \|d\|^2 .$$

■

El siguiente es un esquema general del Algoritmo que se utilizará para resolver el problema (1):

Algoritmo General

Dados $x_0, F(x_0), M \in \mathbb{N}, 0 < \tau_{min} < \tau_{max} < 1, 0 \leq \epsilon < 1, 0 < \gamma < 1, \{\eta_k\}$ tal que $\eta_k > 0$ y $\sum_{k=0}^{\infty} \eta_k = \eta < \infty$.

Sea $k \leftarrow 0$.

Paso 1: Si $\|F(x_k)\| \leq \epsilon \ max\{1, \|F(x_0)\|\}$ parar.

Paso 2: Calcular una dirección de búsqueda d_k.

Paso 3: Buscar α_k y definir $x_{k+1} = x_k + \alpha_k d_k$ usando el Algoritmo de búsqueda lineal.

Actualizar $k \leftarrow k + 1$ y volver al Paso 1.

Proposición 3.2. *Para todo $k \in \mathbb{N}$, consideremos el conjunto*

$$U_k = max\{f(x_{(k-1)M+1}), \ldots, f(x_{kM})\}$$

y definimos $\nu(k) \in \{(k-1)M + 1, \ldots, kM\}$ el índice tal que $f(x_{\nu(k)}) = U_k$ entonces, para todo $k = 1, 2, \ldots$

$$f(x_{\nu(k+1)}) \leq f(x_{\nu(k)}) + \eta$$

donde

$$\eta = \sum_{i=0}^{\infty} \eta_i$$

Demostración. Por el Algoritmo de búsqueda lineal tenemos

$$f(x_{k+1}) = f(x_k + \alpha_+ d) \leq \bar{f}_k + \eta_k - \gamma \alpha_+^2 \|d\|^2$$

o

$$f(x_{k+1}) = f(x_k - \alpha_- d) \leq \bar{f}_k + \eta_k - \gamma \alpha_-^2 \|d\|^2$$

donde $\bar{f}_k = max\{f(x_k), \ldots, f(x_{max\{0,k-M+1\}})\}$, entonces

$$\begin{aligned} f(x_{kM+1}) &\leq max\left\{f(x_{(k-1)(M+1)}), \ldots, f(x_{kM})\right\} + \eta_{kM} - \gamma \alpha_{kM}^2 \|d_{kM}\|^2 \\ &= U_k + \eta_{kM} - \gamma \alpha_{kM}^2 \|d_{kM}\|^2 \end{aligned}$$

como $\gamma > 0$, entonces $-\gamma \alpha_{kM}^2 \|d_{kM}\|^2 \leq 0$, con lo cual

$$f(x_{kM+1}) \leq U_k + \eta_{kM}.$$

Análogamente,

$$\begin{aligned} f(x_{kM+2}) &\leq max\left\{f(x_{(k-1)(M+2)}), \ldots, f(x_{kM+1})\right\} + \eta_{kM+1} - \gamma \alpha_{kM+1}^2 \|d_{kM+1}\|^2 \\ &= max\left\{U_k, f(x_{kM+1})\right\} + \eta_{kM+1} - \gamma \alpha_{kM+1}^2 \|d_{kM+1}\|^2 \end{aligned}$$

como $\gamma > 0$, entonces $-\gamma \alpha_{kM+1}^2 \|d_{kM+1}\|^2 \leq 0$, con lo cual

$$f(x_{kM+2}) \leq max\left\{U_k, f(x_{kM+1})\right\} + \eta_{kM+1}$$

y como vimos que $f(x_{kM+1}) \leq U_k + \eta_{kM}$ obtenemos que

$$f(x_{kM+2}) \leq U_k + \eta_{kM} + \eta_{kM+1}.$$

Haciendo un proceso inductivo:

$$\begin{aligned} f(x_{kM+1}) &\leq U_k + \eta_{kM} \\ f(x_{kM+2}) &\leq U_k + \eta_{kM} + \eta_{kM+1} \\ f(x_{kM+3}) &\leq U_k + \eta_{kM} + \eta_{kM+1} + \eta_{kM+2} \end{aligned}$$

$$\vdots$$

$$f(x_{kM+(l-1)}) \leq U_k + \eta_{kM} + \ldots + \eta_{kM+(l-2)}$$

luego $f(x_{kM+l}) \leq U_k + \eta_{kM} + \ldots + \eta_{kM+(l-1)} - \gamma\alpha^2_{kM+(l-1)} \left\|d_{kM+(l-1)}\right\|^2$.

Obtenemos que:

$$f(x_{kM+l}) \leq U_k + \sum_{j=0}^{l-1} \eta_{kM+j} - \gamma\alpha^2_{kM+(l-1)} \left\|d_{kM+(l-1)}\right\|^2 \text{ para } l = 1, 2, \ldots$$

Por hipótesis, definimos $\nu(k) \in \{(k-1)M+1, \ldots, kM\}$ el índice tal que $f(x_{\nu(k)}) = U_k$, entonces $\nu(k+1) \in \{kM+1, \ldots, (k+1)M\} = \{kM+1, \ldots, kM+M\}$ es el índice tal que $f(x_{\nu(k+1)}) = U_{k+1}$.

$$U_{k+1} = f(x_{\nu(k+1)}) \leq U_k + \sum_{j=0}^{M-1} \eta_{kM+j} - \gamma\alpha^2_{\nu(k+1)-1} \left\|d_{\nu(k+1)-1}\right\|^2$$

entonces, para todo $k = 1, 2, \ldots$

$$f(x_{\nu(k+1)}) \leq f(x_{\nu(k)}) + \sum_{j=0}^{M-1} \eta_{kM+j} - \gamma\alpha^2_{\nu(k+1)-1} \left\|d_{\nu(k+1)-1}\right\|^2.$$

Como $\gamma > 0$, entonces $-\gamma\alpha^2_{\nu(k+1)-1} \left\|d_{\nu(k+1)-1}\right\|^2 \leq 0$, con lo cual

$$\begin{aligned} f(x_{\nu(k+1)}) &\leq f(x_{\nu(k)}) + \sum_{j=0}^{M-1} \eta_{kM+j} \\ &\leq f(x_{\nu(k)}) + \sum_{j=0}^{\infty} \eta_{kM+j} \\ &\leq f(x_{\nu(k)}) + \eta \end{aligned}$$

como queriamos demostrar. ■

Proposición 3.3. *Se cumple que*

$$\lim_{k\to\infty} \alpha^2_{\nu(k)-1} \left\|d_{\nu(k)-1}\right\|^2 = 0.$$

Demostración. Vimos en la demostración anterior que para todo $k = 1, 2, \ldots$

$$f(x_{\nu(k+1)}) \leq f(x_{\nu(k)}) + \sum_{j=0}^{M-1} \eta_{kM+j} - \gamma\alpha^2_{\nu(k+1)-1} \left\|d_{\nu(k+1)-1}\right\|^2$$

Si consideramos la última desigualdad para $k = 1, 2, \ldots, L$ obtenemos que:

$k = 1$,

$$f(x_{\nu(2)}) \leq f(x_{\nu(1)}) + \sum_{j=0}^{M-1} \eta_{M+j} - \gamma\alpha^2_{\nu(2)-1} \left\|d_{\nu(2)-1}\right\|^2,$$

$k = 2$,

$$f(x_{\nu(3)}) \leq f(x_{\nu(2)}) + \sum_{j=0}^{M-1} \eta_{2M+j} - \gamma\alpha^2_{\nu(3)-1} \left\|d_{\nu(3)-1}\right\|^2,$$

$$\vdots$$

$k = L - 1$,

$$f(x_{\nu(L)}) \leq f(x_{\nu(L-1)}) + \sum_{j=0}^{M-1} \eta_{(L-1)M+j} - \gamma\alpha^2_{\nu(L)-1} \left\|d_{\nu(L)-1}\right\|^2,$$

$k = L$,

$$f(x_{\nu(L+1)}) \leq f(x_{\nu(L)}) + \sum_{j=0}^{M-1} \eta_{LM+j} - \gamma\alpha^2_{\nu(L+1)-1} \left\|d_{\nu(L+1)-1}\right\|^2.$$

Por lo tanto,

$$\begin{aligned}
f(x_{\nu(L+1)}) &\leq f(x_{\nu(L-1)}) + \sum_{j=0}^{M-1} \eta_{(L-1)M+j} - \gamma\alpha^2_{\nu(L)-1} \left\|d_{\nu(L)-1}\right\|^2 + \\
&\sum_{j=0}^{M-1} \eta_{LM+j} - \gamma\alpha^2_{\nu(L+1)-1} \left\|d_{\nu(L+1)-1}\right\|^2 \\
&\leq \ldots \leq f(x_{\nu(1)}) + \sum_{j=0}^{M-1} \eta_{M+j} - \gamma\alpha^2_{\nu(2)-1} \left\|d_{\nu(2)-1}\right\|^2 + \ldots + \\
&\sum_{j=0}^{M-1} \eta_{(L-1)M+j} - \gamma\alpha^2_{\nu(L)-1} \left\|d_{\nu(L)-1}\right\|^2 + \sum_{j=0}^{M-1} \eta_{LM+j} - \gamma\alpha^2_{\nu(L+1)-1} \left\|d_{\nu(L+1)-1}\right\|^2 \\
&= f(x_{\nu(1)}) + \sum_{j=0}^{M-1} \eta_{M+j} + \ldots + \sum_{j=0}^{M-1} \eta_{(L-1)M+j} + \sum_{j=0}^{M-1} \eta_{LM+j} - \gamma\alpha^2_{\nu(2)-1} \left\|d_{\nu(2)-1}\right\|^2 - \\
&\ldots - \gamma\alpha^2_{\nu(L)-1} \left\|d_{\nu(L)-1}\right\|^2 - \gamma\alpha^2_{\nu(L+1)-1} \left\|d_{\nu(L+1)-1}\right\|^2.
\end{aligned}$$

Reescribiendo las sumatorias bajo un mismo símbolo de sumatoria obtenemos que:

$$\sum_{j=0}^{M-1} \eta_{1M+j} + \ldots + \sum_{j=0}^{M-1} \eta_{2M+j} + \ldots + \sum_{j=0}^{M-1} \eta_{(L-1)M+j} + \sum_{j=0}^{M-1} \eta_{LM+j} = \sum_{j=M}^{(L+1)M-1} \eta_j$$

pues:

$$\begin{aligned}
\sum_{j=M}^{(L+1)M-1} \eta_j &= \eta_M + \eta_{M+1} + \ldots + \eta_{(L+1)M-1} \\
&= \eta_M + \eta_{M+1} + \ldots + \eta_{LM+M-1}
\end{aligned}$$

y

$$\sum_{j=0}^{M-1} \eta_{1M+j} + \ldots + \sum_{j=0}^{M-1} \eta_{2M+j} + \ldots + \sum_{j=0}^{M-1} \eta_{(L-1)M+j} + \sum_{j=0}^{M-1} \eta_{LM+j} =$$

$$\eta_{1M} + \ldots + \eta_{1M+M-1} + \eta_{2M} + \ldots + \eta_{2M+M-1} + \ldots + \eta_{(L-1)M} + \ldots$$

$$+\eta_{(L-1)M+(M-1)} + \eta_{LM} + \ldots + \eta_{LM+M-1} =$$

$$\eta_M + \ldots + \eta_{2M-1} + \eta_{2M} + \ldots + \eta_{3M-1} + \ldots + \eta_{LM-M} + \ldots$$

$$+\eta_{LM-M+M-1} + \eta_{LM} + \ldots + \eta_{LM+M-1}.$$

Analizando las restas:

$$-\gamma\alpha^2_{\nu(2)-1} \left\|d_{\nu(2)-1}\right\|^2 - \gamma\alpha^2_{\nu(3)-1} \left\|d_{\nu(3)-1}\right\|^2 - \ldots - \gamma\alpha^2_{\nu(L)-1} \left\|d_{\nu(L)-1}\right\|^2 -$$
$$\gamma\alpha^2_{\nu(L+1)-1} \left\|d_{\nu(L+1)-1}\right\|^2 = -\gamma \sum_{j=2}^{L+1} \alpha^2_{\nu(j)-1} \left\|d_{\nu(j)-1}\right\|^2.$$

Con lo cual,

$$f(x_{\nu(L+1)}) \leq f(x_{\nu(1)}) + \sum_{j=M}^{(L+1)M-1} \eta_j - \gamma \sum_{j=2}^{L+1} \alpha^2_{\nu(j)-1} \left\|d_{\nu(j)-1}\right\|^2.$$

Despejando,

$$\gamma \sum_{j=2}^{L+1} \alpha^2_{\nu(j)-1} \left\|d_{\nu(j)-1}\right\|^2 \leq f(x_{\nu(1)}) + \sum_{j=M}^{(L+1)M-1} \eta_j - f(x_{\nu(L+1)})$$

como f es positiva,

$$\begin{aligned}\gamma \sum_{j=2}^{L+1} \alpha^2_{\nu(j)-1} \left\|d_{\nu(j)-1}\right\|^2 &\leq f(x_{\nu(1)}) + \sum_{j=M}^{(L+1)M-1} \eta_j \\ &\leq f(x_{\nu(1)}) + \sum_{j=M}^{\infty} \eta_j \\ &= f(x_{\nu(1)}) + \eta.\end{aligned}$$

Como $\gamma \sum_{j=2}^{L+1} \alpha^2_{\nu(j)-1} \left\|d_{\nu(j)-1}\right\|^2$ está acotada superiormente por $f(x_{\nu(1)}) + \eta$ y es creciente, la serie

$$\sum_{j=2}^{\infty} \alpha^2_{\nu(j)-1} \left\|d_{\nu(j)-1}\right\|^2$$

es convergente. Con lo cual,

$$\lim_{k\to\infty} \alpha^2_{\nu(k)-1} \left\| d_{\nu(k)-1} \right\|^2 = 0$$

■

Proposición 3.4. *La sucesión $\{x_k\}$ generada por el Algoritmo General está contenida en el conjunto*

$$\Omega = \{x \in \mathbb{R}^n \text{ tal que } f(x) \leq f(x_{\nu(1)}) + \eta\}.$$

Demostración. En la demostración de la Proposición (3.3) obtuvimos la siguiente desigualdad:

$$f(x_{\nu(j)}) \leq f(x_{\nu(1)}) + \eta \text{ para todo } j \geq 1.$$

Por lo tanto, para $k \in \mathbb{N}$ existe $j_0 \in \mathbb{N}$, $j_0 \geq 1$ tal que

$$f(x_k) \leq f(x_{\nu(1)}) + \eta$$

y como

$$f(x_{\nu(j_0)}) \leq f(x_{\nu(1)}) + \eta$$

tenemos que

$$f(x_k) \leq f(x_{\nu(1)}) + \eta.$$

La desigualdad anterior vale para todo k, con lo cual, la sucesión $\{x_k\}$ está contenida en Ω.

■

3.2. Análisis para el caso M=1

Analizaremos el caso en el que la búsqueda lineal libre de derivadas utilizada en el Algoritmo de búsqueda lineal considera el caso $M = 1$.

Por la presencia del η_k la búsqueda lineal sigue siendo no monótona, pero impone un comportamiento casi monótono de la función de mérito cuando x_k está cerca de la solución.

Proposición 3.5. *Si tomamos $M = 1$ en el Algoritmo de búsqueda lineal, entonces:*

- *la sucesión $\{x_k\}$ generada por el Algoritmo General, está contenida en*

$$\Omega = \{x \in \mathbb{R}^n \text{ tal que } f(x) \leq f(x_0) + 2\eta\}.$$

- *la serie $\sum_{k=1}^{\infty} \alpha_k^2 \|d_k\|^2$ es convergente.*

Demostración. Vimos en la Proposición 3.4 que

$$\{x_k\} \subset \Omega = \{x \in \mathbb{R}^n \ tal \ que \ f(x) \leq f(x_{\nu(1)}) + \eta\}$$

Tomando $M = 1$ en el Algoritmo obtenemos que $f(x) \leq f(x_1) + \eta$ y

$$f(x_1) = f(x_0 + \alpha_0 d_0) \leq f(x_0) + \eta_0 - \gamma \alpha_0^2 \|d_0\|^2,$$

con lo cual,

$$\{x_k\} \subset \Omega = \{x \in \mathbb{R}^n \ tal \ que \ f(x) \leq f(x_0) + 2\eta\}.$$

También vimos en la Demostración de la Proposición 3.3 que la serie

$$\sum_{j=1}^{\infty} \alpha_{\nu(j+1)-1}^2 \left\|d_{\nu(j+1)-1}\right\|^2$$

es convergente. Tomando $M = 1$ en el Algoritmo obtenemos que la serie

$$\sum_{k=1}^{\infty} \alpha_k^2 \|d_k\|^2$$

es convergente. ■

Observación 3.6. Cuando $M = 1$ en el Algoritmo de búsqueda lineal, el esquema de búsqueda es el siguiente:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \eta_k - \gamma \alpha_k^2 \|d_k\|^2$$

y se verifica que $\eta_k \to 0$ y $\alpha_k^2 \|d_k\|^2 \to 0$. Entonces se observa que la búsqueda lineal impone un comportamiento casi monótono cuando el x_k está cerca de la solución.

4. Método de quasi-Newton libre de derivadas usando la fórmula de actualización de Broyden (DF-QNB)

Recordemos que el problema es determinar una solución $x^* \in \mathbb{R}^n$ que verifique

$$F(x) = 0$$

donde $F : \mathbb{R}^n \to \mathbb{R}^m$ es una función continuamente diferenciable y $m \leq n$.

En esta sección vamos a definir un método quasi-Newton basado en la fórmula de actualización de Broyden para las matrices, con la búsqueda libre de derivadas definida en el Algoritmo de búsqueda lineal y estableceremos resultados de convergencia.

Vamos a definir la dirección de búsqueda como una solución aproximada del sistema lineal

$$B_k s = -F(x_k)$$

y usaremos una **fórmula de actualización de Broyden**:

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)\, s_k^t}{s_k^t s_k} \tag{9}$$

donde:

$$s_k = x_{k+1} - x_k$$
$$y_k = F(x_{k+1}) - F(x_k).$$

El objetivo es resolver el sistema lineal teoricamente, en forma exacta cuando sea posible y de manera aproximada en otro caso.

4.1. Algoritmo DF-QNB

Dados $x_0 \in \mathbb{R}^n$, $B_0 \in \mathbb{R}^{m \times n}$, $0 < \gamma < 1$, $0 \leq \theta_0 < 1$, $\Delta > 0$, $0 < \tau_{min} < \tau_{max} < 1$, $0 \leq \epsilon < 1$.

Sea $k \leftarrow 0$.

Paso 1: Si $\|F(x_k)\| \leq \epsilon\, max\{1, \|F(x_0)\|\}$, parar.

Paso 2: Calcular la dirección d_k.

Paso 2.1: Buscar d tal que

$$B_k d + F(x_k) = 0 \text{ y } \|d\| \leq \Delta$$

Si tal dirección es encontrada, definir: $d_k = d$, $\theta_{k+1} = \theta_k$ e ir al Paso 3.

Paso 2.2: Buscar una solución d de

$$\min_{d \in \mathbb{R}^n} \|B_k d + F(x_k)\|$$

Si d satisface que

$$\|B_k d + F(x_k)\| \leq \theta_k \|F(x_k)\| \text{ y } \|d\| \leq \Delta$$

definir $d_k = d, \theta_{k+1} = \theta_k$ e ir al Paso 3.

Paso 2.3: Si no se encuentra una dirección d que cumpla las condiciones anteriores, considerar $d_k = 0$, $x_{k+1} = x_k$, $\theta_{k+1} = \frac{\theta_k + 1}{2} > \theta_k$ e ir al Paso 5.

UNIVERSIDAD NACIONAL DE LA PLATA ★ FACULTAD DE CIENCIAS EXACTAS ★ DEPARTAMENTO DE MATEMÁTICA BIBLIOTECA
8868

Paso 3: Buscar α_k y $x_{k+1} = x_k + \alpha_k d_k$ usando el Algoritmo de búsqueda lineal.

Paso 4: Actualizar B_{k+1} usando la fórmula de actualización de Broyden.

Paso 5: Actualizar $k \leftarrow k + 1$ e ir al Paso 1.

4.2. Análisis de convergencia global

A continuación analizaremos teoremas que establecen las hipótesis necesarias para la obtención de resultados de convergencia global.

Teorema 4.1. *Supongamos que el Algoritmo DF-QNB genera una sucesión infinita* $\{x_k\}$, *que*

$$\lim_{k\to\infty} \langle (B_k - J(x_k))\, d_k, F(x_k)\rangle = 0 \tag{10}$$

y que existe $k_0 \in \mathbb{N}$ *tal que, para todo* $k \geq k_0$, $\theta_k = \overline{\theta} < 1$, *entonces, todo punto límite de* $\{x_k\}_{k\in K}$ *es una solución de* $F(x) = 0$, *donde* $K = \{\nu(1) - 1, \nu(2) - 1, \nu(3) - 1, \ldots\}$.

Demostración. Sea x^* un punto límite de $\{x_k\}_{k\in K}$. Queremos ver que x^* cumple que $F(x^*) = 0$.

Como x^* es punto límite de $\{x_k\}_{k\in K}$, existe un conjunto $K_1 \subset K$ tal que $\lim_{k\in K_1} x_k = x^*$.
Por la Proposición 3.3 tenemos que $\lim_{k\in K} \alpha_k^2 \|d_k\|^2 = 0$. Como vale sobre $k \in K$ y $K_1 \subset K$, vale también sobre $k \in K_1$ entonces

$$\lim_{k\in K_1} \alpha_k^2 \|d_k\|^2 = 0.$$

Pueden pasar dos casos:

a) $\lim_{k\in K_1} \alpha_k \neq 0$

b) $\lim_{k\in K_1} \alpha_k = 0.$

Analizamos cada uno de esos casos por separado:

a) Si $\lim_{k\in K_1} \alpha_k \neq 0 \Rightarrow \lim_{k\in K_1} \|d_k\|^2 = 0 \Rightarrow \lim_{k\in K_1} \|d_k\| = 0.$

Usaremos la última implicación

$$\lim_{k\in K_1} \|d_k\| = 0 \tag{11}$$

La dirección d_k se calcula en el Paso 2 del Algoritmo DF-QNB. Pueden pasar 2 situaciones en la búsqueda de tal dirección:

1) $\{k \in \mathbb{N} : d_k$ es obtenida como solución del sistema $B_k d + F(x_k) = 0\}$ es finito.

2) $\{k \in \mathbb{N} : d_k$ es obtenida como solución del sistema $B_k d + F(x_k) = 0\}$ es infinito.

1) Que $\{k \in \mathbb{N} : d_k$ es obtenida como solución del sistema $B_k d + F(x_k) = 0\}$ sea finito, indica que el Algoritmo pasó finitas veces por el Paso 2.1, con lo cual existe $k_1 \in K_1$ tal que para todo $k \geq k_1, k \in K_1$, d_k verifica la fórmula

$$\|B_k d + F(x_k)\| \leq \theta_k \|F(x_k)\| \quad \forall k \geq k_1 \tag{12}$$

$\forall k \in K_1, k \geq max\{k_0, k_1\}$ tenemos que

$$\langle B_k d_k, F(x_k)\rangle \leq \tfrac{\bar{\theta}^2 - 1}{2} \|F(x_k)\|^2$$

pues:

$$\begin{aligned}
\|B_k d + F(x_k)\|^2 &= \langle B_k d + F(x_k), B_k d + F(x_k)\rangle \\
&= \langle B_k d, B_k d\rangle + \langle B_k d, F(x_k)\rangle + \langle F(x_k), B_k d\rangle + \langle F(x_k), F(x_k)\rangle \\
&= \langle B_k d, B_k d\rangle + 2\langle B_k d, F(x_k)\rangle + \langle F(x_k), F(x_k)\rangle \\
&= \|B_k d\|^2 + 2\langle B_k d, F(x_k)\rangle + \|F(x_k)\|^2
\end{aligned}$$

como $\|B_k d\|^2 \geq 0$,

$$2\langle B_k d, F(x_k)\rangle + \|F(x_k)\|^2 \leq \|B_k d\|^2 + 2\langle B_k d, F(x_k)\rangle + \|F(x_k)\|^2$$

y por (12),

$$\begin{aligned}
2\langle B_k d, F(x_k)\rangle + \|F(x_k)\|^2 &\leq \theta_k^2 \|F(x_k)\|^2 \\
&\leq \bar{\theta}^2 \|F(x_k)\|^2
\end{aligned}$$

entonces,

$$\begin{aligned}
2\langle B_k d, F(x_k)\rangle &\leq \bar{\theta}^2 \|F(x_k)\|^2 - \|F(x_k)\|^2 \\
&= (\bar{\theta}^2 - 1) \|F(x_k)\|^2
\end{aligned}$$

con lo cual,

$$\langle B_k d, F(x_k)\rangle \leq \frac{\bar{\theta}^2 - 1}{2} \|F(x_k)\|^2 .$$

Esto implica que,

$$\langle (B_k - J(x_k))\, d_k, F(x_k)\rangle + \langle J(x_k) d_k, F(x_k)\rangle \leq \frac{\bar{\theta}^2 - 1}{2} \|F(x_k)\|^2 ,$$

pues:

$$\begin{aligned}\langle B_k d_k, F(x_k)\rangle &= \langle B_k d_k, F(x_k)\rangle - \langle J(x_k)d_k, F(x_k)\rangle + \langle J(x_k)d_k, F(x_k)\rangle \\ &= \langle B_k d_k - J(x_k)d_k, F(x_k)\rangle + \langle J(x_k)d_k, F(x_k)\rangle \\ &\leq \frac{\bar{\theta}^2 - 1}{2} \|F(x_k)\|^2 .\end{aligned}$$

Tomando límites para $k \in K_1, k \geq max\{k_0, k_1\}$ en

$$\langle B_k d_k - J(x_k)d_k, F(x_k)\rangle + \langle J(x_k)d_k, F(x_k)\rangle \leq \frac{\bar{\theta}^2 - 1}{2} \|F(x_k)\|^2$$

usando la continuidad de $F(x)$ y de $J(x)$ y los límites (10) y (11), obtenemos que

$$0 \leq \frac{\bar{\theta}^2 - 1}{2} \|F(x^*)\|^2 .$$

Como $\bar{\theta} < 1$, tenemos que $\|F(x^*)\| = 0$.

2) Como $\{k \in \mathbb{N} : d_k$ es obtenida como solución del sistema $B_k d + F(x_k) = 0\}$ es infinito, entonces existe K_2 subconjunto de K_1 tal que para todo $k \in K_2$, d_k se obtuvo de resolver el sistema $B_k d + F(x_k) = 0$. Entonces:

$$\langle (B_k - J(x_k))\, d_k, F(x_k)\rangle + \langle J(x_k)d_k, F(x_k)\rangle = -\|F(x_k)\|^2 \quad \forall k \in K_2$$

pues:

$$\begin{aligned}\langle (B_k - J(x_k))\, d_k, F(x_k)\rangle + \langle J(x_k)d_k, F(x_k)\rangle &= \langle B_k d_k - J(x_k)d_k, F(x_k)\rangle + \langle J(x_k)d_k, F(x_k)\rangle \\ &= \langle B_k d_k, F(x_k)\rangle - \langle J(x_k)d_k, F(x_k)\rangle + \langle J(x_k)d_k, F(x_k)\rangle \\ &= \langle B_k d_k, F(x_k)\rangle \\ &= \langle -F(x_k), F(x_k)\rangle \\ &= -\langle F(x_k), F(x_k)\rangle \\ &= -\|F(x_k)\|^2 .\end{aligned}$$

Tomando límites para $k \in K_2, k \geq k_0$, usando la continuidad de $F(x)$ y de $J(x)$ y (10) y (11) obtenemos que $\|F(x^*)\| = 0$.

b) Veamos el caso $\lim_{k \in K_1} \alpha_k = 0$

Como $\{d_k\}$ es acotada para todo $k \in \mathbb{N}$ ($\|d_k\| \leq \Delta$, para todo $k \in \mathbb{N}$ por el Algoritmo), entonces $\{d_k\}$ es acotada para todo $k \in K_1$. Por lo tanto, la sucesión $\{d_k\}_{k \in K_1}$ tiene una subsucesión convergente, es decir, existe $K_2 \subset K_1$ y $\bar{d} \in \mathbb{R}^n$ tal que $\lim_{k \in K_2} d_k = \bar{d}$.

En la búsqueda lineal, para elegir el paso α_k, el Algoritmo DF-QNB testea las siguientes desigualdades:

$$f(x_k + \alpha_+ d_k) \leq \bar{f}_k + \eta_k - \gamma \alpha_+^2 \|d_k\|^2 \tag{13}$$

$$f(x_k - \alpha_- d_k) \leq \bar{f}_k + \eta_k - \gamma \alpha_-^2 \|d_k\|^2 \tag{14}$$

en el Algoritmo de búsqueda lineal.

Como $\lim_{k \in K_2} \alpha_k = 0$, existe $\bar{\alpha} \in K_2$ tal que $\alpha_k < 1$ para todo $k \in K_2, k \geq \bar{k}$, entonces, para esas iteraciones k, existen pasos α_+ y α_- que no satisfacen (13) y (14) y $\lim_{k \in K_2} \alpha_k^+ = \lim_{k \in K_2} \alpha_k^- = 0$ (achico el paso del α_k^+, como $\alpha_k \to 0$, entonces $\alpha_k^+ \to 0$).

Entonces se tienen las siguientes desigualdades:

$$f(x_k + \alpha_k^+ d_k) > \bar{f}_k + \eta_k - \gamma \left(\alpha_k^+\right)^2 \|d_k\|^2 \tag{15}$$

$$f(x_k - \alpha_k^- d_k) > \bar{f}_k + \eta_k - \gamma \left(\alpha_k^-\right)^2 \|d_k\|^2 . \tag{16}$$

Trabajando con la ecuación (15) obtenemos que:

$$\begin{aligned} f(x_k + \alpha_k^+ d_k) &> \bar{f}_k + \eta_k - \gamma \left(\alpha_k^+\right)^2 \|d_k\|^2 \\ &> f(x_k) - \gamma \left(\alpha_k^+\right)^2 \|d_k\|^2 \end{aligned}$$

pues $\bar{f}_k = max\{f(x_k), \ldots, f(x_{max\{0,k-M+1\}})\}$ y $\eta_k > 0$.

$$\begin{aligned} f(x_k + \alpha_k^+ d_k) &> f(x_k) - \gamma \left(\alpha_k^+\right)^2 \|d_k\|^2 \\ f(x_k + \alpha_k^+ d_k) - f(x_k) &> -\gamma \left(\alpha_k^+\right)^2 \|d_k\|^2 \\ \frac{f(x_k + \alpha_k^+ d_k) - f(x_k)}{\alpha_k^+} &> -\gamma \left(\alpha_k^+\right) \|d_k\|^2 . \end{aligned}$$

Como $\|d_k\|^2 \leq \Delta^2$ entonces $-\|d_k\|^2 \geq -\Delta^2$ con lo cual $-\|d_k\|^2 \gamma \left(\alpha_k^+\right)^2 \geq -\Delta^2 \gamma \left(\alpha_k^+\right)^2$. Usamos esta última expresión para acotar y obtenemos:

$$\frac{f(x_k + \alpha_k^+ d_k) - f(x_k)}{\alpha_k^+} > -\gamma \left(\alpha_k^+\right) \Delta^2 .$$

Por el Teorema del Valor Medio, existe $\xi_k^+ \in [0, 1]$ tal que

$$\left\langle g(x_k + \xi_k^+ \alpha_k^+ d_k), d_k \right\rangle > -\gamma \alpha_k^+ \Delta^2 .$$

Trabajando de la misma manera que hicimos con (15), pero ahora con (16) obtenemos que

$$\begin{aligned} f(x_k - \alpha_k^- d_k) &> \bar{f}_k + \eta_k - \gamma \left(\alpha_k^-\right)^2 \|d_k\|^2 \\ &> f(x_k) - \gamma \left(\alpha_k^-\right)^2 \|d_k\|^2 \end{aligned}$$

pues $\bar{f}_k = max\{f(x_k), \ldots, f(x_{max\{0,k-M+1\}})\}$ y $\eta_k > 0$.

$$\begin{aligned} f(x_k - \alpha_k^- d_k) &> f(x_k) - \gamma \left(\alpha_k^-\right)^2 \|d_k\|^2 \\ f(x_k - \alpha_k^- d_k) - f(x_k) &> -\gamma \left(\alpha_k^-\right)^2 \|d_k\|^2 \\ \frac{f(x_k - \alpha_k^- d_k) - f(x_k)}{\alpha_k^-} &> -\gamma \left(\alpha_k^-\right) \|d_k\|^2 \\ - \frac{f(x_k) - f(x_k - \alpha_k^- d_k)}{\alpha_k^-} &> -\gamma \left(\alpha_k^-\right) \|d_k\|^2 \\ \frac{f(x_k) - f(x_k - \alpha_k^- d_k)}{\alpha_k^-} &< \gamma \left(\alpha_k^-\right) \|d_k\|^2 \end{aligned}$$

como $\|d_k\|^2 \leq \Delta^2$ entonces

$$\frac{f(x_k) - f(x_k - \alpha_k^- d_k)}{\alpha_k^-} < \gamma \left(\alpha_k^-\right) \Delta^2.$$

Por el Teorema del Valor Medio, existe $\xi_k^- \in [0,1]$ tal que

$$\left\langle g(x_k - \xi_k^- \alpha_k^- d_k), d_k \right\rangle < \gamma \alpha_k^- \Delta^2.$$

Obtuvimos:

$$\left\langle g(x_k + \xi_k^+ \alpha_k^+ d_k), d_k \right\rangle > -\gamma \alpha_k^+ \Delta^2 \tag{17}$$

$$\left\langle g(x_k - \xi_k^- \alpha_k^- d_k), d_k \right\rangle < \gamma \alpha_k^- \Delta^2. \tag{18}$$

Tomando límite en (17) cuando $k \to \infty$, $k \in K_2$ obtenemos que:

$$\left\langle g(x^*), \bar{d} \right\rangle \geq 0 \tag{19}$$

pues $\lim_{k \to \infty} \alpha_k^+ = 0$.

Tomando límite en (18) cuando $k \to \infty$, $k \in K_2$ obtenemos que:

$$\left\langle g(x^*), \bar{d} \right\rangle \leq 0 \tag{20}$$

pues $\lim_{k \to \infty} \alpha_k^- = 0$.

Por lo tanto, por (19) y (20)

$$\left\langle g(x^*), \bar{d} \right\rangle = 0$$

$$\langle g(x^*), \bar{d} \rangle = \langle J(x^*)^t F(x^*), \bar{d} \rangle = 0.$$
$$\Rightarrow \langle J(x^*)\bar{d}, F(x^*) \rangle = 0. \tag{21}$$

Ahora, como hicimos antes en el caso a), hay que considerar dos casos:

i) $\{ k \in \mathbb{N} : d_k$ es obtenida como solución del sistema $B_k d + F(x_k) = 0 \}$ es finito.

ii) $\{ k \in \mathbb{N} : d_k$ es obtenida como solución del sistema $B_k d + F(x_k) = 0 \}$ es infinito.

La demostración es análoga al caso $\lim_{k \in K_1} \alpha_k \neq 0$ y usando la hipótesis del teorema y la ecuación (21) obtenemos que $\|F(x^*)\| = 0$, pues:

$$\begin{aligned}
\lim_{k\to\infty} \langle (B_k - J(x_k))\, d_k, F(x_k) \rangle &= 0 \\
\lim_{k\to\infty} \langle B_k d_k - J(x_k) d_k, F(x_k) \rangle &= 0 \\
\lim_{k\to\infty} \langle B_k d_k, F(x_k) \rangle - \langle J(x_k) d_k, F(x_k) \rangle &= 0 \\
\lim_{k\to\infty} \langle B_k d_k, F(x_k) \rangle - \lim_{k\to\infty} \langle J(x_k) d_k, F(x_k) \rangle &= 0 \\
\lim_{k\to\infty} \langle -F(x_k), F(x_k) \rangle - \langle J(x^*)\bar{d}, F(x^*) \rangle &= 0
\end{aligned}$$

Por (21)

$$\begin{aligned}
-\langle F(x^*), F(x^*) \rangle - 0 &= 0 \\
-\|F(x^*)\|^2 &= 0 \\
\|F(x^*)\| &= 0.
\end{aligned}$$

■

Si no podemos garantizar que $\theta_k = \bar{\theta} < 1$ para todo $k \geq k_0$, entonces podemos garantizar el siguiente resultado:

Teorema 4.2. *Supongamos que en el Algoritmo DF-QNB, θ_k se incrementa un número infinito de veces y definimos:*

$$K_* = \{k \in \mathbb{N} \text{ tal que } \theta_{k+1} > \theta_k\}.$$

Supongamos además que

$$\lim_{k \in K_*} \|B_k - J(x_k)\| = 0$$

entonces, todo punto límite x^* de la sucesión $\{x_k\}_{k \in K_*}$ es una solución de F(x) = 0 o es un minimizador global de $\|F(x^*) + J(x^*)(x - x^*)\|$.

Demostración. Si $F(x^*) = 0$ ya estaría pues x^* sería solución de $F(x) = 0$.

Supongamos que $F(x^*) \neq 0$ entonces $\|F(x^*)\| > 0$. Queremos ver que x^* es minimizador global de $\|F(x^*) + J(x^*)(x - x^*)\|$.

Supongamos por el absurdo que x^* **no** es minimizador global de $\|F(x^*) + J(x^*)(x - x^*)\|$. En particular no es minimizador global de $\|F(x^*) + J(x^*)(x - x^*)\|$ sujeto a $\|x - x^*\| = \|d\| \leq \frac{\Delta}{2}$. Por lo tanto existe d tal que $\|d\| \leq \frac{\Delta}{2}$ y

$$\|F(x^*) + J(x^*)d\| < \|F(x^*)\|$$

con lo cual,

$$\frac{\|F(x^*) + J(x^*)d\|}{\|F(x^*)\|} < 1$$

entonces existe $r \in (0, 1)$ tal que

$$\frac{\|F(x^*) + J(x^*)d\|}{\|F(x^*)\|} \leq r < 1.$$

Por hipótesis sabemos que $\lim_{k \in K_*} \|B_k - J(x_k)\| = 0$, entonces

$$\frac{\|B_k (x^* + d - x_k) + F(x_k)\|}{\|F(x_k)\|} \underset{k \in K_*}{\to} \frac{\|J(x^*)d + F(x^*)\|}{\|F(x^*)\|} \leq r.$$

Por lo tanto podemos afirmar que

$$\frac{\|B_k (x^* + d - x_k) + F(x_k)\|}{\|F(x_k)\|} \leq \frac{r+1}{2}$$

para todo $k \in K_*$ suficientemente grande.

Como x^* es punto de acumulación de $\{x_k\}_{k \in K_*}$, para todo k suficientemente grande,

$$\|x^* - x_k + d\| \leq \Delta$$

pues

$$\|x^* - x_k + d\| \leq \|x^* - x_k\| + \|d\| \leq \frac{\Delta}{2} + \frac{\Delta}{2} = \Delta.$$

Entonces encontramos d que cumple que $\|B_k d + F(x_k)\| \leq \theta_k \|F(x_k)\|$ con $k \in K_*$ y $\|d\| \leq \Delta$, con lo cual $\theta_{k+1} = \theta_k$ produciendo un absurdo, pues para esas iteraciones de K_*, el θ_k es incrementado.

Por lo tanto, x^* es minimizador global de $\|F(x^*) + J(x^*)(x - x^*)\|$. ■

4.2.1. Análisis de convergencia global cuando M=1

A continuación analizaremos resultados de convergencia global cuando en el Algoritmo de búsqueda lineal consideramos $M = 1$.

Lema 4.3. *Supongamos que el conjunto* $\Omega = \{x \in \mathbb{R}^n \text{ tal que } f(x) \leq f(x_0) + 2\eta\}$ *es acotado y que* $J(x)$ *es Lipschitz continua en* Ω. *Si se verifica que*

$$\sum_{k=1}^{\infty} \|s_k\|^2 = \sum_{k=1}^{\infty} \alpha_k^2 \|d_k\|^2 < \infty$$

entonces

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} \rho_i^2 = 0$$

donde

$$\rho_k = \frac{\|(A_{k+1} - B_k)\, d_k\|}{\|d_k\|}.$$

considerando

$$A_{k+1} = \int_0^1 J(x_k + ts_k)\ dt.$$

En particular, existe una subsucesión de $\{\rho_k\}$ *tendiendo a cero.*

Demostración. Se encuentra demostrado en el Lema 2.6 de [5] ■

Teorema 4.4. *Asumimos que el Algoritmo DF-QNB genera una sucesión infinita* $\{x_k\}$, *que consideramos M=1 en el Algoritmo de búsqueda lineal y que se cumplen las hipótesis del Lema 4.3. Si existe* k_0 *tal que* $\theta_k = \bar{\theta} < 1$ *para todo* $k \geq k_0$ *entonces existe un punto límite* x^* *de* $\{x_k\}_{k\in\mathbb{N}}$ *que es solución de* $F(x) = 0$.

Demostración. Notemos que $A_{k+1}s_k = y_k$ pues

$$\begin{aligned} A_{k+1}s_k &= \left(\int_0^1 J(x_k + ts_k)dt\right) s_k \\ &= \int_0^1 J(x_k + ts_k)s_k dt \\ &= F(x_k + ts_k)\Big|_0^1 \\ &= F(x_k + s_k) - F(x_k) \\ &= F(x_{k+1}) - F(x_k) = y_k. \end{aligned}$$

Usando que $A_{k+1}s_k = y_k$ podemos reescribir (9) de la siguiente manera:

$$B_{k+1} = B_k + \frac{(A_{k+1} - B_k)s_k s_k^t}{\|s_k\|^2},$$

pues

$$B_k + \frac{(A_{k+1} - B_k)s_k s_k^t}{\|s_k\|^2} = B_k + \frac{(A_{k+1}s_k - B_k s_k)s_k^t}{s_k^t s_k} = B_k + \frac{(y_k - B_k s_k)s_k^t}{s_k^t s_k} = B_{k+1}$$

Por el Lema 4.3, existe un conjunto $\widetilde{K} \subset \mathbb{N}$ tal que $\lim_{k\in\widetilde{K}} \frac{\|(A_{k+1} - B_k)d_k\|}{\|d_k\|} = 0.$

Sea x^* punto límite de $\{x_k\}_{k\in\widetilde{K}}$. Tenemos que

$$|\langle (B_k - J(x_k))\, d_k, F(x_k)\rangle| \leq [\|(A_{k+1} - B_k)\, d_k\| + \|(J(x_k) - A_{k+1})\, d_k\|]\, \|F(x_k)\|$$

pues usando la desigualdad de Hölder,

$$\begin{aligned} |\langle (B_k - J(x_k))\, d_k, F(x_k)\rangle| &\leq \|(B_k - J(x_k))\, d_k\|\, \|F(x_k)\| \\ &= \|B_k d_k - J(x_k)d_k\|\, \|F(x_k)\|\,, \end{aligned}$$

sumamos y restamos $A_{k+1}d_k$

$$\begin{aligned} |\langle (B_k - J(x_k))\, d_k, F(x_k)\rangle| &\leq \|B_k d_k - J(x_k)d_k + A_{k+1}d_k - A_{k+1}d_k\|\, \|F(x_k)\| \\ &= \|B_k d_k - A_{k+1}d_k + A_{k+1}d_k - J(x_k)d_k\|\, \|F(x_k)\| \\ &= \|(B_k - A_{k+1})d_k + (A_{k+1} - J(x_k))d_k\|\, \|F(x_k)\| \\ &= \|-(A_{k+1} - B_k)d_k - (J(x_k) - A_{k+1})d_k\|\, \|F(x_k)\| \\ &\leq [\|-(A_{k+1} - B_k)d_k\| + \|-(J(x_k) - A_{k+1})d_k\|]\, \|F(x_k)\| \\ &\leq [\|(A_{k+1} - B_k)d_k\| + \|(J(x_k) - A_{k+1})d_k\|]\, \|F(x_k)\|\,. \end{aligned}$$

Por Proposición 3.5, se sabe que $\|F(x_k)\|$ es acotada:

$$|\langle (B_k - J(x_k))d_k, F(x_k)\rangle| \leq C\,[\|(A_{k+1} - B_k)d_k\| + \|(J(x_k) - A_{k+1})d_k\|] \tag{22}$$

Notar que

$$\|(A_{k+1} - B_k)d_k\| \to 0,$$

pues $\lim_{k\in\widetilde{K}} \frac{\|(A_{k+1} - B_k)d_k\|}{\|d_k\|} = 0$ y $\{d_k\}$ es acotada.

Además, por definición de A_{k+1} y usando que $J(x)$ es Lipschitz continua tenemos que

$$\|(J(x_k) - A_{k+1})\, d_k\| \to 0,$$

pues

$$\begin{aligned}\|(J(x_k) - A_{k+1})\, d_k\| &= \left\| \left(J(x_k) - \int_0^1 J(x_k + ts_k)dt \right) d_k \right\| \\ &= \left\| \left(\int_0^1 J(x_k)\, dt - \int_0^1 J(x_k + ts_k)\, dt \right) d_k \right\| \\ &= \left\| \left(\int_0^1 (J(x_k) - J(x_k + ts_k))\, dt \right) d_k \right\|.\end{aligned}$$

Como $J(x)$ es Lipschitz continua tenemos que

$$\begin{aligned}\|(J(x_k) - A_{k+1})\, d_k\| &\leq \left(\int_0^1 \|J(x_k) - J(x_k) + ts_k\|\, dt \right) \|d_k\| \\ &\leq \left(\int_0^1 L\, \|x_k - x_k - ts_k\|\, dt \right) \|d_k\| \\ &= \left(\int_0^1 Lt\, \|s_k\|\, dt \right) \|d_k\| \\ &= L\, \|s_k\| \int_0^1 t\, dt\, \|d_k\| = L\, \|s_k\| \frac{t^2}{2}\Big|_0^1 \|d_k\| \\ &= \frac{L}{2}\, \|s_k\|\, \|d_k\| \leq \frac{L}{2}\, \|s_k\|\, \Delta.\end{aligned}$$

Por la Proposición (3.5) sabemos que la serie $\sum_{k=1}^{\infty} \|s_k\|^2$ es convergente. Por criterio de series convergentes, $\lim_{k\to\infty} \|s_k\|^2 = 0$, entonces $\lim_{k\to\infty} \|s_k\| = 0$. Con lo cual,

$$0 \leq \lim_{k\to\infty} \|(J(x_k) - A_{k+1})\, d_k\| \leq \lim_{k\to\infty} \frac{L}{2}\Delta\, \|s_k\| = \frac{L}{2}\Delta \lim_{k\to\infty} \|s_k\| = 0.$$

Por lo tanto $\lim_{k\to\infty} \|(J(x_k) - A_{k+1})\, d_k\| = 0$.

Tomamos límite para $k \in \widetilde{K}$ en (22) obtenemos que

$$\begin{aligned}0 &\leq \lim_{k\in\widetilde{K}} |\langle (B_k - J(x_k))\, d_k, F(x_k)\rangle| \\ &\leq \lim_{k\in\widetilde{K}} C\, [\|(A_{k+1} - B_k)\, d_k\| + \|(J(x_k) - A_{k+1})\, d_k\|] \\ &= C \left[\lim_{k\in\widetilde{K}} \|(A_{k+1} - B_k) d_k\| + \lim_{k\in\widetilde{K}} \|(J(x_k) - A_{k+1}) d_k\| \right] = 0\end{aligned}$$

Con lo cual, $\lim_{k\in\widetilde{K}} \langle (B_k - J(x_k))\, d_k, F(x_k)\rangle = 0$.

Siguiendo la demostración del Teorema 4.1 y que $\lim_{k\in\widetilde{K}} \langle (B_k - J(x_k)) d_k, F(x_k)\rangle = 0$, se puede demostrar que $F(x^*) = 0$.

Por lo tanto, existe un punto límite de $\{x_k\}_{k\in\mathbb{N}}$ que es solución de $F(x) = 0$. ■

Teorema 4.5. *Supongamos que el Algoritmo DF-QNB genera una sucesión infinita* $\{x_k\}_{k\in\mathbb{N}}$, *que consideramos* $M = 1$ *en el Algoritmo de búsqueda lineal, que el conjunto* $\Omega = \{x \in \mathbb{R}^n \text{ tal que } f(x) \leq f(x_0) + 2\eta\}$ *es acotado y que existe* $k_0 \in \mathbb{N}$ *tal que* $\theta_k = \bar{\theta} < 1$ *para todo* $k \geq k_0$. *Entonces,*

$$\lim_{k\to\infty} F(x_k) = 0.$$

Es decir, todo punto límite de la sucesión $\{x_k\}_{k\in\mathbb{N}}$ *es solución de* $F(x) = 0$.

Demostración. Como se cumplen las hipótesis del Teorema (4.4), existe un punto límite x^* de $\{x_k\}_{k\in\mathbb{N}}$ tal que $F(x^*) = 0$.

Sea K_1 subconjunto infinito de $\mathbb{N}$ tal que $\lim_{k\in K_1} x_k = x^*$.

Entonces por continuidad de F

$$\lim_{k\in K_1} F(x_k) = F(x^*) = 0.$$

Para todo $k \in \mathbb{N}$ consideramos

$$U_k = \text{máx}\{f(x_{(k-1)M+1}), \ldots, f(x_{kM})\}$$

y definimos

$$\nu(k) \in \{(k-1)M + 1, \ldots, kM\}$$

el índice tal que $f(x_{\nu(k)}) = U_k$.

Como consideramos $M = 1$ en el Algoritmo de búsqueda lineal, tenemos que:

$$U_k = f(x_k) = \frac{1}{2}\|F(x_k)\|^2$$

$$\text{y } \nu(k) = k,$$

con lo cual $\lim_{k\in K_1} U_k = 0$ pues vimos que $\lim_{k\in K_1} F(x_k) = 0$.

Sea $K_1 = \{k_1, k_2, k_3, \ldots\}$ tal que $k_i < k_j$ si $i < j$ y $\lim_{j\to\infty} k_j = \infty$.

Tomemos k_j un elemento fijo del conjunto K_1. Para todo $i \in \mathbb{N}$, $i > k_j$ tenemos que

$$f(x_i) \leq f(x_{k_j}) + \eta_{k_j} + \eta_{k_j+1} + \ldots + \eta_{i-1}$$

pues tenemos que:

$$f(x_{\nu(k+1)}) \leq f(x_{\nu(k)}) + \sum_{j=0}^{M-1} \eta_{kM+j} - \gamma\alpha^2_{\nu(k+1)-1}\left\|d_{\nu(k+1)-1}\right\|^2$$

como consideramos $M = 1$ en el Algoritmo de búsqueda lineal,

$$f(x_{k+1}) \leq f(x_k) + \eta_k - \gamma\alpha_k^2 \|d_k\|^2 .$$

Para $k = k_j$,

$$f(x_{k_j+1}) \leq f(x_{k_j}) + \eta_{k_j} - \gamma\alpha_{k_j}^2 \left\|d_{k_j}\right\|^2$$
$$f(x_{k_j+1}) \leq f(x_{k_j}) + \eta_{k_j}$$

Para $k = k_j + 1$,

$$f(x_{k_j+2}) \leq f(x_{k_j+1}) + \eta_{k_j+1} - \gamma\alpha_{k_j+1}^2 \left\|d_{k_j+1}\right\|^2$$
$$f(x_{k_j+2}) \leq f(x_{k_j+1}) + \eta_{k_j+1}$$

y como $f(x_{k_j+1}) \leq f(x_{k_j}) + \eta_{k_j}$ tenemos que

$$f(x_{k_j+2}) \leq f(x_{k_j}) + \eta_{k_j} + \eta_{k_j+1}$$

En general,

$$f(x_i) \leq f(x_{k_j}) + \eta_{k_j} + \eta_{k_j+1} + \ldots + \eta_{i-1}$$

para todo $i \in \mathbb{N}$, $i > k_j$, entonces

$$f(x_i) \leq f(x_{k_j}) + \sum_{l=k_j}^{i-1} \eta_l.$$

Por lo tanto, como $U_k = f(x_k)$ tenemos que

$$U_i \leq U_{k_j} + \sum_{l=k_j}^{i-1} \eta_l.$$

Por lo tanto

$$U_i \leq U_{k_j} + \sum_{l=k_j}^{\infty} \eta_l. \tag{23}$$

Tomando supremo en (23) obtenemos:

$$\sup_{i>k_j} U_i \leq U_{k_j} + \sum_{l=k_j}^{\infty} \eta_l \tag{24}$$

pues el lado derecho de (23) no depende de i.

Tomando límite cuando $j \to \infty$ en (24) tenemos que:

$$\lim_{j\to\infty} \sup_{i>k_j} U_i \leq \lim_{j\to\infty} U_{k_j} + \lim_{j\to\infty} \sum_{l=k_j}^{\infty} \eta_l = 0,$$

pues $\lim_{j\to\infty} U_{k_j} = \lim_{k\in K_1} U_k = 0$ y $\lim_{j\to\infty} \sum_{l=k_j}^{\infty} \eta_l = 0$ ya que la serie $\sum_{k=1}^{\infty} \eta_k$ es convergente.

Por otro lado,

$$U_i = f(x_i) = \frac{1}{2} \|F(x_i)\|^2 \geq 0$$

entonces,

$$0 \leq \lim_{j\to\infty} \sup_{i>k_j} U_i \leq 0$$

con lo cual,

$$\lim_{j\to\infty} \sup_{i>k_j} U_i = 0.$$

Es decir,

$$\lim_{i\to\infty} U_i = 0.$$

Por lo tanto,

$$\lim_{k\to\infty} F(x_k) = 0.$$

Con esto, podemos concluir que todo punto límite de $\{x_k\}_{k\in\mathbb{N}}$ es solución de $F(x) = 0$. ■

4.3. Análisis de convergencia local

En esta sección analizaremos que se puede demostrar un resultado del tipo de convergencia local análogo a teoremas probados en [4] y en [7].

Teorema 4.6. *Supongamos que el Algoritmo DF-QNB genera una sucesión infinita* $\{x_k\}_{k\in\mathbb{N}}$ *que tiene a* x^* *como punto de acumulación. Consideremos* $M = 1$ *en el Algoritmo de búsqueda lineal y que* $\Omega = \{x \in \mathbb{R}^n \text{ tal que } f(x) \leq f(x_0) + 2\eta\}$ *es acotado. Si existe* $\delta > 0$ *tal que* x^* *es la única solución del problema* $F(x) = 0$ *en la bola* $\|x - x^*\| \leq \delta$ *y si existe* $k_0 \in \mathbb{N}$ *tal que* $\theta_k = \bar{\theta} < 1$ *para todo* $k \geq k_0$ *entonces* $\{x_k\}_{k\in\mathbb{N}}$ *converge a* x^*.

Demostración. Tomamos K_1 subconjunto infinito de $\mathbb{N}$ tal que

$$\lim_{k\in K_1} x_k = x^*$$

entonces $\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0$ pues:

Como $x_{k+1} = x_k + \alpha_k d_k$, tenemos que

$$\|x_{k+1} - x_k\|^2 = \|\alpha_k d_k\|^2 = \alpha_k^2 \|d_k\|^2,$$

por la Proposición 3.5 la serie $\sum_{k=1}^{\infty} \alpha_k^2 \|d_k\|^2$ es convergente, entonces $\lim_{k\to\infty} \|x_{k+1} - x_k\|^2 = 0$, con lo cual,

$$\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0,$$

entonces existe $k_1 \in \mathbb{N}$ tal que $\|x_{k+1} - x_k\| \leq \frac{\delta}{2}$ para todo $k \geq k_1$, $k \in \mathbb{N}$.

Sea $S = \{x \in \mathbb{R}^n \ tal \ que \ \frac{\delta}{2} \leq \|x - x^*\| \leq \delta\}$.

- S no contiene ningún punto límite de $\{x_k\}_{k\in\mathbb{N}}$ pues, si suponemos que $\bar{x}$ es punto límite de $\{x_k\}_{k\in\mathbb{N}}$ tal que $\bar{x} \in S$, por Teorema 4.5 todo punto límite de $\{x_k\}_{k\in\mathbb{N}}$ es solución de $F(x) = 0$, con lo cual $F(\bar{x}) = 0$. Obtendriamos un absurdo, pues por hipótesis, x^* es la única solución al problema $F(x) = 0$ en la bola $\|x - x^*\| \leq \delta$.

- S no contiene un número infinito de iteraciones x_k pues, si suponemos que S contiene infinitas iteraciones x_k, como S es compacto, $\{x_k\}_{k\in\mathbb{N}}$ admite punto de acumulación que lo llamamos $\widehat{x}$, por Teorema 4.5 todo punto límite de $\{x_k\}_{k\in\mathbb{N}}$ es solución de $F(x) = 0$, con lo cual $F(\widehat{x}) = 0$ con lo que obtenemos nuevamente un absurdo.
Por lo tanto, como S no contiene un número infinito de iteraciones x_k, existe $k_2 \in \mathbb{N}$ tal que $x_k \notin S$ para todo $k \geq k_2$.

Sea $k_3 \geq max\{k_1, k_2\}$; $k_3 \in K_1$ tal que $\|x_k - x^*\| \leq \frac{\delta}{2}$ para todo $k \in K_1, k \geq k_3$, ya que $\lim_{k\in K_1} x_k = x^*$.

Además,

$$\|x_{k_3+1} - x^*\| = \|x_{k_3+1} - x_{k_3} + x_{k_3} - x^*\| \leq \|x_{k_3+1} - x_{k_3}\| + \|x_{k_3} - x^*\| \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

pues teníamos que $\lim_{k\in K_1} x_k = x^*$ y $\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0$.

Con esto podemos afirmar que $x_{k_3+1} \notin S$ ($x_k \notin S$ para todo $k \geq k_2$ y $k_3 + 1 > k_3 \geq k_2$).

Por lo tanto, x_{k_3+1} está en la bola de radio menor o igual a $\frac{\delta}{2}$ ($\|x_{k_3+1} - x^*\| \leq \delta$ y $x_{k_3+1} \notin S$).

Haciendo un paso más obtenemos que:

$$\|x_{k_3+2} - x^*\| = \|x_{k_3+2} - x_{k_3+1} + x_{k_3+1} - x^*\| \leq \|x_{k_3+2} - x_{k_3+1}\| + \|x_{k_3+1} - x^*\| \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

pues $\|x_{k+1} - x_k\| \leq \frac{\delta}{2}$ para todo $k \geq k_1$ y $k_3 + 2 \geq k_3 \geq k_1$ y $\|x_{k_3+1} - x^*\| \leq \frac{\delta}{2}$ vale por el análisis realizado en el paso anterior.

Por lo tanto, x_{k_3+2} está en la bola de radio menor o igual a $\frac{\delta}{2}$ pues $\|x_{k_3+2} - x^*\| \leq \delta$ y $x_{k_3+2} \notin S$.

Siguiendo el proceso inductivo obtenemos que

$$\|x_k - x^*\| \leq \frac{\delta}{2} \text{ para todo } k \geq k_3, \ k \in \mathbb{N}.$$

entonces, todos los puntos de acumulación $\widetilde{x}$ de $\{x_k\}_{k\in\mathbb{N}}$, cumplen que $\|\widetilde{x} - x^*\| \leq \frac{\delta}{2}$. Por hipótesis el conjunto definido por $0 < \|x - x^*\| \leq \frac{\delta}{2}$ no contiene soluciones de $F(x) = 0$.

Por lo tanto $\{x \in \mathbb{R}^n : 0 < \|x - x^*\| \leq \frac{\delta}{2}\}$ no contiene puntos límites. Si los tuviera, por el Teorema (4.5) serían solución del sistema $F(x) = 0$ contradiciendo la hipótesis de este teorema.

Con lo cual, el único punto límite en la bola definida por $\{x \in \mathbb{R}^n : \|x - x^*\| \leq \frac{\delta}{2}\}$ es x^*.

Por lo tanto, podemos afirmar que toda la sucesión $\{x_k\}_{k\in\mathbb{N}}$ converge a x^*. ■

5. Conclusiones y trabajo futuro

En este trabajo se estudiaron los resultados de convergencia global para el método DF-QNB presentados en [3].

Se pudo demostrar un teorema que completa el estudio de convergencia global. El resultado obtenido nos permitió demostrar un teorema de convergencia local, análogo a un resultado presentado en [4] y en [7] para el método DF-QNB cuando tomamos $M = 1$ en el Algoritmo de búsqueda lineal.

Como continuidad del estudio actual, basados en trabajos sobre convergencia local con búsqueda no monótona, se analizará la posibilidad de demostrar un teorema de convergencia local cuando $M > 1$ en el Algoritmo de búsqueda lineal. Así como también se incluirá el análisis de la velocidad de convergencia del método DF-QNB, previo estudio y análisis de los resultados presentados en [1] y en [7].

Referencias

[1] E. G. Birgin, N. Krejic, J. M. Martinez. *Globally convergent inexact quasi-Newton methods for solving nonlinear systems.* Numerical Algorithms, **32** (2003), 249-260.

[2] J. E. Dennis jr, R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Englewood Cliffs, 1983.

[3] N. Echebest, M. L. Schuverdt, R. P. Vignau. *Two derivative-free methods for solving underdetermined nonlinear systems of equations.* Computational & Applied Mathematics, **30** (2010), 217-245.

[4] W. La Cruz, J. M. Martinez, M. Raydan. *Spectral residual method without gradient information for solving large-scale nonlinear systems of equation: Theory and experiments.* Mathematics of Computation, **75** (2006), 1429-1448.

[5] D. H. Li, M. Fukushima. *A derivative-free line search and global convergence of Broyden-like method for nonlinear equations.* Opt. Meth. and Software, **13** (2000), 181-201.

[6] D. E. Luenberger. *Programación lineal y no lineal.* Addison-Wesley Iberoamericana, 1989.

[7] J. M. Martinez. *Quasi-Inexact-Newton methods with global convergence for solving constrained nonlinear systems.* Nonlinear Analysis, **30** (1997), 1-7.

[8] J. M. Martinez, S. A. Santos. *Métodos computacionais de optimizaçao, 20º Colóquio Brasileiro de Matemática.* IMPA, 24-28 Julho 1995.

[9] J. Nocedal, S. Wright. *Numerical Optimization.* Springer, 1999.