

# Estudio numérico del método de Newton con búsqueda lineal no monótona

Leidy Alexandra Florez Ñañez

Universidad del Cauca  
Facultad de Ciencias Naturales, Exactas y de la Educación  
Programa de Matemáticas  
Departamento de Matemáticas  
Popayán  
2018

# Estudio numérico del método de Newton con búsqueda lineal no monótona

Estudiante:

Leidy Alexandra Florez Ñañez

Director

Favián Enrique Arenas Aparicio

Universidad del Cauca  
Facultad de Ciencias Naturales, Exactas y de la Educación  
Programa de Matemáticas  
Departamento de Matemáticas  
Popayán  
2018

## Nota de aceptación

---

---

---

---

Director \_\_\_\_\_

**Mg. Favián Arenas Aparicio**

Jurado \_\_\_\_\_

**Mg. Ana Milena Plaza Buitrago**

Jurado \_\_\_\_\_

**Mg. Hevert Vivas**

**Fecha de sustentación: 11 de septiembre de 2018, Popayán.**

# Índice general

<b>Índice general</b>	<b>4</b>
<b>1. Introducción</b>	<b>5</b>
<b>2. Método de búsqueda lineal no monótona</b>	<b>8</b>
2.1. Teorema de convergencia . . . . .	9
2.1.1. Hipótesis de convergencia . . . . .	9
2.1.2. Lemas importantes . . . . .	10
<b>3. Comparación Numérica</b>	<b>20</b>
<b>4. Comentarios Finales</b>	<b>29</b>
<b>Bibliografía</b>	<b>30</b>

# Capítulo 1

## Introducción

La optimización puede definirse como la ciencia de determinar las “mejores” soluciones a ciertos problemas matemáticamente definidos, que a menudo son modelos de la realidad física. Implica el estudio de criterios de optimalidad para problemas, la determinación de métodos algorítmicos de solución, el estudio de la estructura de tales métodos y la experimentación por ordenador, tanto en condiciones experimentales como en problemas de la vida real [4].

Uno de los problemas fundamentales en la optimización es el de desarrollar procedimientos computacionales para encontrar puntos extremos de una función  $f$ , que llamaremos función objetivo [6], este problema se puede ver de la siguiente manera:

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}), \quad (1.1)$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es continuamente diferenciable.

En la vida real, existen situaciones que involucran problemas de optimización, algunos de minimización, otros de maximización<sup>1</sup>, como lo son: diseño de redes de transporte y rutas de envío y asignación de personal a estaciones de trabajo, entre otras. Por ésta razón es de mucha importancia conocer métodos que permitan la resolución del problema (1.1).

La aplicabilidad de los métodos de optimización es generalizada, llegando a casi todas

---

<sup>1</sup>Siempre se puede convertir un problema de máx  $f(x)$  en mín  $-f(x)$ , por eso se acostumbra a tratar solo uno de éstos problemas.

las actividades en las que se procesa la información numérica (Ciencia, Ingeniería, Matemáticas, Economía, Comercio, etc.). Proporcionar una descripción completa de todas estas aplicaciones, sería poco práctico, pero una selección de estas aplicaciones podría incluir [4]:

- Diseño de reactores químicos.
- Diseño de un motor o de un aerógrafo
- Diseño estructural: edificios, puentes, etc.
- Asignación de recursos comerciales, programación, mezcla.

Y aplicaciones a otras ramas del análisis numérico:

- Ajuste de datos.
- Principios de variación en ecuaciones diferenciales parciales.
- Ecuaciones no lineales en ecuaciones diferenciales ordinarias.
- Las funciones penalizadas.

Entre los métodos más utilizados en la literatura, para resolver éste tipo de problemas están: el *método de Newton*, que en su forma pura presenta inconvenientes, por lo que se han propuesto varias modificaciones para asegurar la convergencia global hacia mínimos locales [5], entre éstas está el *método de Newton con búsqueda lineal* el cual utiliza la búsqueda lineal monótona, como estrategia de globalización.

En 1986, *Grippio*, *Lampariello* y *Lucidi* presentaron una nueva estrategia de globalización: la búsqueda lineal no monótona para el método de Newton, la cual se quiere estudiar en éste trabajo, en esencia, consiste en elegir el tamaño del paso de la búsqueda lineal, forzando a que el descenso sea suficiente, comparado con varios puntos ya elegidos. Ésta técnica puede ser vista como una generalización de la regla de *Armijo* [2], en el sentido de que también garantiza que el descenso sea suficiente, sin afectar las propiedades de convergencia [5].

El documento está organizado de la siguiente forma: En el **Capítulo 2**, se presenta la respectiva diferencia que se tiene al escoger el *tamaño de paso* adecuado en cada iteración de la *búsqueda lineal monótona* y la *búsqueda lineal no monótona*, lo que conlleva

a realizar detalladamente el teorema de convergencia para el *método de búsqueda lineal no monótona* para problemas de minimización (1.1) y comprender mejor el algoritmo propuesto. En el **Capítulo 3**, se presentan los resultados numéricos obtenidos al implementar el algoritmo propuesto en [5] y se realiza las debidas comparaciones que se obtuvieron. En el **Capítulo 4**, se hacen algunos comentarios finales y una propuesta para futuros trabajos en el tema.

# Capítulo 2

## Método de búsqueda lineal no monótona

En los métodos de minimización locales, se presenta un procedimiento llamado búsqueda direccional; que consiste en buscar una dirección de descenso  $\mathbf{d}$ , mediante métodos como: el método del gradiente, el método de Newton, métodos de cuasi Newton, entre otros.

Para convertirlos en métodos globales son muy utilizadas las estrategias de: *búsqueda lineal* y *región de confianza* [3]. En particular para la primera, se toma una nueva variable  $\lambda_k$  que se conoce como *tamaño de paso*, la cual indica, a partir de la dirección de descenso antes encontrada, que tan grande o pequeño es ese paso. El tamaño del paso inicial es  $\lambda_0 = 1$  y se va reduciendo a la mitad hasta cuando se cumpla la siguiente condición, llamada *condición de Armijo*:

$$f(\mathbf{x}_k + \lambda_k \mathbf{d}_k) \leq f(\mathbf{x}_k) + \alpha \lambda_k \nabla f(\mathbf{x}_k)^T \mathbf{d}_k. \quad (2.1)$$

El método de *búsqueda lineal no monótona* también es una estrategia de globalización para minimización; de ésta puede decirse que es una generalización de la estrategia de búsqueda lineal. En lugar de la *condición de Armijo* éste método utiliza la siguiente condición:

$$f(\mathbf{x}_k + \lambda_k \mathbf{d}_k) \leq \max_{0 \leq j \leq m(k)} [f(\mathbf{x}_{k-j})] + \alpha \lambda_k \nabla f(\mathbf{x}_k)^T \mathbf{d}_k, \quad (2.2)$$

donde  $m(0) = 0$  y  $0 \leq m(k) \leq \min[m(k-1) + 1, M]$ ,  $k \geq 1$ .

Observemos que la diferencia entre la *condición de Armijo* y la *búsqueda lineal no monótona*, es que la condición de Armijo compara en cada iteración solamente con el valor que toma la función en la anterior iteración, es decir  $f(\mathbf{x}_k)$ , en cambio la búsqueda lineal no monótona, compara con el máximo de los valores que toma la función en iteraciones anteriores, las cuales van a depender del parámetro fijo  $M$ , ésto garantiza que el decrecimiento de la función sea notable y ayuda a que la convergencia sea más rápida.

## 2.1. Teorema de convergencia

En el teorema dado por *Grippo, Lampariello y Lucidi* en [5], el cual se va a mencionar en el **Teorema 2.1**, establece unas hipótesis que van a ayudar a la demostración de algunos lemas importantes, los cuales serán de mucha utilidad a la hora de demostrar la teoría de convergencia. Estas hipótesis están dadas por.

### 2.1.1. Hipótesis de convergencia

Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función continuamente diferenciable y sea  $\{\mathbf{x}_k\}$  una sucesión definida por:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k, \quad \mathbf{d}_k \neq 0$$

**H1.**  $\Omega_0 = \{\mathbf{x} : f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$  es compacto.

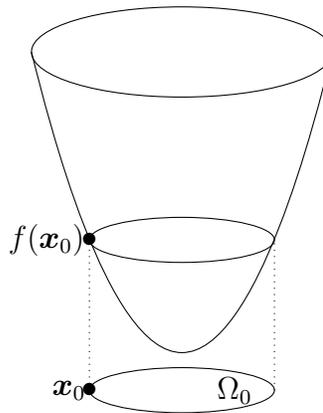


Figura 2.1: Conjunto de nivel inferior  $\Omega_0$

**H2.** Existen números reales positivos  $c_1, c_2$  tales que:

$$\begin{aligned}\nabla f(\mathbf{x}_k)^T \mathbf{d}_k &\leq -c_1 \|\nabla f(\mathbf{x}_k)\|^2, \\ \|\mathbf{d}_k\| &\leq c_2 \|\nabla f(\mathbf{x}_k)\|.\end{aligned}$$

**H3.** Sea  $a > 0$ ,  $\sigma \in (0, 1)$ ,  $\alpha \in (0, 1)$  y sea  $M$  un entero no negativo, se elige  $\lambda_k = \sigma^{h_k} a$ , donde  $h_k$  es el primer entero no negativo para el cual  $\lambda_k$  satisface:

$$f(\mathbf{x}_k + \lambda_k \mathbf{d}_k) \leq \max_{0 \leq j \leq m(k)} [f(\mathbf{x}_{k-j})] + \alpha \lambda_k \nabla f(\mathbf{x}_k)^T \mathbf{d}_k, \quad (2.3)$$

donde  $m(0) = 0$  y  $0 \leq m(k) \leq \min[m(k-1) + 1, M]$ ,  $k \geq 1$ .

Observemos que:  $m(k) \leq m(k-1) + 1$ , es decir que  $m(k)$  puede ser más pequeño que  $m(k-1)$ , o más grande a lo sumo en una unidad.

De aquí en adelante se utilizarán las siguientes notaciones:

1. Sea  $\ell(k)$  un entero tal que:

$$k - m(k) \leq \ell(k) \leq k \quad \text{se cumple} \quad (2.4)$$

$$f(\mathbf{x}_{\ell(k)}) := \max_{0 \leq j \leq m(k)} \{f(\mathbf{x}_{k-j})\}.$$

2.  $\nabla f_k := \nabla f(\mathbf{x}_k)$ .

### 2.1.2. Lemas importantes

A continuación se presentan algunos lemas importantes que van a ayudar en la demostración del **Teorema 2.1**.

**Lema 2.1.** *Supongamos la hipótesis **H3**, entonces la subsucesión  $\{f(\mathbf{x}_{\ell(k)})\}$  es decreciente.*

*Demostración.* La expresión (2.4) permite ver que  $f(\mathbf{x}_{\ell(k)})$  es el valor máximo de las iteraciones elegidas y  $\ell(k)$  es la iteración donde se da ese valor máximo.

Como  $m(k+1) \leq m(k) + 1$ , a partir de aquí se puede escribir:

$$\begin{aligned} f(\mathbf{x}_{\ell(k+1)}) &= \max_{0 \leq j \leq m(k+1)} \{f(\mathbf{x}_{k+1-j})\} \quad \text{por (2.4)} \\ &\leq \max_{0 \leq j \leq m(k)+1} \{f(\mathbf{x}_{k+1-j})\} \\ &= \max \{f(\mathbf{x}_{\ell(k)}), f(\mathbf{x}_{k+1})\}. \end{aligned}$$

De la hipótesis **H3**:

$$f(\mathbf{x}_{k+1}) \leq \max_{0 \leq j \leq m(k)} \{f(\mathbf{x}_{k-j})\} + \alpha \lambda_k \nabla f_k^T \mathbf{d}_k \leq \max_{0 \leq j \leq m(k)} \{f(\mathbf{x}_{k-j})\} = f(\mathbf{x}_{\ell(k)})$$

Luego el  $\max \{f(\mathbf{x}_{\ell(k)}), f(\mathbf{x}_{k+1})\} = f(\mathbf{x}_{\ell(k)})$ , es decir que  $f(\mathbf{x}_{\ell(k+1)}) \leq f(\mathbf{x}_{\ell(k)})$ .

Por lo tanto, la subsucesión  $\{f(\mathbf{x}_{\ell(k)})\}$  es decreciente.

◆

**Lema 2.2.** *Supongamos que se cumple la hipótesis **H3**, entonces la sucesión  $\{\mathbf{x}_k\}$  permanece en  $\Omega_0$ .*

*Demostración.* Por la hipótesis **H3** se tiene que:

$$\begin{aligned} f(\mathbf{x}_1) &\leq f(\mathbf{x}_0) + \alpha \lambda_0 \nabla f_0^T \mathbf{d}_0 \leq f(\mathbf{x}_0) \\ f(\mathbf{x}_2) &\leq \max_{0 \leq j \leq m(1)} \{f(\mathbf{x}_{1-j})\} + \alpha \lambda_1 \nabla f_1^T \mathbf{d}_1 \leq f(\mathbf{x}_0) \\ f(\mathbf{x}_3) &\leq \max_{0 \leq j \leq m(2)} \{f(\mathbf{x}_{2-j})\} + \alpha \lambda_2 \nabla f_2^T \mathbf{d}_2 \leq f(\mathbf{x}_0) \\ &\dots \\ f(\mathbf{x}_{k+1}) &\leq \max_{0 \leq j \leq m(k)} \{f(\mathbf{x}_{k-j})\} + \alpha \lambda_k \nabla f_k^T \mathbf{d}_k \leq f(\mathbf{x}_0) \end{aligned}$$

De aquí que para todo  $k \in \mathbb{N}$  se va a cumplir que  $f(\mathbf{x}_k) \leq f(\mathbf{x}_0)$ . En otras palabras, la sucesión  $\{\mathbf{x}_k\}$  permanece en  $\Omega_0$ .

◆

**Lema 2.3.** *La subsucesión  $\{f(\mathbf{x}_{\ell(k)})\}$  es convergente.*

*Demostración.* Por lo probado en el **Lema 2.2**,  $\{\mathbf{x}_k\}$  permanece en  $\Omega_0$ , es decir, está acotada y como la función  $f$  es continua, entonces por el teorema del valor intermedio [1],

$f$  alcanza su máximo y mínimo, de aquí que  $\{f(\mathbf{x}_k)\}$  es acotada y como  $\{f(\mathbf{x}_{\ell(k)})\}$  está contenida en ella también va a ser acotada, por otro lado, por lo probado en el **Lema 2.1**, la subsucesión es monótona, entonces por el Teorema de Bolzano-Weierstrass [1], la subsucesión es convergente.

◆

**Lema 2.4.** *Si se cumplen las hipótesis **H1** a **H3**, entonces  $\lim_{k \rightarrow \infty} \lambda_{\ell(k)-1} \|\mathbf{d}_{\ell(k)-1}\| = 0$ .*

*Demostración.* Como por el anterior lema,  $\{f(\mathbf{x}_{\ell(k)})\}$  es convergente, entonces supongamos que  $\lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}) = L$ . Además, de (2.3) se obtiene para  $k > M$ :

$$\begin{aligned} f(\mathbf{x}_{\ell(k)}) &= f(\mathbf{x}_{(\ell(k)-1)+1}) \\ &= f(\mathbf{x}_{\ell(k)-1} + \lambda_{\ell(k)-1} \mathbf{d}_{\ell(k)-1}) \\ &\leq \max_{0 \leq j \leq m(\ell(k)-1)} f(\mathbf{x}_{\ell(k)-1-j}) + \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \\ &= f(\mathbf{x}_{\ell(k)-1}) + \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1}, \end{aligned} \quad (2.5)$$

la última igualdad se tiene reemplazando lo definido en (2.4).

Tomando límite cuando  $k \rightarrow \infty$  a ambos lados se tiene:

$$\begin{aligned} \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}) &\leq \lim_{k \rightarrow \infty} \left( \max_{0 \leq j \leq m(\ell(k)-1)} f(\mathbf{x}_{\ell(k)-1-j}) + \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \right) \\ L &\leq \lim_{k \rightarrow \infty} \left( \max_{0 \leq j \leq m(\ell(k)-1)} f(\mathbf{x}_{\ell(k)-1-j}) \right) + \lim_{k \rightarrow \infty} \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \\ L &\leq \max_{0 \leq j \leq m(\ell(k)-1)} \left[ \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)-1-j}) \right] + \lim_{k \rightarrow \infty} \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \\ L &\leq L + \lim_{k \rightarrow \infty} \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \\ L - L &\leq \lim_{k \rightarrow \infty} \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \\ 0 &\leq \lim_{k \rightarrow \infty} \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1}, \end{aligned}$$

pero como  $\lambda_k > 0$  y  $\nabla f_k^T \mathbf{d}_k \leq 0$  por ser una dirección de descenso entonces

$$0 \leq \lim_{k \rightarrow \infty} \alpha \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \leq 0,$$

por el teorema de comparación se deduce que:

$$\lim_{k \rightarrow \infty} \lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} = 0. \quad (2.6)$$

De la hipótesis **H2** tenemos que:

$$\lambda_k \nabla f_k^T \mathbf{d}_k \leq -c_1 \lambda_k \|\nabla f_k\|^2, \quad (2.7)$$

y

$$\begin{aligned} \|\mathbf{d}_k\| \leq c_2 \|\nabla f_k\| &\implies \|\mathbf{d}_k\|^2 \leq c_2^2 \|\nabla f_k\|^2 \\ &\implies \frac{\|\mathbf{d}_k\|^2}{c_2^2} \leq \|\nabla f_k\|^2 \\ &\implies \lambda_k \frac{\|\mathbf{d}_k\|^2}{c_2^2} \leq \lambda_k \|\nabla f_k\|^2. \end{aligned}$$

Como  $c_1 > 0$  entonces  $-c_1 \leq 0$ , multiplicando la anterior desigualdad por  $-c_1$  se tiene:

$$\frac{-c_1}{c_2^2} \lambda_k \|\mathbf{d}_k\|^2 \geq -c_1 \lambda_k \|\nabla f_k\|^2,$$

luego en (2.7) se va a tener:

$$\lambda_k \nabla f_k^T \mathbf{d}_k \leq -c_1 \lambda_k \|\nabla f_k\|^2 \leq \frac{-c_1}{c_2^2} \lambda_k \|\mathbf{d}_k\|^2.$$

Como ésta desigualdad se cumple para todo  $k$ , en particular se va cumplir para  $k = \ell(k) - 1$ , es decir que:

$$\lambda_{\ell(k)-1} \nabla f_{\ell(k)-1}^T \mathbf{d}_{\ell(k)-1} \leq \frac{-c_1}{c_2^2} \lambda_{\ell(k)-1} \|\mathbf{d}_{\ell(k)-1}\|^2,$$

pero sabemos que  $c_1 > 0$ ,  $c_2^2 > 0$ ,  $\lambda_{\ell(k)-1} > 0$  y  $\|\mathbf{d}_{\ell(k)-1}\|^2 \geq 0$  entonces

$$\frac{-c_1}{c_2^2} \lambda_{\ell(k)-1} \|\mathbf{d}_{\ell(k)-1}\|^2 \leq 0,$$

tomando limite cuando  $k \rightarrow \infty$  a ambos lados de la desigualdad y por (2.6) se va a tener:

$$0 \leq \frac{-c_1}{c_2^2} \lim_{k \rightarrow \infty} \lambda_{\ell(k)-1} \|\mathbf{d}_{\ell(k)-1}\|^2 \leq 0,$$

luego, por el teorema de comparación (o sándwich) se deduce que:

$$\lim_{k \rightarrow \infty} \lambda_{\ell(k)-1} \|\mathbf{d}_{\ell(k)-1}\|^2 = 0,$$

ahora como  $\lambda_k < a$ , es decir acotada, entonces también se va a cumplir:

$$\lim_{k \rightarrow \infty} \lambda_{\ell(k)-1} \|\mathbf{d}_{\ell(k)-1}\| = 0.$$

◆

**Lema 2.5.** *Si se cumplen las hipótesis **H1** a **H3**, entonces  $\lim_{k \rightarrow \infty} \lambda_k \|\mathbf{d}_k\| = 0$ .*

*Demostración.* Sea  $\hat{\ell}(k) = \ell(k+M+2)$ , primero demostremos por inducción matemática, para cualquier  $j \geq 1$  se mantiene

$$\lim_{k \rightarrow \infty} \lambda_{\hat{\ell}(k)-j} \|\mathbf{d}_{\hat{\ell}(k)-j}\| = 0 \quad (2.8)$$

y

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)-j}) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}) \quad (2.9)$$

Para  $j = 1$ , por la escogencia de  $\hat{\ell}(k)$  se tiene que  $\{\hat{\ell}(k)\} \subset \{\ell(k)\}$ , de aquí se demuestra (2.8) por el lema anterior. Además por el **Lema 2.1** se demuestra que:

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)}) \text{ existe y } \lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)}) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}).$$

Ahora,

$$\begin{aligned} \mathbf{x}_{\hat{\ell}(k)} &= \mathbf{x}_{\hat{\ell}(k)-1} + \lambda_{\hat{\ell}(k)-1} \mathbf{d}_{\hat{\ell}(k)-1} \\ \mathbf{x}_{\hat{\ell}(k)} - \mathbf{x}_{\hat{\ell}(k)-1} &= \lambda_{\hat{\ell}(k)-1} \mathbf{d}_{\hat{\ell}(k)-1}, \end{aligned}$$

por lema anterior  $\|\mathbf{x}_{\hat{\ell}(k)} - \mathbf{x}_{\hat{\ell}(k)-1}\| \rightarrow 0$  cuando  $k \rightarrow \infty$ , y puesto que  $f$  es uniformemente continua en  $\Omega_0$  se tiene que:

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)-1}) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)}) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}).$$

Por lo tanto se cumple (2.8) y (2.9) para  $j = 1$ .

Supongamos ahora que (2.8) y (2.9) se cumplen para un  $j$  dado. Por la hipótesis **H3**:

$$f(\mathbf{x}_{\hat{\ell}(k)-j}) \leq f(\mathbf{x}_{\ell(\hat{\ell}(k)-j-1)}) + \alpha \lambda_{\hat{\ell}(k)-j-1} \nabla f_{\hat{\ell}(k)-j-1}^T \mathbf{d}_{\hat{\ell}(k)-j-1}.$$

Como por (2.9)

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(\hat{\ell}(k)-j-1)}) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}),$$

tomando limite cuando  $k \rightarrow \infty$  en ambos lados de la desigualdad se llega a:

$$\lim_{k \rightarrow \infty} \lambda_{\hat{\ell}(k)-j-1} \nabla f_{\hat{\ell}(k)-j-1}^T \mathbf{d}_{\hat{\ell}(k)-j-1} = 0,$$

y usando los mismos argumentos empleados para deducir el **Lema 2.4**, se obtiene:

$$\lim_{k \rightarrow \infty} \lambda_{\hat{\ell}(k)-(j+1)} \|\mathbf{d}_{\hat{\ell}(k)-(j+1)}\| = 0. \quad (2.10)$$

Esto indica que, en el caso arbitrario  $j \geq 1$ , (2.8) se mantiene.

Ahora (2.10) también implica que:  $\|\mathbf{x}_{\hat{\ell}(k)-j} - \mathbf{x}_{\hat{\ell}(k)-(j+1)}\| \rightarrow 0$  cuando  $k \rightarrow \infty$ , así que por (2.9) y la continuidad uniforme de  $f$  en  $\Omega_0$ :

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)-(j+1)}) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)-j}) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)})$$

Por lo tanto, (2.8) y (2.9) se mantienen para cualquier  $j \geq 1$  dado.

Por como se definió  $\hat{\ell}$  al principio y por  $k - m(k) \leq \ell(k) \leq k$  es válido  $\hat{\ell}(k) = \ell(k + M + 2) \leq k + M + 2$ , luego:

$$\hat{\ell}(k) - k - 1 \leq M + 1. \quad (2.11)$$

Ahora, para cualquier  $k$ :

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{x}_{\hat{\ell}(k)} - \sum_{j=1}^{\hat{\ell}(k)-k-1} (\mathbf{x}_{\hat{\ell}(k)-j+1} - \mathbf{x}_{\hat{\ell}(k)-j}) \\ &= \mathbf{x}_{\hat{\ell}(k)} - \sum_{j=1}^{\hat{\ell}(k)-k-1} \lambda_{\hat{\ell}(k)-j} \mathbf{d}_{\hat{\ell}(k)-j} \end{aligned}$$

Por (2.11) y por anterior obtenemos:

$$\|\mathbf{x}_{k+1} - \mathbf{x}_{\hat{\ell}(k)}\| = \left\| - \sum_{j=1}^{\hat{\ell}(k)-k-1} \lambda_{\hat{\ell}(k)-j} \mathbf{d}_{\hat{\ell}(k)-j} \right\| \leq \sum_{j=1}^{M+1} \|\lambda_{\hat{\ell}(k)-j} \mathbf{d}_{\hat{\ell}(k)-j}\|.$$

Tomando el límite cuando  $k \rightarrow \infty$  en ambos lados se tiene:

$$\lim_{k \rightarrow \infty} \|\mathbf{x}_{k+1} - \mathbf{x}_{\hat{\ell}(k)}\| = 0,$$

y ya que la subsucesión  $\{f(\mathbf{x}_{\ell(k)})\}$  es convergente por el **Lema 2.3**, se deduce, de la continuidad uniforme de  $f$  en  $\Omega_0$  que:

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_k) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\hat{\ell}(k)}),$$

luego,

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_k) = \lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}). \quad (2.12)$$

Por (2.3) tenemos:

$$f(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_{\ell(k)}) + \alpha \lambda_k \nabla f_k^T \mathbf{d}_k,$$

tomando límites cuando  $k \rightarrow \infty$  y por (2.12) obtenemos

$$\lim_{k \rightarrow \infty} \lambda_k \nabla f_k^T \mathbf{d}_k = 0,$$

lo que implica, como se ha demostrado antes que:

$$\lim_{k \rightarrow \infty} \lambda_k \|\mathbf{d}_k\| = 0.$$

◆

A partir de las hipótesis de convergencia antes descritas y los lemas con su respectiva prueba, se da paso para enunciar y demostrar el **Teorema de convergencia** dado por *Grippo, Lampariello y Lucidi* en [5].

**Teorema 2.1.** *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función continuamente diferenciable y sea  $\{\mathbf{x}_k\}$  una sucesión definida por:*

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k, \quad \mathbf{d} \neq 0$$

*Supongamos que se cumplen las hipótesis de convergencia **H1** a **H3**. Entonces:*

- a) *La sucesión  $\{\mathbf{x}_k\}$  permanece en  $\Omega_0$  y cada punto límite  $\mathbf{x}^*$  satisface  $\nabla f(\mathbf{x}^*) = 0$ ;*
- b) *Ningún punto límite de  $\{\mathbf{x}_k\}$  es un máximo local de  $f$ ;*
- c) *Si el número de puntos estacionarios de  $f$  en  $\Omega_0$  es finito, la sucesión  $\{\mathbf{x}_k\}$  converge.*

*Demostración.* a) El **Lema 2.2** garantiza que la sucesión  $\{\mathbf{x}_k\}$  está acotada y por el teorema de Bolzano-Weierstrass [1], existe un  $\mathbf{x}^*$  tal que es un punto límite de  $\{\mathbf{x}_k\}$ . Probemos que  $\mathbf{x}^*$  satisface  $\nabla f(\mathbf{x}^*) = 0$ .

Del **Lema 2.5** obtuvimos que:

$$\lim_{k \rightarrow \infty} \lambda_k \|\mathbf{d}_k\| = 0.$$

en particular para  $\mathbf{d}_k = -\nabla f_k$  se da:

$$\lim_{k \rightarrow \infty} \lambda_k \|\nabla f_k\| = 0. \quad (2.13)$$

Ahora, existe una subsucesión  $\{\mathbf{x}_k\}_{k \in K} \subset \{\mathbf{x}_k\}_{k \in \mathbb{N}}$  tal que converge a  $\mathbf{x}^*$ , luego por (2.13) se van a tener los siguientes casos:

1.  $\lim_{k \rightarrow \infty, k \in K} \|\nabla f_k\| = 0,$
2.  $\lim_{k \rightarrow \infty, k \in K} \lambda_k = 0.$

Para el caso 1, por la continuidad del gradiente de  $f$  se tiene que:

$$\lim_{k \rightarrow \infty, k \in K} \|\nabla f(\mathbf{x}_k)\| = \|\nabla f\left(\lim_{k \rightarrow \infty, k \in K} \mathbf{x}_k\right)\| = \|\nabla f(\mathbf{x}^*)\| = 0.$$

Por lo tanto,  $\nabla f(\mathbf{x}^*) = 0$ .

Y para el caso 2, por lo supuesto en **H3** existe un índice  $\bar{k}$  suficientemente grande tal que, para todo  $k \geq \bar{k}, k \in K$  no satisface la condición, es decir:

$$f\left(\mathbf{x}_k + \frac{\lambda_k}{\sigma} \mathbf{d}_k\right) > \max_{0 \leq j \leq m(k)} \{f(\mathbf{x}_{k-j})\} + \alpha \frac{\lambda_k}{\sigma} \nabla f_k^T \mathbf{d}_k \geq f(\mathbf{x}_k) + \alpha \frac{\lambda_k}{\sigma} \nabla f_k^T \mathbf{d}_k \quad (2.14)$$

Ahora mediante el Teorema del Valor Medio, podemos encontrar para cualquier  $k \geq \bar{k}, k \in K$  un punto  $u_k = \mathbf{x}_k + \omega_k \left(\frac{\lambda_k}{\sigma}\right) \mathbf{d}_k$ , con  $\omega_k \in (0, 1)$  en el intervalo  $(\mathbf{x}_k, \mathbf{x}_k + \frac{\lambda_k}{\sigma} \mathbf{d}_k)$  tal que:

$$f\left(\mathbf{x}_k + \frac{\lambda_k}{\sigma} \mathbf{d}_k\right) - f(\mathbf{x}_k) = f'(u_k) \left(\mathbf{x}_k + \frac{\lambda_k}{\sigma} \mathbf{d}_k - \mathbf{x}_k\right) \quad (2.15)$$

de (2.14):

$$f\left(\mathbf{x}_k + \frac{\lambda_k}{\sigma} \mathbf{d}_k\right) - f(\mathbf{x}_k) \geq \alpha \frac{\lambda_k}{\sigma} \nabla f_k^T \mathbf{d}_k.$$

Pero como  $f'(u_k) = \nabla f(u_k)$  y por anterior, se tiene en (2.15):

$$\alpha \frac{\lambda_k}{\sigma} \nabla f_k^T \mathbf{d}_k \leq f\left(\mathbf{x}_k + \frac{\lambda_k}{\sigma} \mathbf{d}_k\right) - f(\mathbf{x}_k) = \nabla f(u_k)^T \frac{\lambda_k}{\sigma} \mathbf{d}_k,$$

luego,

$$\alpha \frac{\lambda_k}{\sigma} \nabla f_k^T \mathbf{d}_k \leq \nabla f(u_k)^T \frac{\lambda_k}{\sigma} \mathbf{d}_k$$

de aquí que la relación (2.14) se puede expresar como:

$$\nabla f(u_k)^T \mathbf{d}_k \geq \alpha \nabla f_k^T \mathbf{d}_k \quad (2.16)$$

Sea ahora una subsucesión  $\{\mathbf{x}_k\}_{k \in K_1} \subset \{\mathbf{x}_k\}_{k \in K}$  tal que:

$$\lim_{k \rightarrow \infty, k \in K_1} \mathbf{x}_k = \mathbf{x}^* \quad , \quad \lim_{k \rightarrow \infty, k \in K_1} \frac{\mathbf{d}_k}{\|\mathbf{d}_k\|} = \bar{\mathbf{d}}$$

Por lo demostrado en el **Lema 2.5**,  $u_k \rightarrow \mathbf{x}^*$ , con  $k \in K_1$ , así que, si dividimos ambos lados de (2.16) por  $\|\mathbf{d}_k\|$  y tomamos limite cuando  $k \rightarrow \infty$  obtenemos:

$$\nabla f(\mathbf{x}^*)^T \bar{\mathbf{d}} \geq \alpha \nabla f(\mathbf{x}^*)^T \bar{\mathbf{d}}$$

luego,

$$(1 - \alpha) \nabla f(\mathbf{x}^*)^T \bar{\mathbf{d}} \geq 0.$$

Como  $1 - \alpha > 0$  y  $\nabla f_k^T \mathbf{d}_k \leq 0$  para todo  $k$ , entonces tenemos:

$$\nabla f(\mathbf{x}^*)^T \bar{\mathbf{d}} = 0,$$

lo cual, implica por **H2**,  $\nabla f(\mathbf{x}^*) = 0$ , y ésto completa la demostración de la parte (a).

**b)** Supongamos por contradicción que existe un punto limite  $\bar{\mathbf{x}} \in \{\mathbf{x}_k\}$  que es un máximo local de  $f$ .

Como sabemos que  $\lim_{k \rightarrow \infty} \|\mathbf{x}_{k+1} - \mathbf{x}_{\hat{\ell}(k)}\| = 0$ , entonces debe existir una subsucesión  $\{\mathbf{x}_{\ell(k)}\}_K \subset \{\mathbf{x}_{\ell(k)}\}$  que converge a  $\bar{\mathbf{x}}$ .

Por otro lado, recordando el **Lema 2.1** y **Lema 2.3**,  $\{f(\mathbf{x}_{\ell(k)})\}$  es decreciente y convergente, tenemos:  $\lim_{k \rightarrow \infty} f(\mathbf{x}_{\ell(k)}) = f(\bar{\mathbf{x}})$  y  $f(\mathbf{x}_{\ell(k)}) \geq f(\bar{\mathbf{x}})$  para todo  $k$ . Además, por (2.5) tenemos:

$$f(\mathbf{x}_{\ell(k')}) < f(\mathbf{x}_{\ell(k)}) \quad \text{para, } k' \geq k + M$$

de modo que, para  $k \in K$  suficientemente grande, podemos encontrar un punto  $\mathbf{x}_{\ell(k)}$  en alguna vecindad de  $\bar{\mathbf{x}}$  que satisface  $f(\mathbf{x}_{\ell(k)}) > f(\bar{\mathbf{x}})$ . Lo que contradice el supuesto de que  $\bar{\mathbf{x}}$  es un máximo.

Por lo tanto queda demostrado la parte (b) del teorema.

c) Por teorema ([8], pág. 476) :

Sea  $g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , continuamente diferenciable en el conjunto compacto  $\Omega_0 \subset D$ , y suponga que el conjunto  $P = \{\mathbf{x} \in \Omega_0 : \nabla g(\mathbf{x}) = 0\}$  de puntos estacionarios de  $g$  en  $\Omega_0$  es finito. Sea  $\{\mathbf{x}_k\} \subset \Omega_0$  cualquier sucesión para la cual  $\lim_{k \rightarrow \infty} (\mathbf{x}_k - \mathbf{x}_{k+1}) = 0$  y  $\lim_{k \rightarrow \infty} \nabla g(\mathbf{x}_k) = 0$ . Entonces  $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$  y  $\nabla g(\mathbf{x}^*) = 0$ ,

y por lo demostrado en el **Lema 2.5** se concluye la convergencia de la sucesión  $\{\mathbf{x}_k\}$ .



# Capítulo 3

## Comparación Numérica

En éste capítulo se van a presentar los resultados numéricos obtenidos, a partir de la implementación del algoritmo propuesto en [5], llamado *búsqueda lineal no monótona* en funciones de prueba y, se realizará la respectiva comparación con el *método de búsqueda lineal monótona*.

Además, para algunas funciones de prueba se buscará el valor más conveniente de  $M$  en la *búsqueda lineal no monótona*, para ver si es posible establecer alguna estimación general de éste. Los resultados obtenidos se presentarán en una tabla y por medio de gráficos.

Para escribir el código del algoritmo y de las *funciones de prueba*, se utilizó el lenguaje MATLAB y las pruebas numéricas fueron realizadas en un computador con procesador Intel(R) Core(TM) i5 de 2.40 GHz.

A continuación se va a presentar el algoritmo dado por *Grippo, Lampariello y Lucidi* en [5]:

---

**Algoritmo 1** Método de Newton con búsqueda lineal no monótona

---

**Entrada:** Función  $f$ , vector inicial  $\mathbf{x}_0$ , gradiente  $\nabla f$ , hessiana  $H$ ,  $\alpha \in (0, 1)$ ,  $M \geq 0$ ,  $\sigma \in (0, 1)$ ,  $N > 0$ ,  $c_1 > 0$ ,  $c_2 > 0$ .

**Salida:**  $\mathbf{x}_k$

- 1: Establecer  $k = 0$ ,  $m(0) = 0$  y calcular  $f(\mathbf{x}_0)$ ,  $\nabla f(\mathbf{x}_0)$  y  $H(\mathbf{x}_0)$ .
  - 2: **mientras**  $\|\nabla f(\mathbf{x}_k)\| > 10^{-8}$  y  $k \leq N$  **hacer**
  - 3:   **si**  $\det(H(\mathbf{x}_k)) < 10^{-10}$  **entonces**
  - 4:      $\mathbf{d}(\mathbf{x}_k) = -\nabla f(\mathbf{x}_k)$
  - 5:   **si no**
  - 6:      $\mathbf{d}(\mathbf{x}_k) = -H^{-1}(\mathbf{x}_k)\nabla f(\mathbf{x}_k)$
  - 7:   **si**  $|\nabla f(\mathbf{x}_k)^T \mathbf{d}(\mathbf{x}_k)| < c_1 \|\nabla f(\mathbf{x}_k)\|^2$  o  $\|\mathbf{d}(\mathbf{x}_k)\| > c_2 \|\nabla f(\mathbf{x}_k)\|$  **entonces**
  - 8:      $\mathbf{d}(\mathbf{x}_k) = -\nabla f(\mathbf{x}_k)$
  - 9:   **si no, si**  $\nabla f(\mathbf{x}_k)^T \mathbf{d}(\mathbf{x}_k) > 0$  **entonces**
  - 10:      $\mathbf{d}(\mathbf{x}_k) = -\mathbf{d}(\mathbf{x}_k)$
  - 11:   **fin si**
  - 12: **fin si**
  - 13: Establecer  $\lambda_k = 1$
  - 14: **mientras**  $f(\mathbf{x}_k + \lambda_k \mathbf{d}_k) > \max_{0 \leq j \leq m(k)} [f(\mathbf{x}_{k-j})] + \alpha \lambda_k \nabla f(\mathbf{x}_k)^T \mathbf{d}_k$  **hacer**
  - 15:    $\lambda_k = \lambda_k \sigma$ .
  - 16: **fin mientras**
  - 17: Establecer  $k = k+1$ ,  $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k$ ,  $f(\mathbf{x}_{k+1})$ ,  $m(k) \leq \min\{m(k-1)+1, M\}$ .
  - 18: **fin mientras**
-

Para las pruebas numéricas, se consideraron cinco problemas de minimización dados en [5] y [7], cada problema descrito a continuación tiene, respectivamente, la función que los define, el punto inicial dado  $\mathbf{x}_0$ , la solución  $\mathbf{x}^*$  y el valor de la solución evaluada en la función  $f(\mathbf{x}^*)$ .

### 1. Función extendida de Rosenbrock

$$f(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2],$$

$$\mathbf{x}_0 = (-1.2, 1, \dots, -1.2, 1)^T, \mathbf{x}^* = (1, 1, \dots, 1, 1)^T \text{ y } f(\mathbf{x}^*) = 0.$$

### 2. Función Wood

$$f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4)^2 + 10.1[(x_2 - 1)^2 + (x_4 - 1)^2] + 19.8(x_2 - 1)(x_4 - 1),$$

$$\mathbf{x}_0 = (-3, -1, -3, -1)^T, \mathbf{x}^* = (1, 1, 1, 1)^T \text{ y } f(\mathbf{x}^*) = 0.$$

### 3. Función singular de Powell

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4,$$

$$\mathbf{x}_0 = (3, -1, 0, 1)^T, \mathbf{x}^* = (0, 0, 0, 0)^T \text{ y } f(\mathbf{x}^*) = 0.$$

### 4. Cubo

$$f(x) = 100(x_2 - x_1^3)^2 + (1 - x_1)^2,$$

$$\mathbf{x}_0 = (-1.2, -1)^T, \mathbf{x}^* = (1, 1)^T \text{ y } f(\mathbf{x}^*) = 0.$$

### 5. Función Freudenstein y Roth

$$f(x) = (x_1 - x_2(2 - x_2(5 - x_2)) - 13)^2 + (x_1 - x_2(14 - x_2(1 + x_2)) - 29)^2,$$

$$\mathbf{x}_0 = (0.5, -2)^T, \mathbf{x}^* = (11.4128, -0.896805)^T \text{ y } f(\mathbf{x}^*) \simeq 24.4921.$$



Rosenbrock con n=10													
M	0	1	2	3	4	5	6	7	8	9	10	15	20
Iteraciones	-	-	-	28	54	70	57	-	75	99	108	130	-
Rosenbrock con n=15													
M	0	1	2	3	4	5	6	7	8	9	10	15	20
Iteraciones	33	38	38	38	60	55	167	113	-	187	176	-	-
Rosenbrock con n=20													
M	0	1	2	3	4	5	6	7	8	9	10	15	20
Iteraciones	-	41	42	55	61	55	48	48	48	48	48	48	48
Rosenbrock con n=50													
M	0	1	2	3	4	5	6	7	8	9	10	15	20
Iteraciones	-	-	84	99	84	84	84	84	84	84	84	84	84
Rosenbrock con n=100													
M	0	1	2	3	4	5	6	7	8	9	10	15	20
Iteraciones	-	-	156	164	154	154	-	91	76	76	-	130	-

Tabla 3.1: Resultados de la función extendida de Rosenbrock

Haciendo una comparación de la *búsqueda lineal monótona* con la *búsqueda lineal no monótona*, de la **Tabla 3.1** se puede observar que en la mayoría de los problemas de la función extendida de *Rosenbrock*, se obtuvieron más éxitos al implementar la búsqueda lineal no monótona, y en cada problema, se evidencia el valor del parámetro  $M$ , con el cual la función logra el menor número de iteraciones, como por ejemplo, para  $n = 2$  se tiene con  $M = 3$ ,  $n = 4$  con  $M = 6$  y  $n = 50$  con  $M = 2$ . Un valor óptimo del parámetro  $M$  que se podría tomar para el método de *búsqueda lineal no monótona* y que funcione para la mayoría de las funciones, podría ser  $M = 9$ , ya que todos los valores de  $n$  con ese valor, presentaron convergencia y para algunas funciones se tienen que el ahorro en iteraciones es más del 50%.

Ésto se puede ver mejor en la **Figura 3.1**, donde se toman las diferentes dimensiones de la función extendida de *Rosenbrock* y se mira la variación del número de iteraciones en cada caso, donde no hay gráfica pintada ni de azul ni de rojo se debe entender que el método establecido de ese color no converge en esa dimensión del problema.

También se puede inferir de la **Tabla 3.1** que, dependiendo de la función y de los parámetros que tomemos al implementar el algoritmo de la búsqueda lineal no monótona, cada vez que se aumenta el tamaño de  $M$  se puede obtener mejores resultados, llegando a un punto en algunas funciones que desde un determinado valor de  $M$  siga con las mismas iteraciones, como es el caso de  $n = 3$  que desde  $M = 7$  no obtiene ningún cambio.

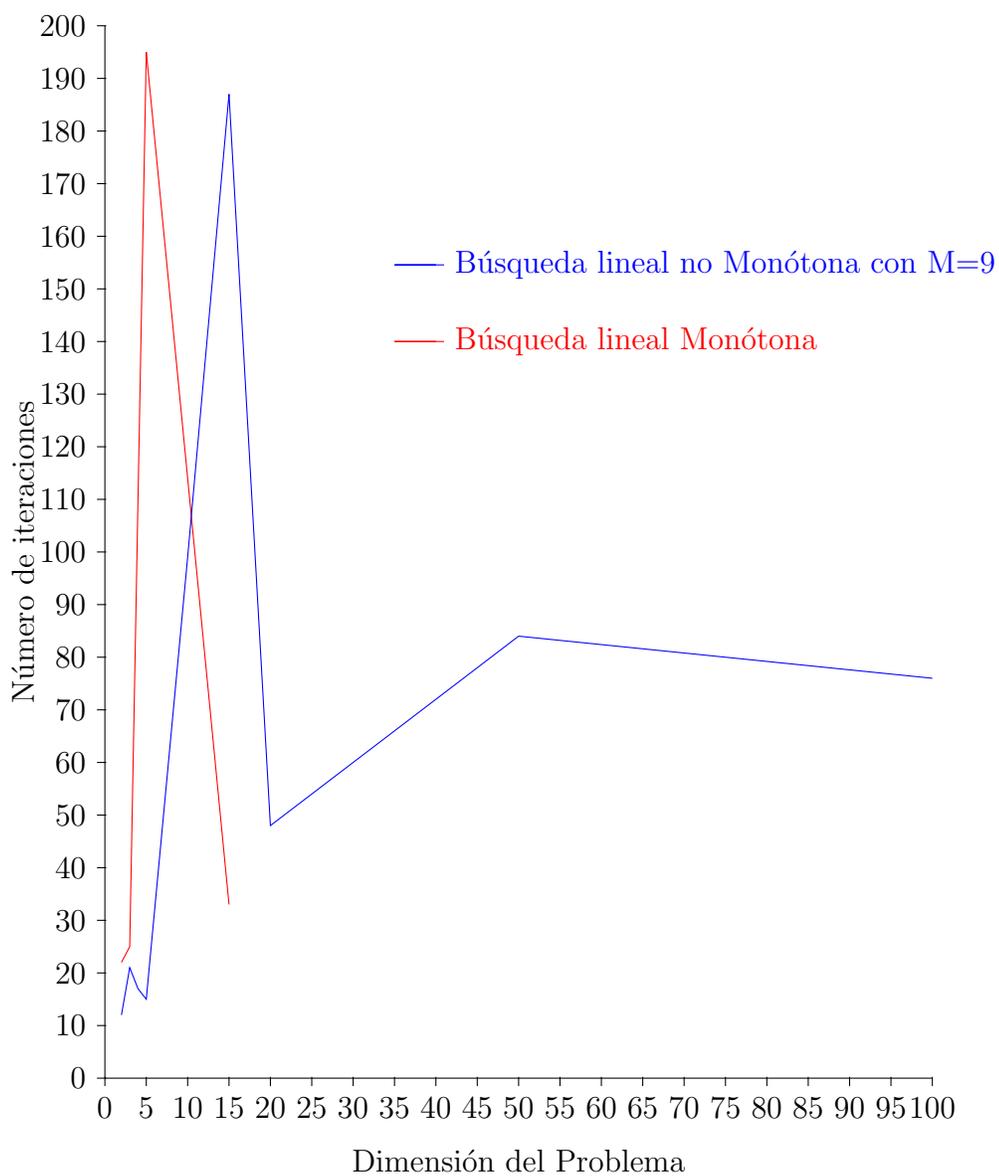


Figura 3.1: **Función extendida de Rosenbrock**

De las **Figuras 3.2, 3.3, 3.4 y 3.5**, se puede observar que al igual que la función extendida de *Rosenbrock*, el ahorro en las iteraciones, sigue siendo casi del 50 % al comparar la *búsqueda lineal monótona* con la *búsqueda lineal no monótona* con  $M = 9$ .

Ahora, notando la **Figura 3.3**, que corresponde a los resultados numéricos de la función singular de *Powell*, el paso completo de Newton satisface las condiciones de Armijo, de modo que los resultados que se obtuvieron son independientes del valor del parámetro  $M$ , esto significa que no importa el valor de  $M$ , el número de iteraciones siempre va a ser el mismo. [5]

Al igual que la función extendida de *Rosenbrock* en algunas dimensiones, de las Figuras también se puede analizar que el **Algoritmo 1** a partir de un cierto valor del parámetro  $M$ , no va a converger más rápido sino que se va a mantener igual las iteraciones, como es el caso por ejemplo de la **Figura 3.2** que corresponde a los resultados numéricos de la función *Wood*, que a partir de  $M = 8$  se mantiene igual, en la **Figura 3.4** desde  $M = 5$  en adelante va a converger en 12 iteraciones y en el caso de la **Figura 3.5** que corresponde a los resultados que se obtuvieron de la función *Freudenstein y Roth*, se tiene que para el caso de Newton con la *condición de Armijo* no converge, y para cualquier tamaño del parámetro  $M$ , es decir con *búsqueda lineal no monótona*, siempre va a converger en 7 iteraciones, lo que resulta muy satisfactorio.

En general no se puede establecer cual de los dos métodos es mejor para la elección del *tamaño de paso* adecuado en cada iteración, en cuestión de tiempo se puede decir que la *búsqueda lineal no monótona* es más costosa computacionalmente, puesto que el **Algoritmo 1** al tener que ir guardando los valores de la función de las iteraciones anteriores, los cuales van a ir aumentando hasta tomar el valor del parámetro  $M$ , va a gastar más tiempo que la *búsqueda lineal monótona*, ésta compara en la iteración actual sólo con el valor de la función de la anterior iteración, pero teniendo en cuenta que el método de *búsqueda lineal no monótona* es otra alternativa para globalizar el método de *Newton*, se debe tener presente a la hora de resolver un problema de minimización, ya que como se pudo ver en algunas funciones, éste resulta más beneficioso en cuestión de número de iteraciones, por ésta razón, mi recomendación para utilizar éste método, es tomar en particular el valor del parámetro  $M = 9$ , ya que a partir de esa cantidad, la mejora en el desempeño no es significativa y todas las funciones tomadas convergieron con ese tamaño.

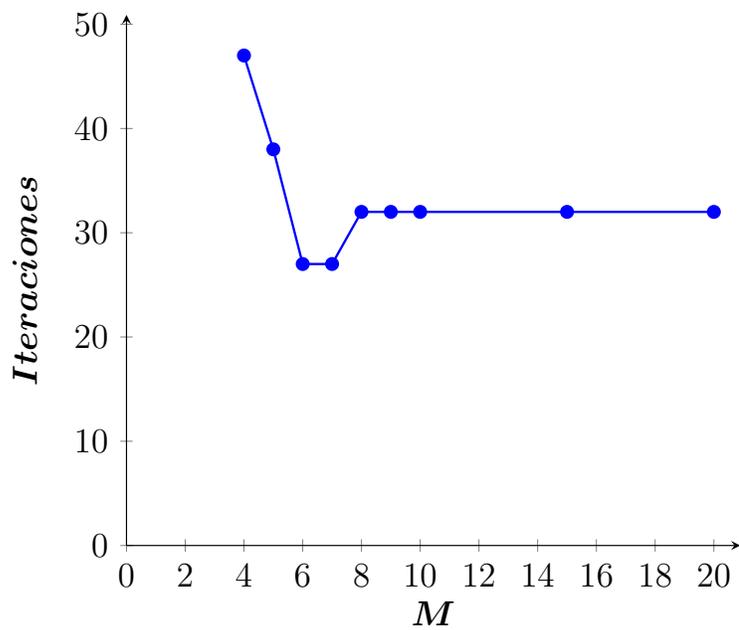


Figura 3.2: Función Wood

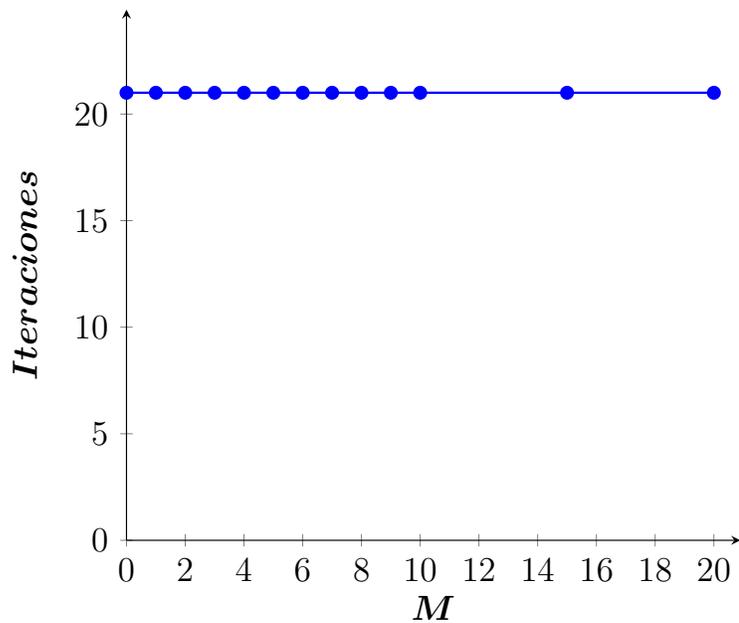


Figura 3.3: Función singular de Powell

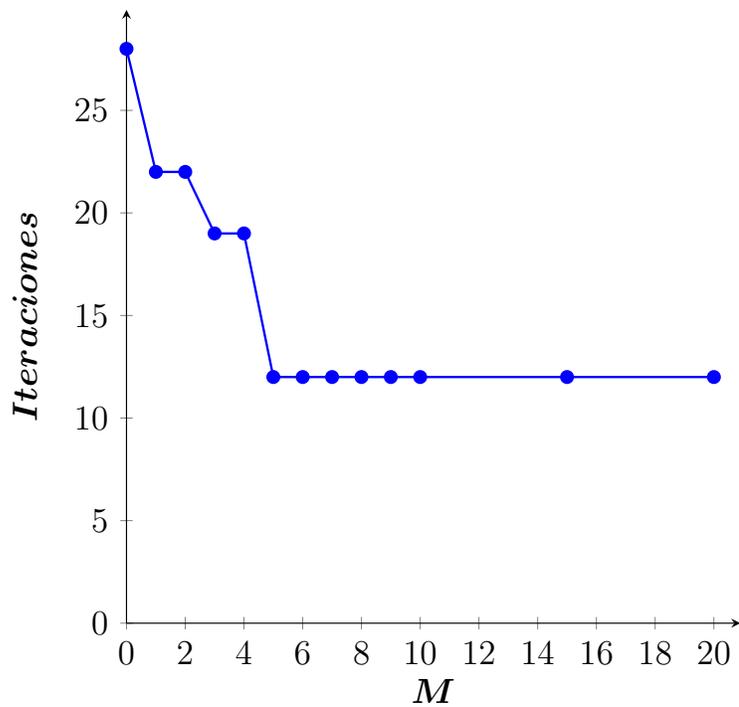


Figura 3.4: **Cubo**

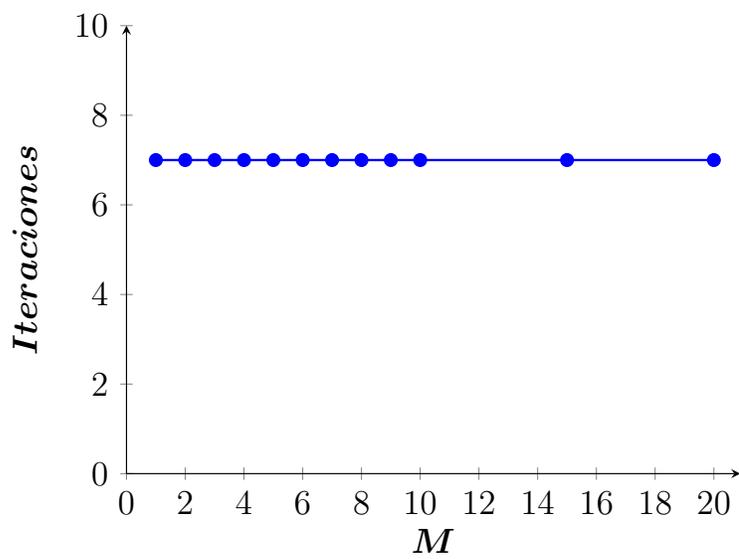


Figura 3.5: **Función Freudenstein y Roth**

# Capítulo 4

## Comentarios Finales

En éste trabajo, al realizar de manera detallada la demostración del teorema de convergencia del *método de búsqueda lineal no monótona*, me ayudó a aplicar la teoría de análisis en sucesiones y comprender mejor el algoritmo propuesto por *Grippo, Lampariello y Lucidi*.

A pesar de que no se puede establecer en general cual de los dos métodos es mejor, de todos los estudios numéricos que se realizaron, se puede decir que para la mayoría de las funciones el *método de búsqueda lineal no monótona* se comporta mejor al resolver un problema de minimización de la forma (1.1), pues éste método garantiza que el descenso de la función en cada iteración sea más notable y rápido.

Para las comparaciones, se implementó el algoritmo propuesto en [5] en diferentes funciones de prueba, las cuales resultaron mejor de lo que se esperaba, ya que el uso de la técnica de *búsqueda lineal no monótona* parece ser valioso especialmente en la etapa inicial e intermedia del proceso de minimización, puesto que en esas etapas ayuda a que la sucesión se acerque más rápido a un minimizador de la función.

Además, las variaciones que se le realizaron al parámetro  $M$ , permite ratificar en algunas funciones la efectividad y velocidad en cuestión de número de iteraciones que puede llegar a obtener el *método de Newton con búsqueda lineal no monótona* comparado con la *búsqueda lineal*.

Finalmente, para ampliar el estudio numérico de la técnica de *búsqueda lineal no monótona*, sería bueno realizar un estudio comparativo del *método de Newton* con la estrategia de globalización de región de confianza [3].

# Bibliografía

- [1] T.M. Apostol, *Análisis matemático*, Reverté, 1996.
- [2] Larry Armijo, *Minimization of functions having lipschitz continuous first partial derivatives.*, Pacific J. Math. **16** (1966), no. 1.
- [3] A.R. Conn, N.I.M. Gould, and P.L. Toint, *Trust region methods*, Society for Industrial and Applied Mathematics, 2000.
- [4] R. Fletcher, *Practical methods of optimization*, no. v. 1, Wiley, 1987.
- [5] L. Grippo, F. Lampariello, and S. Lucidi, *A nonmonotone line search technique for newton's method*, SIAM Journal on Numerical Analysis **23** (1986), no. 4, 707–716.
- [6] Magnus Rudolph Hestenes, *Conjugate direction methods in optimization*, pp. 81–149, Springer New York, 1980.
- [7] Hans Bruun Nielsen, *Uctp - test problems for unconstrained optimization*, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2000.
- [8] J. Ortega and W. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Society for Industrial and Applied Mathematics, 2000.