

## Tarea 04 - Optimización

### Giovanni Gamaliel López Padilla

#### Resumen

Los métodos del descenso del gradiente y de Newton son los métodos más simples en la rama de la optimización. Esto debido a que se basan en principios matemáticos los cuales indican que existe un mínimo global o local en una función con dominio convexo. En este trabajo se analizó la eficiencia de estos dos métodos usando las funciones de Rosembrock, de Wood y de suavizado. Llegando así a que el método del descenso del gradiente tiende a obtener mínimos locales cercanos al mínimo global en cuestión del valor de la función a analizar. Por otro lado el método de Newton obtiene un mínimo local cercano al punto inicial con un número menor de iteraciones a comparación del método del descenso del gradiente.

## Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Marco teórico</b>	<b>2</b>
2.1. Convexidad . . . . .	2
2.2. Condiciones necesarias . . . . .	2
2.3. Direcciones de búsqueda . . . . .	3
2.4. Condiciones de Wolfe . . . . .	3
<b>3. Métodos</b>	<b>4</b>
3.1. Método de descenso del gradiente . . . . .	4
3.2. Método de Newton . . . . .	4
3.3. Condiciones de paro . . . . .	4
3.3.1. Por posición . . . . .	4
3.3.2. Por función . . . . .	5
3.3.3. Por gradiente . . . . .	5
3.4. Algoritmos implementados . . . . .	5
3.4.1. Método del descenso del gradiente . . . . .	5
3.4.2. Método de Newton . . . . .	5
3.5. Función de Rosembrock . . . . .	6
3.6. Función de Wood . . . . .	6
3.7. Función de suavizado . . . . .	7
3.8. Vectores iniciales aleatorios . . . . .	7
<b>4. Resultados</b>	<b>7</b>
4.1. Función de Rosembrock . . . . .	8
4.1.1. 2 dimensiones . . . . .	8
4.1.2. 2 dimensiones con vector aleatorio . . . . .	9
4.1.3. 100 dimensiones . . . . .	9
4.1.4. 100 dimensiones con vector aleatorio . . . . .	10
4.1.5. Estadísticas con vectores aleatorios . . . . .	11
4.2. Función de Wood . . . . .	12
4.2.1. Vector predefinido . . . . .	12

4.3. Vector aleatorio . . . . .	13
4.3.1. Estadísticas . . . . .	14
4.4. Función de suavizado . . . . .	14
4.4.1. Primer caso. . . . .	15
4.4.2. Segundo caso. . . . .	15
4.4.3. Tercer caso. . . . .	16

5. Conclusiones	17
-----------------	----

## 1. Introducción

La optimización es procedimiento que con sus resultados se toman de decisiones en el análisis de sistemas físicos. Para realizar es la optimización de un sistema o situación se debe de contemplar un objetivo, el cual debe ser caracterizado por una función cuantitativa.

El resultado de realizar la optimización a un sistema puede representarse como un ahorro de tiempo, energía o cualquier objeto que pueda ser reflejado en un número. El objetivo del proceso de una optimización es obtener un conjunto de números o características que representen un mínimo o máximo de objeto el cual esta siendo caracterizado. Esto puede ser representado como se encuentra en la ecuación 1.

$$\min_{x \in \mathbb{R}} f(x) \quad (1)$$

donde  $f$  es la caracterización cuantitativa del problema y  $x$  son los objetos que interactúan con el sistema.

Una optimización local es una solución al problema de optimización en una vecindad alrededor del valor de  $x$  encontrado. En cambio una optimización global es aquella solución que es menor o mayor con respecto a todas las demás. La solución de un proceso de optimización no siempre encontrará los valores en que el sistema se sitúe una optimización global.

## 2. Marco teórico

### 2.1. Convexidad

El concepto de convexidad en los problemas de optimización es de gran importancia. Este concepto aporta una mayor facilidad al resolver un problema. La convexidad puede ser aplicado a conjunto o funciones. Se dice que  $S$  es un conjunto convexo si un segmento de línea conecta cualquier par de puntos en  $S$ . Sean  $x, y \in S$ , entonces  $\alpha x + (1 - \alpha)y \in S$  donde  $\alpha \in [0, 1]$ . Se dice que una función es convexa si su dominio  $S$  es convexo y cumplen con la propiedad escrita en la ecuación 2.

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \quad (2)$$

Donde  $x, y \in S$  y  $\alpha \in [0, 1]$ .

### 2.2. Condiciones necesarias

Suponiendo que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  y que es continuamente diferenciable y  $p \in \mathbb{R}^n$ . Entonces, tenemos que

$$f(x + p) = f(x) + \nabla f(x + tp)^T p$$

donde  $t \in (0, 1)$ . De igual manera, si  $f$  es doblemente continua diferenciable, entonces se tiene que:

$$\nabla f(x+p) = \nabla f(x) \int_0^1 \nabla^2 f(x+tp) p dt$$

por lo tanto

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x) p$$

Con estas hechas, se puede demostrar que si  $x^*$  es un punto estacionario entonces el gradiente y el hessiano de  $f$  tiene las características mostradas en la ecuación 3.

$$\nabla f(x^*) = 0 \quad \nabla^2 f(x^*) \succ 0 \quad (3)$$

También se puede demostrar que si  $\nabla^2 f$  es continua en una vecindad alrededor de  $x^*$  y que  $\nabla f(x^*) = 0$  y  $\nabla^2 f(x^*)$  es positiva definida. Entonces  $x^*$  es un mínimo local de  $f$ .

### 2.3. Direcciones de búsqueda

La dirección mas eficiente usando un método de descenso es usar una dirección  $p$  descrita en la ecuación 4 en cada paso.

$$p_k = -\nabla f_k \quad (4)$$

En cada iteración de la línea de búsqueda se implementa un cambio en la posición siguiendo la ecuación 5.

$$x_{k+1} = x_k + \alpha_k p_k \quad (5)$$

Donde  $\alpha_k$  es un escalar positivo llamado tamaño de paso.

### 2.4. Condiciones de Wolfe

Existen maneras de verificar si el  $\alpha_k$  elegido para el  $k$ -ésimo paso es el óptimo para seguir en la dirección  $p_k$ . La condición de decrecimiento suficiente esta descrita en la ecuación 6.

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k \quad (6)$$

La interpretación de este resultado es que la función  $f$  debe ser proporcional al tamaño de paso  $\alpha_k$  y la derivada direccional  $\nabla f_k^T p_k$  para una constante  $c_1 \in (0, 1)$ . Esta condición también es conocida como la condición de Armijo.

La condición de decrecimiento suficiente no es la única que se tiene que contemplar, esto debido a que existen  $\alpha_k$  muy pequeñas que satisfacen a la desigualdad. Es por ello que se tiene que contemplar la condición de curvatura. La condición de curvatura esta definida en la ecuación 7.

$$\nabla f(x_k + \alpha_k p_k)^T \geq c_2 \nabla f_k^T p_k \quad (7)$$

Donde  $c_2 \in (0, 1)$ .

La búsqueda de un  $\alpha_k$  que cumpla las condiciones de las ecuaciones 6 y 7 esta descrito en el algoritmo 1.

**Algorithm 1:** Búsqueda de un  $\alpha$  que cumpla las condiciones de las ecuaciones 6 y 7

---

```

1  $\alpha_0 \leftarrow 0$        $\alpha_i \leftarrow 1$        $\beta \leftarrow \infty$ 
2 repeat
3   if armijo conditon( $\alpha_i$ ) or curvature contidion( $\alpha_i$ ) then
4     if armijo condition( $\alpha_i$ ) then
5        $\beta \leftarrow \alpha_i$ 
6        $\alpha_i = \frac{\beta + \alpha}{2}$ 
7     else if curvature condition ( $\alpha_i$ ) then
8        $\alpha \leftarrow \alpha_i$ 
9       if  $\beta$  equals  $\infty$  then
10         $\alpha_i \leftarrow 2\alpha$ 
11      else
12         $\alpha_i = \frac{\beta + \alpha}{2}$ 
13   else
14     break
15 return  $\alpha_i$ 

```

---

### 3. Métodos

#### 3.1. Método de descenso del gradiente

La dirección que se tomada con el método del descenso del gradiente es el negativo del gradiente, esto es que la ecuación de cada paso en el método esta descrita en la ecuación 8.

$$x_{k+1} = x_k - \alpha_k \nabla f_k \quad (8)$$

donde se dice que  $\alpha_k$  es fijo si para todo  $k$  el valor de  $\alpha$  no cambia. Por lo que se convierte en un parametro del método.

#### 3.2. Método de Newton

La direccion que es tomada en el método de Newton es obtenida al resolver el sistema matricial de la ecuación 9.

$$\nabla^2 f_k d_k = -\nabla f_k \quad (9)$$

donde  $-d_k$  es la dirección de descenso que se tomara siguiendo la ecuación 10.

$$x_{k+1} = x_k - \alpha_k d_k \quad (10)$$

donde  $\alpha_k$  sería obtenido en cada iteración siguiendo el algoritmo 1.

#### 3.3. Condiciones de paro

##### 3.3.1. Por posición

Una de las condiciones tomadas para parar los métodos de Newton y descenso del gradiente es comprobando si la norma de la diferencia entre la posición  $k$  y  $k + 1$  es menor igual a una tolerancia establecida. Esta condición es descrita en la ecuación 11.

$$\|x_{k+1} - x_k\| \leq \tau \quad (11)$$

### 3.3.2. Por función

Otra manera de comprobar si ya no existe algún movimiendo notable en el método es revisando el valor absoluto de la diferencia de la función evaluada en  $x_{k+1}$  y  $x_k$ . La condición esta descrita en la ecuación 12.

$$|f(x_{k+1}) - f(x_k)| \leq \tau \quad (12)$$

### 3.3.3. Por gradiente

Como estamos en la búsqueda de puntos estacionarios de  $f$ , es razonable que una condición de paro este relacionada al gradiente de la function. La condición usando al gradiente de la función esta descrita en la ecuación 13.

$$\|\nabla f(x_{k+1})\| \leq \tau \quad (13)$$

## 3.4. Algoritmos implementados

### 3.4.1. Método del descenso del gradiente

El algoritmo 2 se implemento para el método del descenso del gradiente. El algoritmo hace uso de las condiciones de paro descritas en 3.3.

---

**Algorithm 2:** Método del descenso de gradiente con  $\alpha$  fija.

---

**Input:**  $x_0, f, \nabla f, \alpha$   
**Output:**  $x_{i+1}, \nabla f(x_{i+1})$

```

1  $x_j \leftarrow x_0$ 
2 repeat
3    $x_i \leftarrow x_j$ 
4    $d_i \leftarrow -\nabla f(x_i)$ 
5    $x_{i+1} \leftarrow x_i + \alpha d_i$ 
6   check stop conditions( $f, x_i, x_{i+1}, \nabla f(x_{i+1})$ )
7 return  $x_{i+1}, \nabla f(x_{i+1})$ 
```

---

### 3.4.2. Método de Newton

El algoritmo 3 se implemento para el método de Newton. El mismo hace de las condiciones de Wolfe descritas en la sección 2.4 y las condiciones de paro descritas en 3.3.

---

**Algorithm 3:** Algoritmo de Newton usando las condiciones de Wolfe.

---

**Input:**  $x_0, f, \nabla f$   
**Output:**  $x_{i+1}, \nabla f(x_{i+1})$

```

1  $x_j \leftarrow x_0$ 
2 repeat
3    $x_i \leftarrow x_j$ 
4    $d_i \leftarrow \text{solve}(\nabla^2 f(x_i)d_i = \nabla f(x_i))$ 
5    $\alpha_i \leftarrow \text{Wolfe conditions}()$ 
6    $x_{i+1} \leftarrow x_i - \alpha_i d_i$ 
7   check stop conditions( $f, x_i, x_{i+1}, \nabla f(x_{i+1})$ )
8 return  $x_{i+1}, \nabla f(x_{i+1})$ 
```

---

### 3.5. Función de Rosembrock

La función de Rosembrock se define en la ecuación 14.

$$f(x) = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (1 - x_i) \quad (14)$$

donde  $x \in \mathbb{R}^n$

Con la función de Rosembrock definida, se tiene que su gradiente es calculado con en la ecuación 15.

$$\nabla f(x) = \begin{cases} -400x_i(x_{i+1} - x_i^2) & \text{para } i = 1 \\ 200(x_i - x_{i-1}^2 - 400x_i(x_{i+1} - x_i^2) - 2(1 - x_i)) & \text{para } 1 < i < n \\ 200(x_i - x_{i-1}) & \text{para } i = n \end{cases} \quad (15)$$

Con la función de Rosembrock definida, se tiene que su hessiano es calculado con la ecuación 16.

$$\nabla^2 f(x) = \begin{cases} \frac{\partial^2 f}{\partial x_i^2} = -200 & \text{para } i = 1 \\ \frac{\partial^2 f}{\partial x_i^2} = 1200x_i^2 - 400x_{i+1} + 202 & \text{para } 1 < i < n \\ \frac{\partial^2 f}{\partial x_i \partial x_{i+1}} = -400x_i & \text{para } 0 \leq i < n \\ \frac{\partial^2 f}{\partial x_i^2} = 200 & \text{para } i = n \end{cases} \quad (16)$$

El vector inicial predefinido que se usará con la función de Rosembrock esta definido en la ecuación 17.

$$x = [-1.2 \quad 1 \quad 1 \quad \dots \quad -1.2 \quad 1]^T \quad (17)$$

### 3.6. Función de Wood

La función de Wood se define en la ecuación 18.

$$f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3 - x_4)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1) \quad (18)$$

Donde  $x \in \mathbb{R}^4$ .

Con la función de Wood definida, podemos obtener le gradiente de la función de Wood. El resultado del gradiente de la función de Wood se encuentra en la ecuación 19.

$$\nabla f(x) = \begin{cases} \frac{\partial f}{\partial x_1} = 400(x_1^2 - x_2)x_1 + 2(x_1 - 1) \\ \frac{\partial f}{\partial x_2} = -200(x_1^2 - x_2) + 20.2(x_2 - 1) + 19.8(x_4 - 1) \\ \frac{\partial f}{\partial x_3} = 2(x_3 - 1) + 360(x_3^2 - x_4)x_3 \\ \frac{\partial f}{\partial x_4} = -180(x_3^2 - x_4) + 20.2(x_4 - 1) + 19.8(x_2 - 1) \end{cases} \quad (19)$$

De igual forma, se puede obtener el hessiano de la función de Wood. El resultado del Hessiano se encuentra en la ecuación 20.

$$\nabla^2 f(x) = \begin{cases} \frac{\partial^2 f}{\partial x_1^2} = 400(x_1^2 - x_2) + 800x_1^2 + 2 \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -400x_1 \\ \frac{\partial^2 f}{\partial x_2^2} = 220.2 \\ \frac{\partial^2 f}{\partial x_4 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_4} = 19.8 \\ \frac{\partial^2 f}{\partial x_3^2} = 720x_3^2 + 360(x_3^2 - x_4) + 2 \\ \frac{\partial^2 f}{\partial x_4 \partial x_3} = \frac{\partial^2 f}{\partial x_3 \partial x_4} = -360x_3 \\ \frac{\partial^2 f}{\partial x_4^2} = 200.2 \end{cases} \quad (20)$$

El vector inicial predefinido para el problema de la función de Wood se encuentra definido en la ecuación 21.

$$x = \begin{bmatrix} -3 & -1 & -3 & -1 \end{bmatrix}^T \quad (21)$$

### 3.7. Función de suavizado

La función de suavizado se define en la ecuación 22.

$$f(x) = \sum_{i=1}^n (x_i - y_i)^2 + \lambda \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 \quad (22)$$

donde  $x, y \in \mathbb{R}^n$ ,  $\lambda > 0$ . El vector  $y$  se propuso obtenerse conforme a la ecuación .

$$y = t_i^2 + \eta \quad t_i = \frac{2i}{n-1} \quad i = 0, 1, \dots, n-1 \quad \eta \in \mathcal{N}(0, \sigma) \quad (23)$$

Calculando el gradiente de la función de suavizado se obtiene lo mostrado en la ecuación 24.

$$\nabla f(x) = \begin{cases} \frac{\partial f}{\partial x_1} = 2(x_1 - y_1) - 2\lambda(x_1 - x_0) \\ \frac{\partial f}{\partial x_i} = 2(x_i - y_i) + 2\lambda(2x_i - x_{i+1} - x_{i-1}) \text{ para } 1 < i < n \\ \frac{\partial f}{\partial x_n} = 2(x_n - y_n) + 2\lambda(x_n - x_{n-1}) \end{cases} \quad (24)$$

Para la prueba del

### 3.8. Vectores iniciales aleatorios

Para realizar una evaluación a los métodos se implemento la creación de vectores aleatorios. Cada elemento del vector sigue una distribución normal con media 0 y sigma 0.5 ( $\mathcal{N}(0, 0.5)$ ).

$$x = \{x_i \in \mathcal{N}(0, 0.5)\} \quad (25)$$

## 4. Resultados

Los parámetros escogidos para los métodos de descenso de gradiente y de Newton usando la función de Rosembrock, Wood y suavizado se encuentran descritos en la tabla 1.

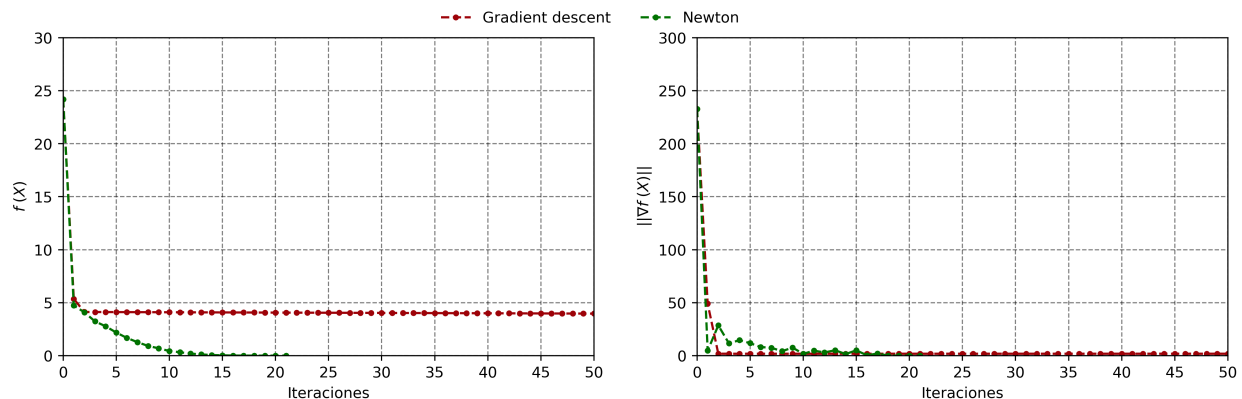
Función	Métodos	Tipo de vector	Tamaño del vector	Tolerancia	$c_1$	$c_2$	$\lambda_0$
Rosenbrock	Newton	Predefinido	2	$1 \times 10^{-6}$	0.1	0.9	
			100	$1 \times 10^{-12}$	0.1	0.9	
		Aleatorio	2	$1 \times 10^{-7}$	0.1	0.9	
			100	$1 \times 10^{-8}$	0.1	0.9	
		Predefinido	2	$1 \times 10^{-6}$			$1 \times 10^{-3}$
			100	$1 \times 10^{-6}$			$1 \times 10^3$
		Aleatorio	2	$1 \times 10^{-6}$			$1 \times 10^4$
			100	$1 \times 10^{-8}$			$1 \times 10^3$
Wood	Newton	Predefinido	4	$1 \times 10^{-6}$	0.1	0.9	
		Aleatorio	4	$1 \times 10^{-8}$	0.1	0.9	
	Descenso del gradiente	Predefinido	4	$1 \times 10^{-6}$			$1 \times 10^{-4}$
		Aleatorio	4	$1 \times 10^{-6}$			$1 \times 10^{-4}$

**Tabla 1:** Parametros usados para evaluar las funciones de Wood y Rosembrock para el método de Newton y el el descenso del gradiente dado un vector definido y aleatorio dado por las ecuaciones 17, 21 y 25.

## 4.1. Función de Rosembrock

### 4.1.1. 2 dimensiones

En la figura 1 se muestran los resultados de cada iteración obtenidos por la función de Rosembrock para el caso de  $n = 2$ .



**Figura 1:** Iteraciones de los métodos del descenso del gradiente y de Newton aplicados a la función de Rosembrock para  $n=2$ .

En la tabla 2 se muestran los resultados de la primer y última iteracion para cada método para la función de Rosembrock para el caso de  $n = 2$ .

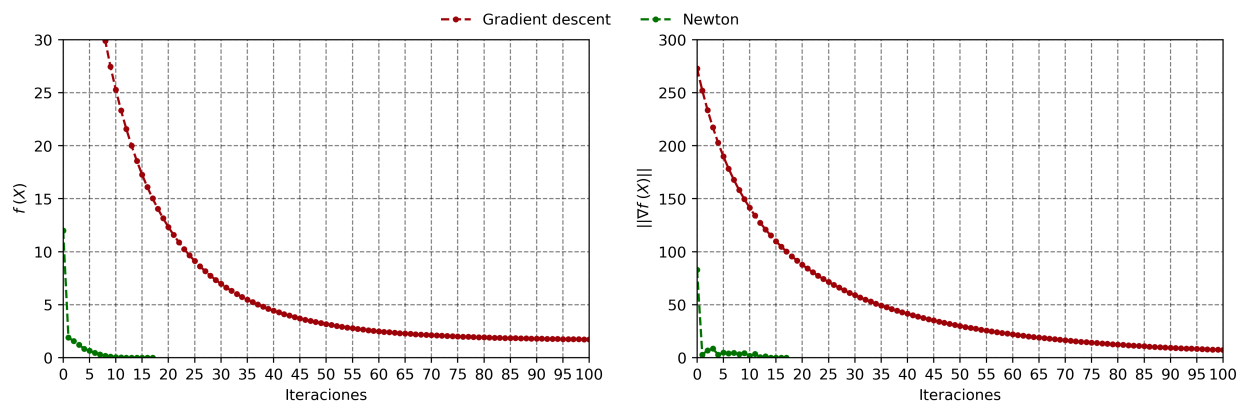


Método	Iteraciones	$f(x_0)$	$f(x_n)$	$\nabla f(x_0)$	$\nabla f(x_n)$
Descenso del gradiente	13931	24.2	0.000002	232.8676	0.001410
Newton	21	24.2	$7.68 \times 10^{-24}$	232.8676	$1.216668 \times 10^{-10}$

**Tabla 2:** Resultados de la primer y última iteración del método del descenso del gradiente y de Newton para la función de Rosembrock para dos dimensiones.

#### 4.1.2. 2 dimensiones con vector aleatorio

En la figura 2 se muestran los resultados de cada iteración obtenidos por la función de Rosembrock para el caso de  $n = 2$ .



**Figura 2:** Iteraciones de los métodos del descenso del gradiente y de Newton aplicados a la función de Rosembrock para  $n=2$ .

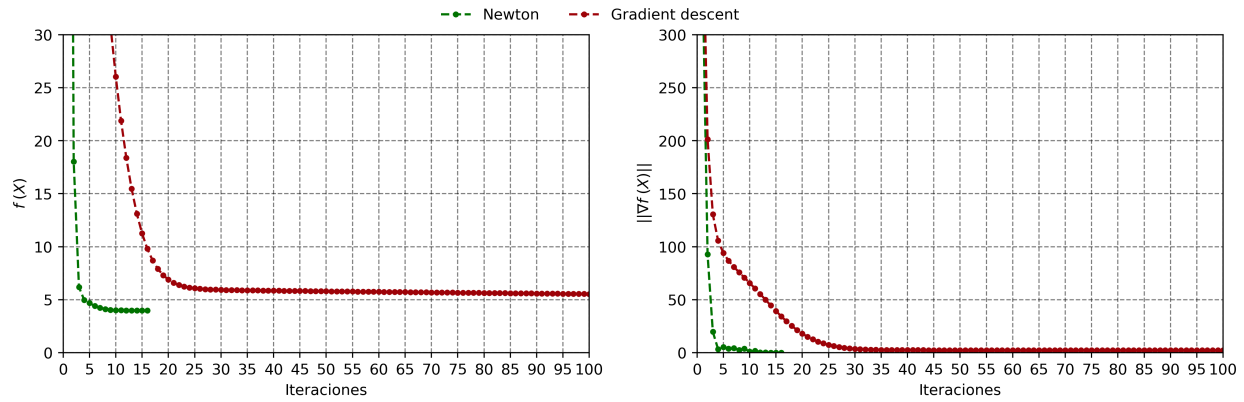
En la tabla 3 se muestran los resultados de la primer y última iteración para cada método.

Método	Iteraciones	$f(x_0)$	$f(x_n)$	$\nabla f(x_0)$	$\nabla f(x_n)$
Descenso del gradiente	76754	66.316393	0.000233	272.967314	0.013819
Newton	17	11.98872	$5.628 \times 10^{-23}$	82.95111	$3.115 \times 10^{-10}$

**Tabla 3:** Resultados de la primer y última iteración del método del descenso del gradiente y de Newton para la función de Rosembrock para dos dimensiones.

#### 4.1.3. 100 dimensiones

En la figura 3 se muestran los resultados de cada iteración obtenidos por la función de Rosembrock para el caso de  $n = 100$ .



**Figura 3:** Iteraciones de los métodos del descenso del gradiente y de Newton aplicados a la función de Rosembrock para  $n=100$ .

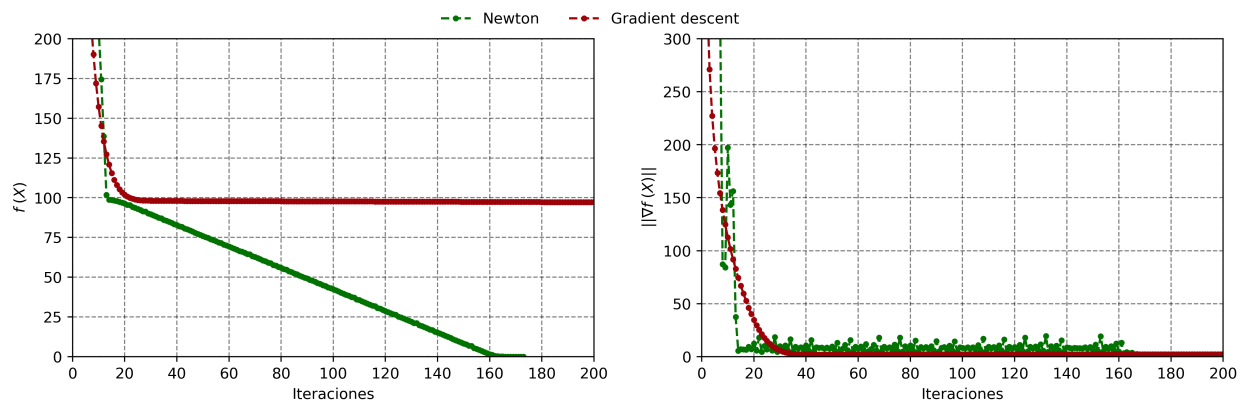
En la tabla 4 se muestran los resultados de la primer y última iteracion para cada método.

Método	Iteraciones	$f(x_0)$	$f(x_n)$	$\nabla f(x_0)$	$\nabla f(x_n)$
Descenso del gradiente	7337	532.40	3.986624	1125.2478	0.009986
Newton	16	532.40	3.986624	1125.2478	$3.15602 \times 10^{-13}$

**Tabla 4:** Resultados de la primer y última iteración del método del descenso del gradiente y de Newton para la función de Rosembrock para 100 dimensiones.

#### 4.1.4. 100 dimensiones con vector aleatorio

En la figura 4 se muestran los resultados de cada iteración obtenidos por la función de Rosembrock para el caso de  $n = 100$ .



**Figura 4:** Iteraciones de los métodos del descenso del gradiente y de Newton aplicados a la función de Rosembrock para  $n=100$ .

En la tabla 5 se muestran los resultados de la primer y última iteracion para cada método.

Método	Iteraciones	$f(x_0)$	$f(x_n)$	$\nabla f(x_0)$	$\nabla f(x_n)$
Descenso del gradiente	7337	532.40	3.986624	1125.2478	0.009986
Newton	173	3824.008	$7.5681 \times 10^{-30}$	2682.654	$1.3651 \times 10^{-13}$

**Tabla 5:** Resultados de la primer y última iteración del método del descenso del gradiente y de Newton para la función de Rosembrock para 100 dimensiones.

#### 4.1.5. Estadísticas con vectores aleatorios

Se crearon 400 vectores aleatorios siguiendo la ecuación 25. Se organizaron conjuntos de 100 vectores para que cada conjunto sea el vector inicial en cada método para la función de Rosembrock de  $n = 2$  y  $n = 100$ .

##### 4.1.5.1 2 dimensiones

###### 4.1.5.1.1 Método de Newton

En la tabla 6 se concentran los valores medios, desviación estándar, mínimo y máximo de la función y el gradiente de Rosembrock en los puntos estacionarios encontrados con el método de Newton.

	$f(x)$	$  \nabla f(x)  $
n	100	100
Media	0.051269	0.189574
$\sigma$	0.300291	1.165927
Mínimo	0.000000	0.000000
Máximo	2.238695	8.806987

**Tabla 6:** Media, desviación estándar, mínimo y máximo de los resultados obtenidos con vectores iniciales aleatorios usando el método de Newton en la función de Rosembrock para  $n = 2$ .

###### 4.1.5.1.2 Método del descenso del gradiente

En la tabla 7 se concentran los valores medios, desviación estándar, mínimo y máximo de la función y el gradiente de Rosembrock en los puntos estacionarios encontrados con el método del descenso del gradiente.

	$f(x)$	$  \nabla f(x)  $
n	100	100
Media	$2.333063 \times 10^{-04}$	$1.381915 \times 10^{-02}$
$\sigma$	$5.679958 \times 10^{-09}$	$1.702874 \times 10^{-07}$
Mínimo	$2.332966 \times 10^{-04}$	$1.381886 \times 10^{-02}$
Máximo	$2.333152 \times 10^{-04}$	$1.381942 \times 10^{-02}$

**Tabla 7:** Media, desviación estándar, mínimo y máximo de los resultados aleatorios usando el método de Newton en la función de Rosembrock para  $n = 2$ .

#### 4.1.5.2 100 dimensiones

##### 4.1.5.2.1 Método de Newton

En la tabla 8 se concentran los valores medios, desviación estandar, mínimo y máximo de la función y el gradiente de Rosembrock en los puntos estacionarios encontrados con el método de Newton.

	$f(x)$	$  \nabla f(x)  $
n	100	100
Media	426.440128	290.794079
$\sigma$	1002.282274	722.327174
Mínimo	0.000000	0.000000
Máximo	5344.705735	3423.160112

**Tabla 8:** Media, desviación estandar, mínimo y máximo de los resultados aleatorios usando el método de Newton en la función de Rosembrock para  $n = 100$ .

##### 4.1.5.2.2 Método del descenso del gradiente

En la tabla 9 se concentran los valores medios, desviación estandar, mínimo y máximo de la función y el gradiente de Rosembrock en los puntos estacionarios encontrados con el método de Newton.

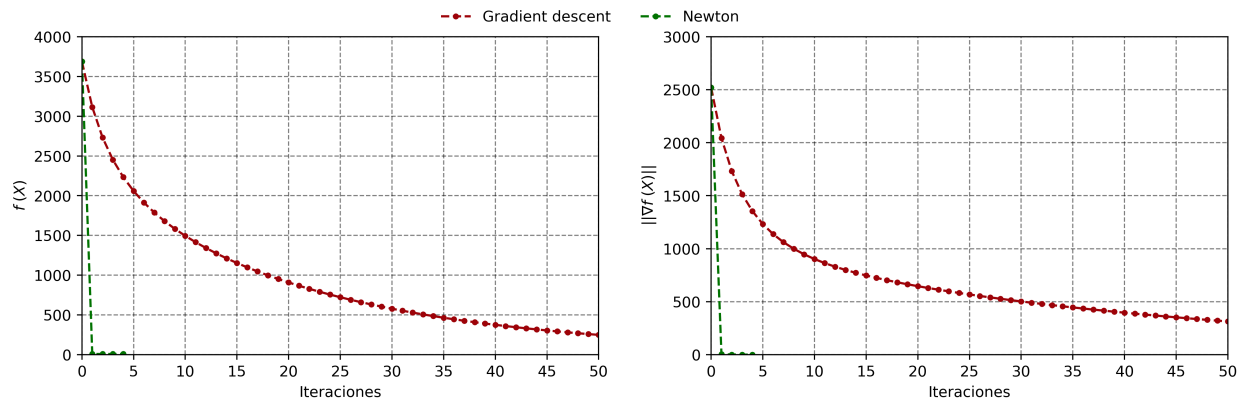
	$f(x)$	$  \nabla f(x)  $
n	100	100
Media	$2.391974x10^{-1}$	$9.992327x10^{-5}$
$\sigma$	$9.515404x10^{-1}$	$1.602367x10^{-8}$
Mínimo	$1.000360x10^{-8}$	$9.988768x10^{-5}$
Máximo	3.986624	$9.994944x10^{-5}$

**Tabla 9:** Media, desviación estandar, mínimo y máximo de los resultados aleatorios usando el método de Newton en la función de Rosembrock para  $n = 100$ .

## 4.2. Función de Wood

### 4.2.1. Vector predefinido

En la figura 5 se muestran los resultados de cada iteración obtenidos por la función de Wood partiendo del vector de la ecuación 21.



**Figura 5:** Iteraciones de los métodos del descenso del gradiente y de Newton aplicados a la función de Wood.

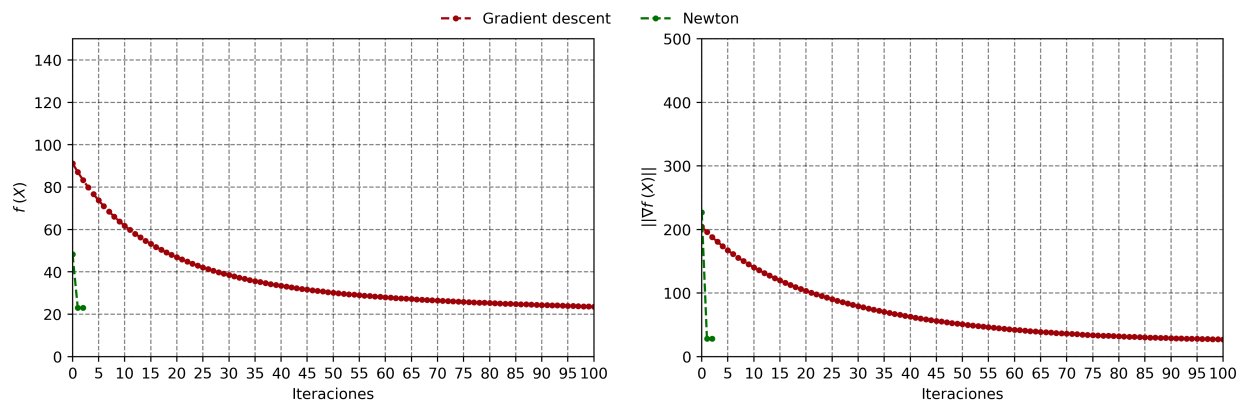
En la tabla 10 se muestran los resultados de la primer y última iteración para cada método.

Método	Iteraciones	$f(x_0)$	$f(x_n)$	$\nabla f(x_0)$	$\nabla f(x_n)$
Descenso del gradiente	3094	0.000081	0.000002	2520.7046	0.019937
Newton	4	3688.000	7.876963	2520.7046	0.012216

**Tabla 10:** Resultados de la primer y última iteración del método del descenso del gradiente y de Newton para la función de Rosembrock para dos dimensiones.

### 4.3. Vector aleatorio

En la figura 6 se muestran los resultados de cada iteración obtenidos por la función de Wood partiendo del vector de la ecuación 25.



**Figura 6:** Iteraciones de los métodos del descenso del gradiente y de Newton aplicados a la función de Wood.

En la tabla 11 se muestran los resultados de la primer y última iteración para cada método.

Método	Iteraciones	$f(x_0)$	$f(x_n)$	$\nabla f(x_0)$	$\nabla f(x_n)$
Descenso del gradiente	72711	91.214553	0.000278	2520.7046	0.019999
Newton	2	48.331395	23.010416	227.054327	28.070864

**Tabla 11:** Resultados de la primer y última iteración del método del descenso del gradiente y de Newton para la función de Rosembrock para dos dimensiones.

#### 4.3.1. Estadísticas

Se ejecutaron los métodos de Newton y descenso del gradiente con 100 vectores aleatorios. Cada vector aleatorio esta creado en base a la ecuación 25.

##### 4.3.1.1 Método de Newton

En la tabla 12 se muestra la media, desviación estandar, mínimo y máximo de la función y el gradiente de Wood evaluados en el punto estacionario encontrado por el método de Newton.

	$f(x)$	$  \nabla f(x)  $
n	100	100
Media	14.12856	13.14793
$\sigma$	14.38304	16.05534
Mínimo	0	0
Máximo	40.22492	72.09054

**Tabla 12:** Media, desviación estandar, mínimo y máximo de los resultados obtenidos con vectores iniciales aleatorios usando el método de Newton en la función de Wood.

##### 4.3.1.2 Método del descenso de Gradiente

En la tabla 13 se muestra la media, desviación estandar, mínimo y máximo de la función y el gradiente de Wood evaluados en el punto estacionario encontrado por el método del descenso del gradiente.

	$f(x)$	$  \nabla f(x)  $
n	100	100
Media	0.315318	0.019965
$\sigma$	1.551148	0.000167
Mínimo	0.000278	0.019122
Máximo	7.876494	0.020000

**Tabla 13:** Media, desviación estandar, mínimo y máximo de los resultados obtenidos con vectores iniciales aleatorios usando el método del descenso del gradiente en la función de Wood.

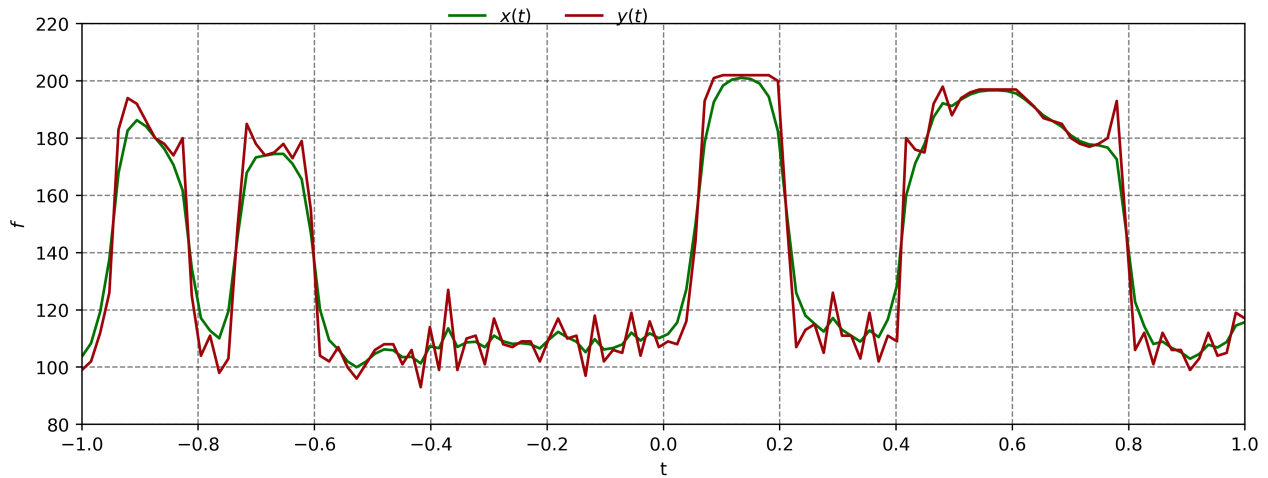
## 4.4. Función de suavizado

Para esta función se usaron dos tipos de vectores para el parámetro  $y$ . EL primer tipo de vector es uno cuyos elementos se encuentran contenidos en los primero 128 número del archivo [y.txt](#). El segundo tipo de vector es aquel que esta definido por la ecuación 25. En todos los casos el vector inicial es aleatori y esta definido por la ecuación 25. Se usaron como parámetros de  $\lambda$  los valores 1, 10 y 1000.

#### 4.4.1. Primer caso.

##### 4.4.1.1 Vector predefinido

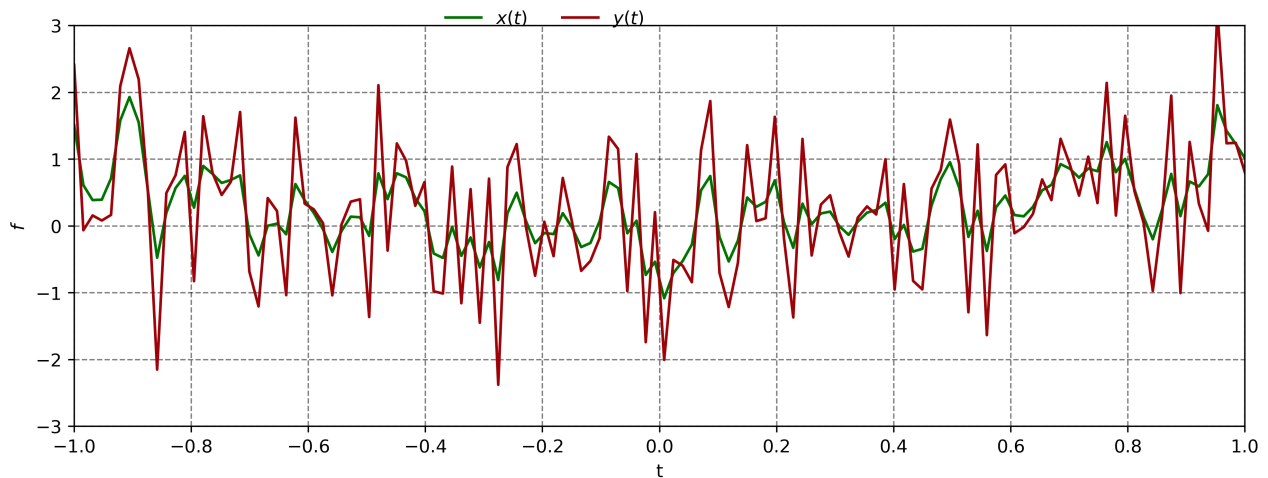
En la figura 7 se muestran los resultados de las funciones  $x(t)$  y el vector predefinido en el archivo `y.txt` para el caso  $\lambda = 1$ .



**Figura 7:** Comparación de los valores de entrada del archivo `y.txt` ( $y(t)$ ) con los valores obtenidos ( $x(t)$ ) al realizar el suavizado usando el método del descenso de gradiente.

##### 4.4.1.2 Vector aleatorio

En la figura 8 se muestran los resultados de las funciones  $x(t)$  y el vector aleatorio  $y$  para el caso  $\lambda = 1$ .

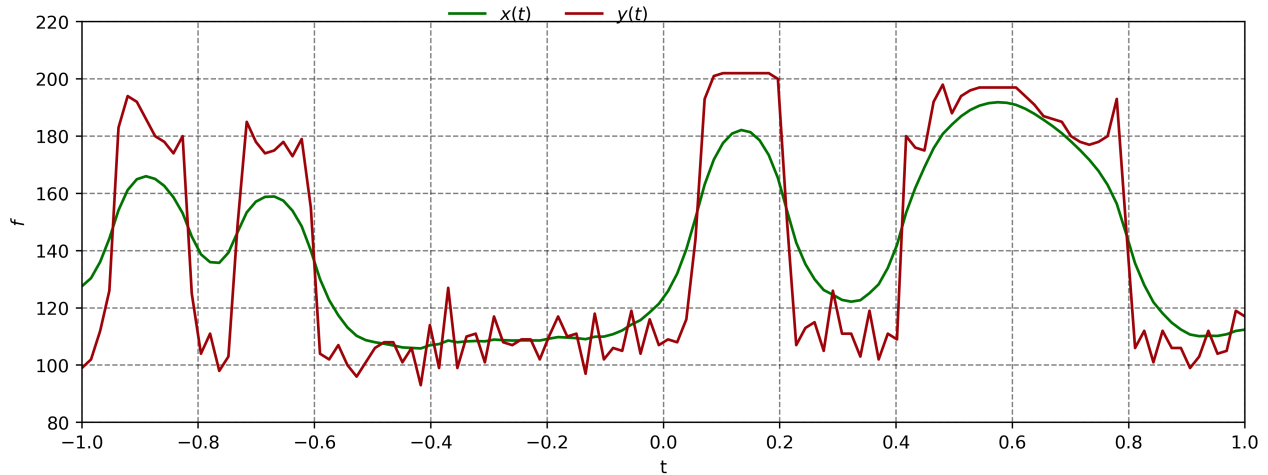


**Figura 8:** Comparación del vector aleatorio ( $y$ ) con los valores obtenidos ( $x$ ) al realizar el suavizado usando el método del descenso de gradiente.

#### 4.4.2. Segundo caso.

##### 4.4.2.1 Vector y predefinido

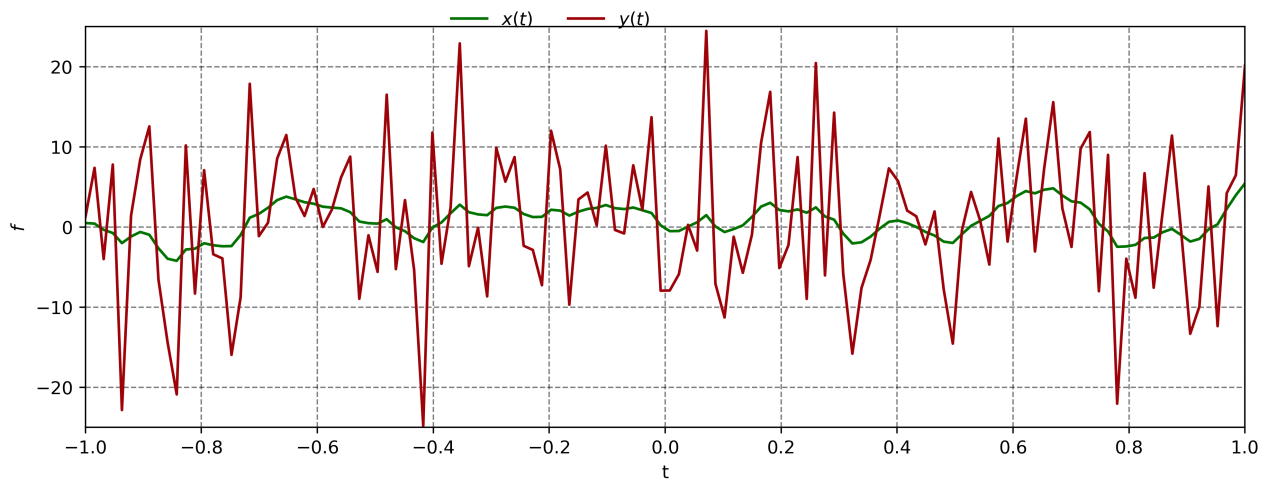
En la figura 9 se muestran los resultados de las funciones  $x(t)$  y el vector predefinido en el archivo `y.txt` para el caso  $\lambda = 10$ .



**Figura 9:** Comparación de los valores de entrada del archivo `y.txt` ( $y(t)$ ) con los valores obtenidos ( $x(t)$ ) al realizar el suavizado usando el método del descenso de gradiente.

#### 4.4.2.2 Vector aleatorio

En la figura 10 se muestran los resultados de las funciones  $x(t)$  y el vector aleatorio  $y$  para el caso  $\lambda = 10$ .



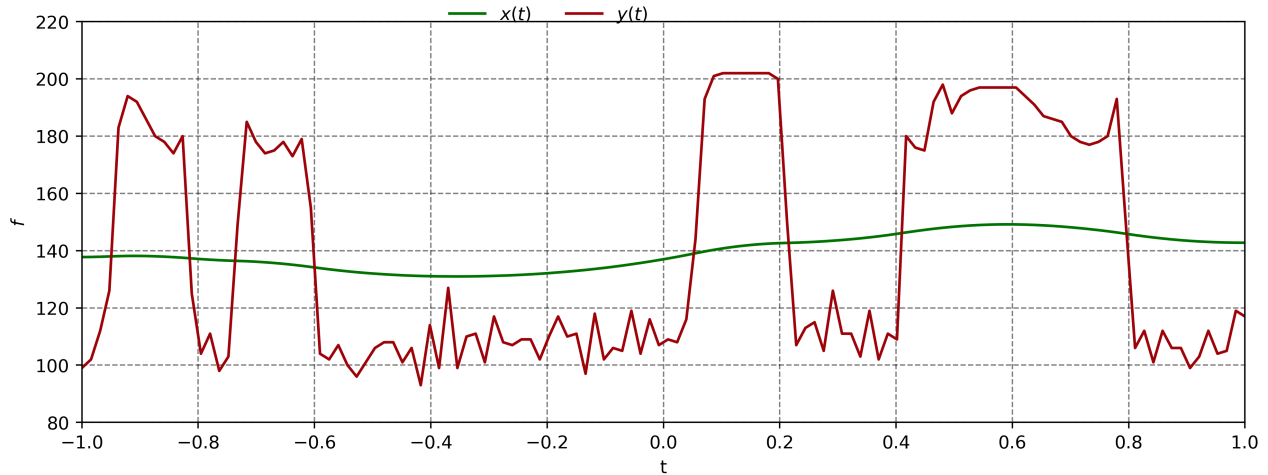
**Figura 10:** Comparación del vector aleatorio ( $y(t)$ ) con los valores obtenidos ( $x(t)$ ) al realizar el suavizado usando el método del descenso de gradiente.

#### 4.4.3. Tercer caso.

##### 4.4.3.1 Vector predefinido

En la figura 11 se muestran los resultados de las funciones  $x(t)$  y el vector predefinido en el archivo `y.txt` para el caso  $\lambda = 1000$ .

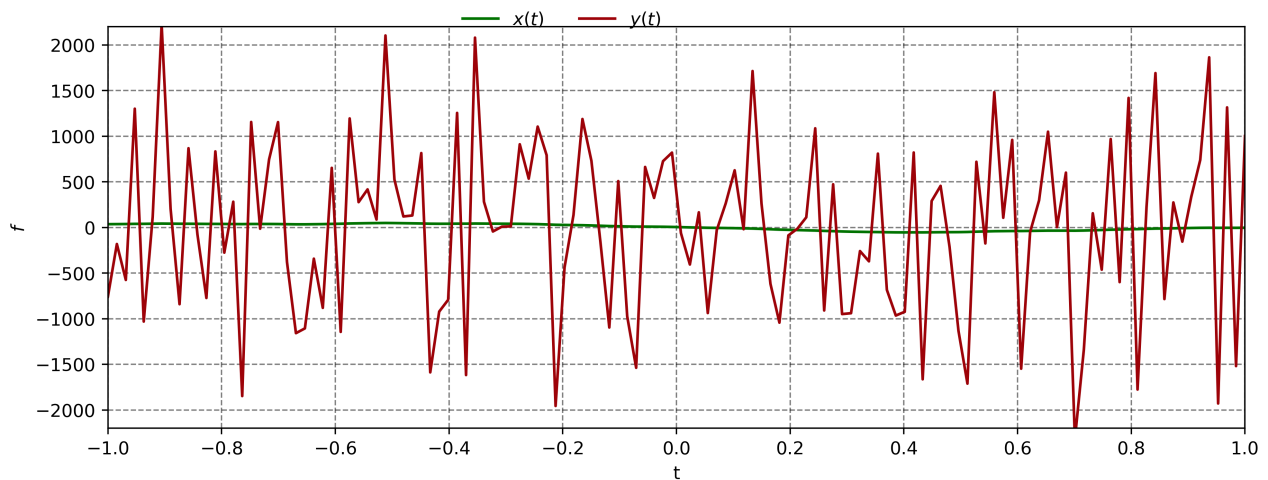




**Figura 11:** Comparación de los valores de entrada del archivo `y.txt` ( $y(t)$ ) con los valores obtenidos ( $x(t)$ ) al realizar el suavizado usando el método del descenso de gradiente.

#### 4.4.3.2 Vector y aleatorio

En la figura 12 se muestran los resultados de las funciones  $x(t)$  y el vector aleatorio  $y$  para el caso  $\lambda = 1000$ .



**Figura 12:** Comparación del vector aleatorio ( $y(t)$ ) con los valores obtenidos ( $x(t)$ ) al realizar el suavizado usando el método del descenso de gradiente.

## 5. Conclusiones

Con los resultados expuestos anteriormente se pueden obtener las siguientes conclusiones:

- El método de Newton realiza una cantidad menor de iteraciones a comparación del método del descenso del gradiente. Esto es debido a la elección continua del tamaño de paso que se calcula en cada iteración. Esto es reflejado en las tablas 2, 3, 4, 5, 10 y 11
- El método del descenso del gradiente logra obtener con mayor frecuencia un mínimo global. Si se obtuvo un mínimo local este no tiene un valor muy lejano al mínimo global. Esto puede verse en las tablas 6, 9, 8, 9, 12 y 13. En cambio el método de Newton tiende a detenerse en mínimos locales que estén cercanos al punto inicial dado.