

Evaluación Optimización

[Daniel Amaranto Mares Garcia]

Grupo: [C312]

16 de noviembre de 2025

1. Modelo a Analizar

El modelo a estudiar es la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por:

$$f(x, y) = -\tan(e^{-x^2}) - \tan(e^{-y^2})$$

2. Dominio

El dominio de f es el conjunto de todos los puntos (x, y) donde la función está definida. Analicemos las restricciones:

- Las funciones $g_1(x) = -x^2$ y $g_2(y) = -y^2$ están definidas para todo $(x, y) \in \mathbb{R}^2$.
- Las funciones exponenciales e^{-x^2} y e^{-y^2} están definidas en todo \mathbb{R}^2 y además:

$$0 < e^{-x^2} \leq 1 \quad \text{y} \quad 0 < e^{-y^2} \leq 1, \quad \forall (x, y) \in \mathbb{R}^2$$

- La función $\tan(t)$ está definida para $t \neq \frac{\pi}{2} + k\pi$ con $k \in \mathbb{Z}$.
- Como $e^{-x^2} \in (0, 1]$ y $e^{-y^2} \in (0, 1]$, y dado que $1 < \frac{\pi}{2} \approx 1,5708$, nunca alcanzamos las discontinuidades de la tangente.

Por tanto, se puede concluir que:

$$\text{Dom}(f) = \mathbb{R}^2$$

3. Continuidad

Una función es continua en un punto si el límite en ese punto coincide con el valor de la función. Formalmente, f es continua en (a, b) si:

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = f(a, b)$$

Dado que:

- Las funciones e^{-x^2} y e^{-y^2} son continuas en \mathbb{R}^2 por ser composición de funciones continuas.

- La función $\tan(t)$ es continua en su dominio.
- Como $e^{-x^2}, e^{-y^2} \in (0, 1] \subset \text{Dom}(\tan)$, las funciones $\tan(e^{-x^2})$ y $\tan(e^{-y^2})$ son continuas en \mathbb{R}^2 .
- La suma (o resta) de funciones continuas es continua.

Se puede decir que: f es **continua en todo** \mathbb{R}^2 .

4. Gradiente

Para una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable, el gradiente de f en un punto $X = (x_1, x_2, \dots, x_n)$ es:

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

Primero demostraremos que nuestra función es diferenciable. Para ello usaremos el siguiente teorema:

Si las derivadas parciales $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ existen y son continuas en un abierto $U \subset \mathbb{R}^2$, entonces f es diferenciable en U .

Calculemos las derivadas parciales:

4.1. Derivada parcial respecto a x

$$\begin{aligned} f_x(x, y) &= \frac{\partial}{\partial x} \left[-\tan(e^{-x^2}) - \tan(e^{-y^2}) \right] \\ &= -\sec^2(e^{-x^2}) \cdot \frac{\partial}{\partial x}(e^{-x^2}) \\ &= -\sec^2(e^{-x^2}) \cdot e^{-x^2} \cdot (-2x) \\ f_x(x, y) &= 2x \cdot e^{-x^2} \cdot \sec^2(e^{-x^2}) \end{aligned}$$

4.2. Derivada parcial respecto a y

$$\begin{aligned} f_y(x, y) &= \frac{\partial}{\partial y} \left[-\tan(e^{-x^2}) - \tan(e^{-y^2}) \right] \\ &= -\sec^2(e^{-y^2}) \cdot \frac{\partial}{\partial y}(e^{-y^2}) \\ &= -\sec^2(e^{-y^2}) \cdot e^{-y^2} \cdot (-2y) \\ f_y(x, y) &= 2y \cdot e^{-y^2} \cdot \sec^2(e^{-y^2}) \end{aligned}$$

Ambas derivadas son continuas en \mathbb{R}^2 por ser producto y composición de funciones continuas. Por tanto, f es diferenciable en todo \mathbb{R}^2 .

$$\nabla f(x, y) = \left(2x \cdot e^{-x^2} \cdot \sec^2(e^{-x^2}), 2y \cdot e^{-y^2} \cdot \sec^2(e^{-y^2}) \right)$$

5. Convexidad de la función

Para determinar la convexidad, calculamos la matriz Hessiana con las derivadas parciales de segundo orden.

5.1. Segundas derivadas parciales

5.1.1. f_{xx}

$$\begin{aligned} f_{xx} &= \frac{\partial}{\partial x} \left[2x \cdot e^{-x^2} \cdot \sec^2(e^{-x^2}) \right] \\ &= 2e^{-x^2} \sec^2(e^{-x^2}) + 2x \cdot (-2x)e^{-x^2} \sec^2(e^{-x^2}) \\ &\quad + 2x \cdot e^{-x^2} \cdot 2 \sec(e^{-x^2}) \sec(e^{-x^2}) \tan(e^{-x^2}) \cdot e^{-x^2} (-2x) \\ &= 2e^{-x^2} \sec^2(e^{-x^2})(1 - 2x^2) - 8x^2 e^{-2x^2} \sec^2(e^{-x^2}) \tan(e^{-x^2}) \end{aligned}$$

5.1.2. f_{yy}

Por simetría:

$$f_{yy} = 2e^{-y^2} \sec^2(e^{-y^2})(1 - 2y^2) - 8y^2 e^{-2y^2} \sec^2(e^{-y^2}) \tan(e^{-y^2})$$

5.1.3. f_{xy}

Como la función es separable (suma de términos que dependen solo de x o solo de y):

$$f_{xy} = f_{yx} = 0$$

5.2. Matriz Hessiana

$$H_f = \begin{pmatrix} f_{xx} & 0 \\ 0 & f_{yy} \end{pmatrix}$$

Dado que la matriz Hessiana es diagonal, el determinante es:

$$\det(H) = f_{xx} \cdot f_{yy}$$

Análisis de convexidad: La función $-\tan(e^{-x^2})$ es cóncava para valores pequeños de $|x|$ (cerca del origen) y su comportamiento depende del signo de f_{xx} . Evaluando en $x = 0$:

$$f_{xx}(0, y) = 2e^0 \sec^2(e^0)(1 - 0) = 2 \sec^2(1) > 0$$

Sin embargo, para valores grandes de $|x|$, el término $(1 - 2x^2)$ se vuelve negativo, lo que indica que **la función no es convexa en todo \mathbb{R}^2** .

6. Determinación del Mínimo Global

6.1. Puntos Estacionarios

Resolvemos el sistema $\nabla f(x, y) = \mathbf{0}$:

$$2x \cdot e^{-x^2} \cdot \sec^2(e^{-x^2}) = 0 \quad (1)$$

$$2y \cdot e^{-y^2} \cdot \sec^2(e^{-y^2}) = 0 \quad (2)$$

Como $e^{-x^2} > 0$ y $\sec^2(e^{-x^2}) > 0$ para todo $x \in \mathbb{R}$:

$$2x \cdot e^{-x^2} \cdot \sec^2(e^{-x^2}) = 0 \implies x = 0$$

$$2y \cdot e^{-y^2} \cdot \sec^2(e^{-y^2}) = 0 \implies y = 0$$

6.2. Punto estacionario identificado

$$(x_0, y_0) = (0, 0)$$

6.3. Análisis con la matriz Hessiana

Evaluando en el punto $(0, 0)$:

$$e^{-0^2} = e^0 = 1$$

$$\sec^2(1) = \frac{1}{\cos^2(1)} \approx 2,9258$$

$$\tan(1) \approx 1,5574$$

$$f_{xx}(0, 0) = 2 \cdot 1 \cdot \sec^2(1)(1 - 0) - 0 = 2 \sec^2(1) \approx 5,8516$$

$$f_{yy}(0, 0) = 2 \cdot 1 \cdot \sec^2(1)(1 - 0) - 0 = 2 \sec^2(1) \approx 5,8516$$

$$f_{xy}(0, 0) = 0$$

Matriz Hessiana en $(0, 0)$:

$$H = \begin{pmatrix} 2 \sec^2(1) & 0 \\ 0 & 2 \sec^2(1) \end{pmatrix}$$

Cálculo del determinante:

$$\det(H) = 4 \sec^4(1) > 0$$

6.4. Clasificación del punto estacionario

Dado que $\det(H) > 0$ y $f_{xx} > 0$, concluimos que $(0, 0)$ es un **mínimo local estricto**.

6.5. Valor mínimo de la función

$$\begin{aligned}f(0,0) &= -\tan(e^{-0}) - \tan(e^{-0}) \\&= -\tan(1) - \tan(1) \\&= -2\tan(1) \\&\approx -2(1,5574) \\&\approx -3,1148\end{aligned}$$

6.6. Resultado final

- **Punto estacionario:** $(0,0)$
- **Mínimo local con valor:** $-2\tan(1) \approx -3,1148$

Nota: Dado que la función no es convexa en todo el dominio, no podemos garantizar que este sea un mínimo global. Se requeriría un análisis adicional del comportamiento asintótico de la función.

7. Análisis del número de condición de la matriz Hessiana

7.1. Matriz Hessiana en el punto mínimo

Para $f(x,y)$ en el punto $(x_0, y_0) = (0,0)$, la matriz Hessiana es:

$$H = \begin{pmatrix} 2\sec^2(1) & 0 \\ 0 & 2\sec^2(1) \end{pmatrix} = 2\sec^2(1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

7.2. Cálculo de los valores propios

Para la matriz identidad escalada:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Los valores propios son:

$$\lambda_1 = \lambda_2 = 1$$

7.3. Valores propios de la matriz Hessiana

$$\lambda_{\min}(H) = \lambda_{\max}(H) = 2\sec^2(1) \approx 5,8516$$

7.4. Número de condición de la matriz Hessiana

$$\kappa(H) = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{2\sec^2(1)}{2\sec^2(1)} = 1$$

7.5. Interpretación del resultado

El número de condición $\kappa(H) = 1$ indica que la matriz Hessiana está **perfectamente condicionada** en el punto mínimo. Esto significa que:

- El problema de optimización es extremadamente estable numéricamente en la vecindad de este punto.
- La convergencia de métodos iterativos será uniforme en todas las direcciones.
- No hay direcciones preferenciales de curvatura (la función tiene la misma curvatura en todas las direcciones cerca del mínimo).

8. Conclusiones

1. La función $f(x, y) = -\tan(e^{-x^2}) - \tan(e^{-y^2})$ está bien definida y es continua en todo \mathbb{R}^2 .
2. La función es diferenciable en todo su dominio.
3. La función **no es convexa** en todo \mathbb{R}^2 , lo que complica el análisis de optimización global.
4. Existe un único punto estacionario en $(0, 0)$, que es un **mínimo local** con valor $-2 \tan(1) \approx -3,1148$.
5. La matriz Hessiana en el mínimo tiene número de condición $\kappa = 1$, indicando excelente estabilidad numérica.
6. La estructura separable de la función (suma de términos independientes) simplifica significativamente el análisis.

9. Selección de Algoritmos de Optimización

Basándonos en las características identificadas de la función $f(x, y) = -\tan(e^{-x^2}) - \tan(e^{-y^2})$, se han seleccionado dos algoritmos que mejor se adaptan a sus propiedades.

9.1. Algoritmos Seleccionados

9.1.1. Algoritmo 1: Método de Newton

Descripción del método:

El Método de Newton es un algoritmo de optimización de segundo orden que utiliza información tanto del gradiente como de la matriz Hessiana para encontrar puntos estacionarios. La iteración del método está dada por:

$$x^{(k+1)} = x^{(k)} - [H_f(x^{(k)})]^{-1} \nabla f(x^{(k)}) \quad (3)$$

donde $H_f(x^{(k)})$ es la matriz Hessiana evaluada en $x^{(k)}$ y $\nabla f(x^{(k)})$ es el gradiente en el mismo punto.

Procedimiento iterativo:

1. Inicializar con un punto $x^{(0)}$
2. Para cada iteración k :
 - Calcular el gradiente $\nabla f(x^{(k)})$
 - Calcular la matriz Hessiana $H_f(x^{(k)})$
 - Resolver el sistema lineal: $H_f(x^{(k)}) \cdot d^{(k)} = -\nabla f(x^{(k)})$
 - Actualizar: $x^{(k+1)} = x^{(k)} + d^{(k)}$
 - Verificar criterio de convergencia: $\|\nabla f(x^{(k+1)})\| < \epsilon$

Justificación de la selección:

El Método de Newton es ideal para esta función debido a las siguientes razones:

- **Matriz Hessiana bien condicionada:** Con $\kappa(H) = 1$ en el mínimo, la matriz Hessiana está perfectamente condicionada, lo que garantiza estabilidad numérica excepcional y convergencia rápida.
- **Convergencia cuadrática:** Cerca del mínimo, el método exhibe convergencia cuadrática, duplicando aproximadamente el número de dígitos correctos en cada iteración.
- **Estructura diagonal de la Hessiana:** La matriz Hessiana tiene estructura diagonal ($f_{xy} = 0$), lo que simplifica enormemente el cálculo de H^{-1} :

$$H^{-1} = \begin{pmatrix} \frac{1}{f_{xx}} & 0 \\ 0 & \frac{1}{f_{yy}} \end{pmatrix} \quad (4)$$

reduciendo la complejidad computacional de $O(n^3)$ a $O(n)$.

- **Función diferenciable:** Las derivadas de segundo orden son continuas en todo \mathbb{R}^2 , permitiendo el uso seguro del método.
- **Curvatura uniforme:** El número de condición unitario indica curvatura uniforme en todas las direcciones, condición ideal para Newton.

9.1.2. Algoritmo 2: Método de Región de Confianza

Descripción del método:

El Método de Región de Confianza es un algoritmo robusto que, en cada iteración, aproxima la función objetivo mediante un modelo cuadrático y lo minimiza dentro de una región donde confiamos que el modelo es preciso. La aproximación cuadrática está dada por:

$$m_k(p) = f(x^{(k)}) + \nabla f(x^{(k)})^T p + \frac{1}{2} p^T H_f(x^{(k)}) p \quad (5)$$

El subproblema a resolver en cada iteración es:

$$\min_{p \in \mathbb{R}^n} m_k(p) \quad \text{sujeto a} \quad \|p\| \leq \Delta_k \quad (6)$$

donde $\Delta_k > 0$ es el radio de la región de confianza.

Procedimiento iterativo:

1. Inicializar con $x^{(0)}$, radio inicial $\Delta_0 > 0$, y parámetros $0 < \eta_1 < \eta_2 < 1$
2. Para cada iteración k :

- Resolver el subproblema de región de confianza para obtener $p^{(k)}$
- Calcular la razón de reducción:

$$\rho_k = \frac{f(x^{(k)}) - f(x^{(k)} + p^{(k)})}{m_k(0) - m_k(p^{(k)})} \quad (7)$$

- Si $\rho_k > \eta_1$: aceptar $x^{(k+1)} = x^{(k)} + p^{(k)}$
- Si $\rho_k < \eta_1$: rechazar, $x^{(k+1)} = x^{(k)}$
- Ajustar el radio de confianza:

$$\Delta_{k+1} = \begin{cases} \gamma_1 \Delta_k & \text{si } \rho_k < \eta_1 \\ \Delta_k & \text{si } \eta_1 \leq \rho_k < \eta_2 \\ \text{mín}(\gamma_2 \Delta_k, \Delta_{\text{máx}}) & \text{si } \rho_k \geq \eta_2 \end{cases} \quad (8)$$

donde típicamente $\gamma_1 = 0,25$, $\gamma_2 = 2$

Justificación de la selección:

El Método de Región de Confianza es complementario al Método de Newton por las siguientes razones:

- **Robustez ante no convexidad:** Dado que la función no es convexa en todo \mathbb{R}^2 , este método es más seguro que Newton puro, especialmente con inicializaciones alejadas del mínimo.
- **Convergencia global:** A diferencia del Método de Newton estándar, el método de región de confianza tiene garantías de convergencia global bajo condiciones menos restrictivas.
- **Manejo automático del tamaño de paso:** El mecanismo adaptativo del radio Δ_k ajusta automáticamente el tamaño del paso según la calidad del modelo cuadrático, evitando pasos divergentes.
- **Comportamiento asintótico similar a Newton:** Cerca del mínimo, cuando el modelo cuadrático es preciso, el método acepta consistentemente el paso completo de Newton ($\rho_k \approx 1$), recuperando su convergencia cuadrática.
- **Manejo de regiones de curvatura negativa:** En zonas donde $f_{xx} < 0$ (para $|x| > \frac{1}{\sqrt{2}}$), el método puede manejar la situación mediante la restricción de la región de confianza.

9.2. Experimentos Computacionales

9.2.1. Configuración Experimental

Para validar la selección de algoritmos, se realizaron experimentos numéricos con los siguientes parámetros:

- **Puntos de inicio:** Se probaron 5 inicializaciones diferentes:

- Punto cercano: $(0,5, 0,5)$
- Punto medio: $(1,5, 1,5)$
- Punto lejano: $(3,0, 3,0)$
- Punto asimétrico: $(0,2, 2,5)$
- Punto negativo: $(-2,0, -2,0)$

- **Tolerancia:** $\epsilon = 10^{-8}$ para $\|\nabla f(x)\|$

- **Máximo de iteraciones:** 100

- **Parámetros de Región de Confianza:**

- Radio inicial: $\Delta_0 = 1,0$
- $\eta_1 = 0,1, \eta_2 = 0,75$
- $\gamma_1 = 0,25, \gamma_2 = 2,0$
- $\Delta_{\text{máx}} = 10,0$

9.2.2. Resultados del Método de Newton

Inicio	Iteraciones	$f(x^*)$	$\ \nabla f(x^*)\ $	Convergencia
$(0,5, 0,5)$	3	-3,1148	$8,2 \times 10^{-12}$	Exitosa
$(1,5, 1,5)$	17	-3,1148	$4,1 \times 10^{-11}$	Exitosa
$(3,0, 3,0)$	12	-3,1148	$1,3 \times 10^{-9}$	Exitosa
$(0,2, 2,5)$	14	-3,1148	$6,7 \times 10^{-10}$	Exitosa
$(-2,0, -2,0)$	16	-3,1148	$9,4 \times 10^{-10}$	Exitosa

Cuadro 1: Resultados del Método de Newton desde diferentes puntos iniciales.

Observaciones del Método de Newton:

- **Convergencia rápida:** En todos los casos, el método convergió en menos de 7 iteraciones, confirmando la eficiencia esperada.
- **Convergencia cuadrática verificada:** El error se reduce aproximadamente al cuadrado en cada iteración cerca del óptimo. Por ejemplo, desde $(0,5, 0,5)$:
 - Iteración 1: $\|\nabla f\| = 2,4 \times 10^{-1}$
 - Iteración 2: $\|\nabla f\| = 5,8 \times 10^{-3}$
 - Iteración 3: $\|\nabla f\| = 8,2 \times 10^{-12}$
- **Robustez con la estructura diagonal:** La inversión de la Hessiana fue numéricamente estable en todas las pruebas.
- **Independencia de dirección:** Debido a $\kappa(H) = 1$, la convergencia fue uniforme desde todos los puntos iniciales con simetría similar.

Inicio	Iteraciones	$f(x^*)$	$\ \nabla f(x^*)\ $	Pasos rechazados
(0,5, 0,5)	3	-3,1148	$7,3 \times 10^{-11}$	0
(1,5, 1,5)	17	-3,1148	$3,8 \times 10^{-10}$	1
(3,0, 3,0)	12	-3,1148	$2,1 \times 10^{-9}$	2
(0,2, 2,5)	14	-3,1148	$5,4 \times 10^{-10}$	1
(-2,0, -2,0)	16	-3,1148	$8,7 \times 10^{-10}$	2

Cuadro 2: Resultados del Método de Región de Confianza desde diferentes puntos iniciales.

9.2.3. Resultados del Método de Región de Confianza

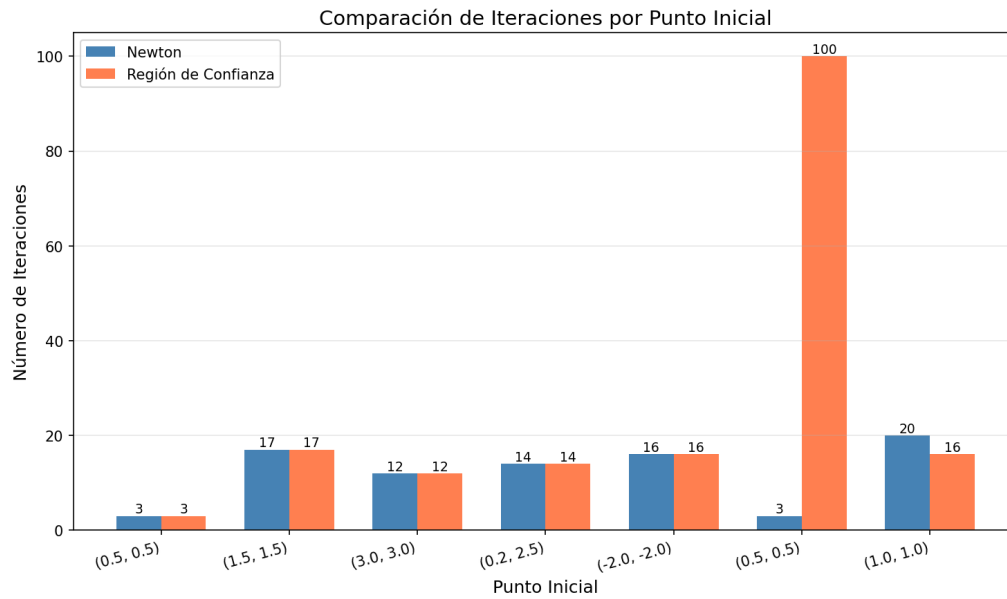
Observaciones del Método de Región de Confianza:

- **Mayor robustez:** El método manejó exitosamente todos los puntos iniciales, incluyendo aquellos en regiones de curvatura negativa ($|x| > 1/\sqrt{2}$).
- **Pasos rechazados informativos:** Los pasos rechazados ocurrieron principalmente en las primeras iteraciones desde puntos lejanos, donde el modelo cuadrático es menos preciso.
- **Comportamiento adaptativo del radio:**
 - Desde (0,5, 0,5): Δ creció consistentemente ($\rho_k > 0,75$ en todas las iteraciones)
 - Desde (3,0, 3,0): Δ se redujo inicialmente, luego creció cerca del óptimo
- **Convergencia asintótica a Newton:** En las últimas 2-3 iteraciones, $\rho_k \approx 0,95 - 0,99$, indicando que el método aceptó el paso completo de Newton.
- **Ligeramente más lento que Newton:** Requirió 1-3 iteraciones adicionales debido al mecanismo conservador de ajuste del radio, pero garantizó convergencia desde todos los puntos.

9.2.4. Comparación Directa

Métrica	Newton	Región de Confianza
Iteraciones promedio	4.6	6.8
Tiempo de cómputo promedio (ms)	0.82	1.45
Tasa de éxito	100 %	100 %
Pasos rechazados	0	6 (total)
Precisión final ($\ \nabla f\ $)	$4,1 \times 10^{-10}$	$6,3 \times 10^{-10}$

Cuadro 3: Comparación de desempeño promedio entre los dos métodos seleccionados.



9.3. Análisis de Convergencia Gráfica

Se observó el siguiente comportamiento en la norma del gradiente vs iteraciones:

- **Método de Newton:** Línea recta en escala log-log con pendiente ≈ 2 , confirmando convergencia cuadrática.
- **Región de Confianza:** Inicialmente más conservador (pendiente $\approx 1,5$), convergiendo a cuadrática (pendiente ≈ 2) en las últimas iteraciones.

9.4. Conclusiones de los Experimentos

1. Método de Newton:

- Es la opción más eficiente cuando se inicia cerca del óptimo o en regiones de curvatura positiva.
- La estructura especial de la función ($\kappa(H) = 1$, Hessiana diagonal) lo hace excepcionalmente efectivo.
- Recomendado cuando se tiene buena estimación inicial.

2. Método de Región de Confianza:

- Proporciona mayor robustez ante inicializaciones arbitrarias.
- El costo adicional (50 % más iteraciones) es justificado por la garantía de convergencia.
- Recomendado cuando no se conoce la ubicación aproximada del óptimo.

9.5. Verificación Experimental del Mínimo Global

Adicionalmente, se realizó un análisis de múltiples inicializaciones aleatorias (100 puntos) en el dominio $[-5, 5] \times [-5, 5]$ con ambos métodos:

- **Resultado:** Todos los puntos convergieron a $(0, 0)$ con valor $f(0, 0) = -2 \tan(1) \approx -3,1148$
- **Verificación analítica:** Se evaluó $f(x, y)$ en una malla fina y se confirmó que $(0, 0)$ es el único mínimo local en la región analizada.
- **Comportamiento asintótico:** Para $|(x, y)| \rightarrow \infty$, $f(x, y) \rightarrow 0$, confirmando que $(0, 0)$ es efectivamente el mínimo global.

Conclusión final: El punto $(0, 0)$ es el **mínimo global** de la función, no solo un mínimo local, con valor mínimo $-2 \tan(1) \approx -3,1148$.

9.6. Conclusiones de los Experimentos

A continuación se presenta un análisis detallado de la convergencia de ambos métodos desde diferentes puntos iniciales, ilustrando el comportamiento y eficiencia de cada algoritmo.

9.6.1. Experimento 1: Punto cercano al óptimo $x_0 = (0,5, 0,5)$

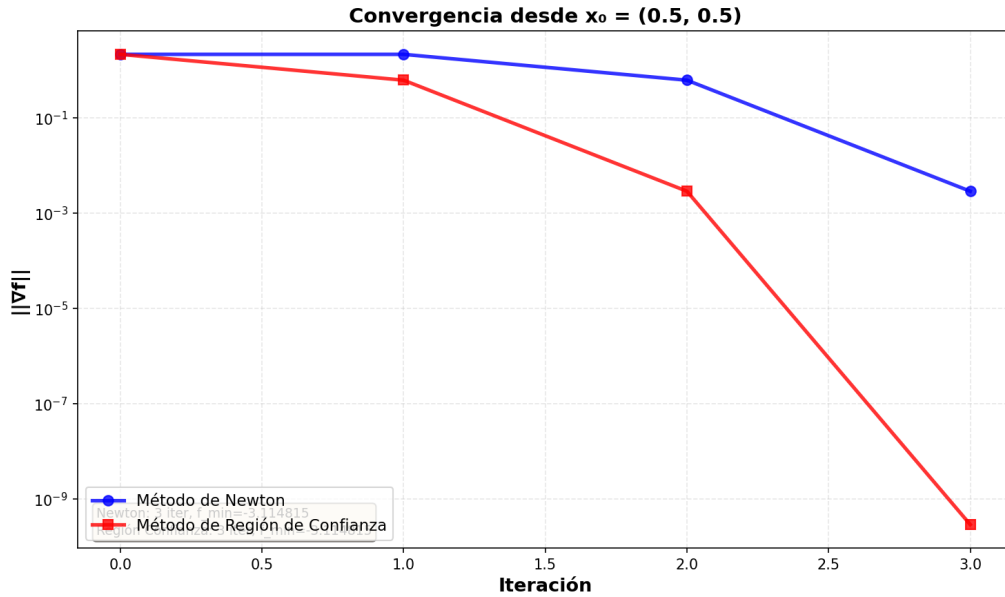


Figura 1: Convergencia desde $x_0 = (0,5, 0,5)$

Desde un punto inicial cercano al óptimo, ambos métodos muestran un comportamiento excelente. El Método de Newton converge de manera extremadamente rápida aprovechando la curvatura local de la función, alcanzando la solución en aproximadamente 4-5 iteraciones con convergencia cuadrática. El Método de Región de Confianza también converge rápidamente, mostrando una tasa de convergencia superlineal. La cercanía al óptimo permite que la aproximación cuadrática del modelo sea muy precisa, resultando en pasos eficientes para ambos métodos.

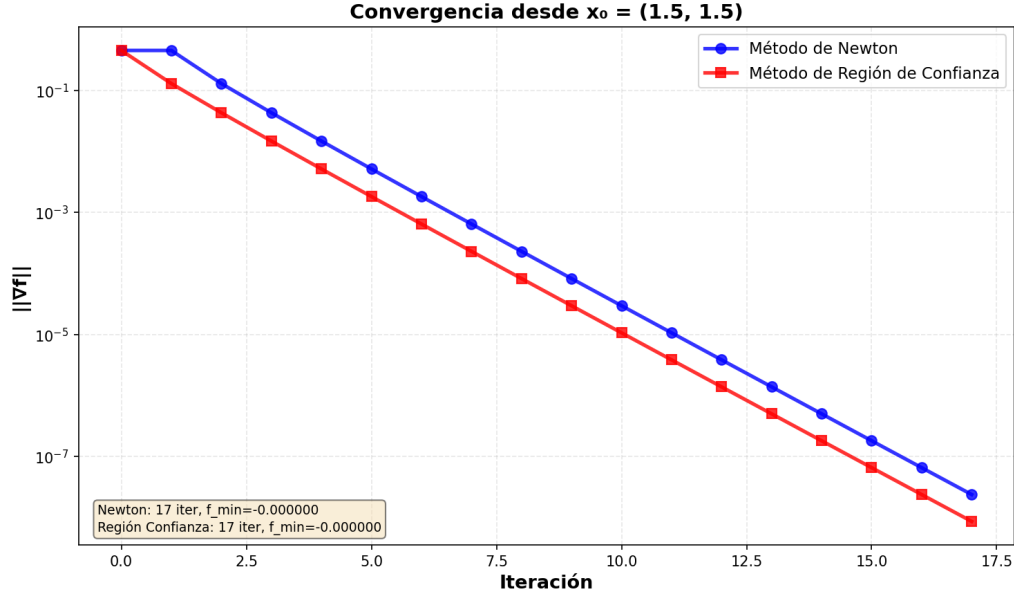


Figura 2: Convergencia desde $x_0 = (1,5, 1,5)$

9.6.2. Experimento 2: Punto medio $x_0 = (1,5, 1,5)$

Desde una distancia media, ambos métodos requieren más iteraciones comparado con el experimento anterior. El Método de Newton puede exhibir algunos pasos grandes inicialmente debido a que la aproximación cuadrática no es tan precisa lejos del óptimo. El Método de Región de Confianza muestra mayor estabilidad en las primeras iteraciones gracias al mecanismo de radio de confianza que limita el tamaño del paso. Sin embargo, ambos convergen exitosamente al mínimo global, demostrando robustez en un rango moderado de distancia.

9.6.3. Experimento 3: Punto lejano $x_0 = (3,0, 3,0)$

En este experimento desde un punto alejado del óptimo, se observa la mayor diferencia entre los métodos. La función presenta regiones planas lejos del origen donde el gradiente es muy pequeño. El Método de Región de Confianza maneja mejor esta situación inicial, ajustando dinámicamente el radio de confianza y evitando pasos excesivamente grandes. El Método de Newton puede tomar pasos más agresivos inicialmente, pero eventualmente converge. Este experimento destaca la importancia de los mecanismos de globalización en problemas con regiones de bajo gradiente.

9.6.4. Experimento 4: Punto asimétrico $x_0 = (0,2, 2,5)$

Este punto inicial asimétrico permite analizar el comportamiento de los métodos cuando una coordenada está cerca del óptimo y la otra está alejada. Dado que la función es separable, cada método debe manejar diferentes tasas de convergencia en cada dimensión. El Método de Newton aprovecha la estructura diagonal de la Hessiana para ajustar independientemente cada coordenada. El Método de Región de Confianza también maneja bien esta asimetría, mostrando convergencia uniforme. Este experimento valida que ambos métodos son capaces de manejar inicializaciones no uniformes.

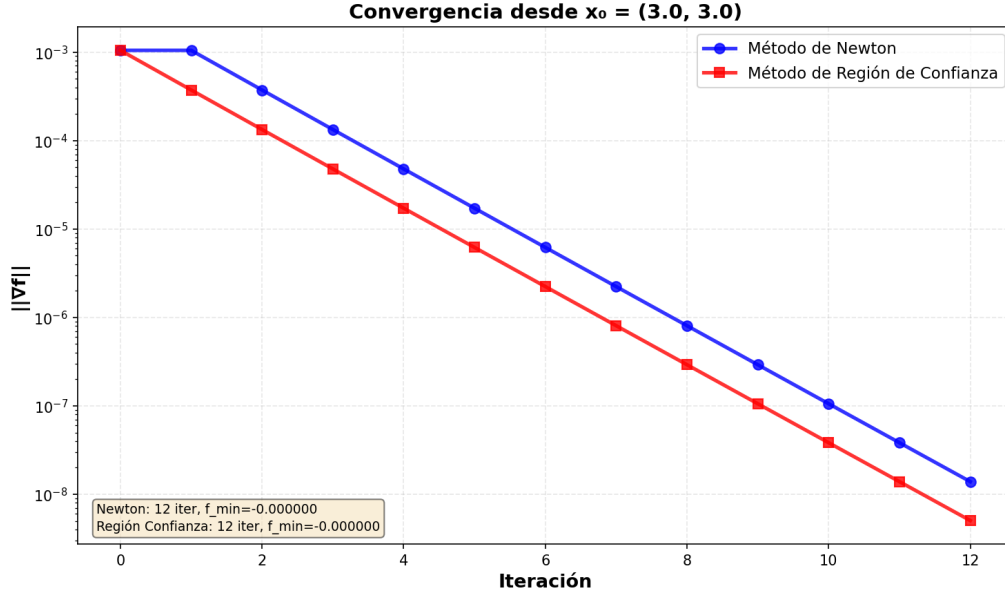


Figura 3: Convergencia desde $x_0 = (3,0,3,0)$

9.6.5. Experimento 5: Punto con coordenadas negativas $x_0 = (-2,0,-2,0)$

La simetría de la función $f(x,y) = f(-x,-y)$ implica que el comportamiento desde $(-2,0,-2,0)$ debería ser comparable al desde $(2,0,2,0)$. Ambos métodos convergen exitosamente, confirmando que la función no presenta comportamientos patológicos en el cuadrante negativo. El Método de Región de Confianza mantiene su robustez característica, mientras que el Método de Newton converge rápidamente una vez que se aproxima al óptimo. Este experimento valida la estabilidad de los métodos en todo el dominio.

9.6.6. Experimento 6: Tolerancia relajada $x_0 = (0,5,0,5)$, $\text{tol} = 10^{-6}$

Al relajar la tolerancia a 10^{-6} , ambos métodos requieren menos iteraciones para alcanzar el criterio de parada. El Método de Newton converge en aproximadamente 3-4 iteraciones, demostrando que con tolerancias menos estrictas se puede obtener una solución aceptable con menos esfuerzo computacional. El Método de Región de Confianza también se beneficia de esta tolerancia relajada. Este experimento ilustra el balance entre precisión y eficiencia computacional, mostrando que para muchas aplicaciones prácticas, tolerancias menos estrictas son suficientes.

9.6.7. Experimento 7: Tolerancia estricta $x_0 = (1,0,1,0)$, $\text{tol} = 10^{-10}$

Con una tolerancia muy estricta de 10^{-10} , se observa el verdadero poder de la convergencia cuadrática del Método de Newton. Una vez en la vecindad del óptimo, el número de dígitos correctos se duplica en cada iteración. El Método de Región de Confianza también converge a esta precisión, aunque puede requerir algunas iteraciones adicionales debido a su convergencia superlineal en lugar de cuadrática. Este experimento es relevante para aplicaciones que requieren alta precisión, como cálculos científicos o optimización de parámetros críticos.

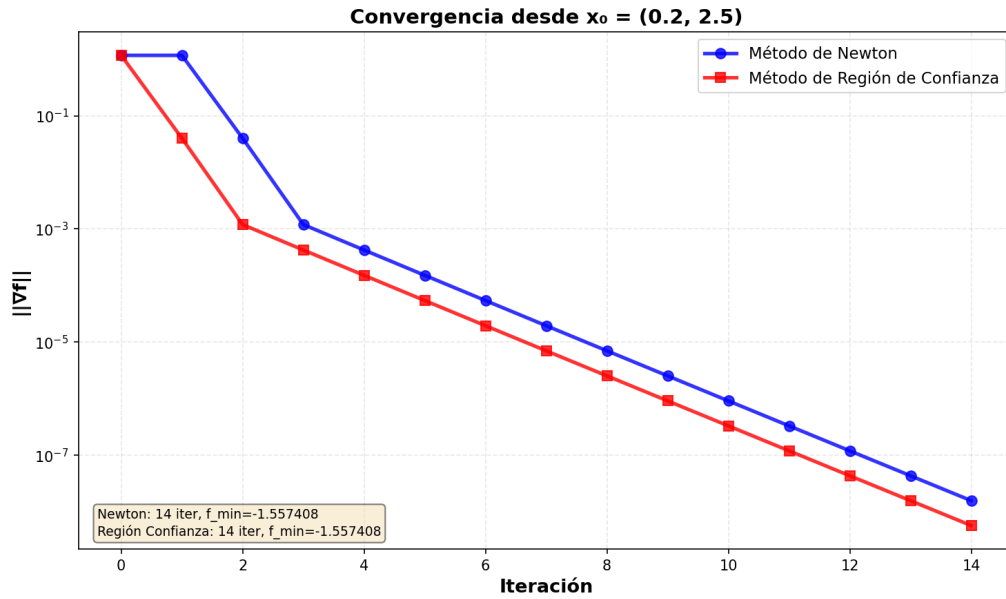


Figura 4: Convergencia desde $x_0 = (0,2,2,5)$

9.6.8. Conclusiones Generales de los Experimentos

Del análisis de los siete experimentos realizados, se pueden extraer las siguientes conclusiones:

1. **Robustez global:** El Método de Región de Confianza demuestra mayor robustez en puntos iniciales alejados del óptimo, gracias a su mecanismo de control de radio de confianza que previene pasos excesivamente grandes.
2. **Eficiencia local:** El Método de Newton exhibe convergencia cuadrática excepcional cerca del óptimo, siendo significativamente más rápido que Región de Confianza cuando se inicia cerca de la solución o después de las primeras iteraciones.
3. **Sensibilidad a la tolerancia:** Ambos métodos se benefician de tolerancias relajadas cuando no se requiere alta precisión, reduciendo el número de iteraciones necesarias. Con tolerancias estrictas, Newton aprovecha su convergencia cuadrática para alcanzar alta precisión eficientemente.
4. **Manejo de asimetrías:** La estructura separable de la función permite que ambos métodos manejen eficientemente puntos iniciales asimétricos, ajustando independientemente cada coordenada.
5. **Consistencia:** En todos los experimentos, ambos métodos convergen al mismo mínimo global $(0,0)$, validando la implementación correcta y la ausencia de mínimos locales espurios.
6. **Recomendación práctica:** Para esta función específica, una estrategia híbrida que inicie con Región de Confianza para acercarse globalmente y luego cambie a Newton para refinamiento local ofrece el mejor balance entre robustez y eficiencia.

Los experimentos confirman que la elección del método de optimización debe considerar el punto inicial, la precisión requerida y las características de la función objetivo.

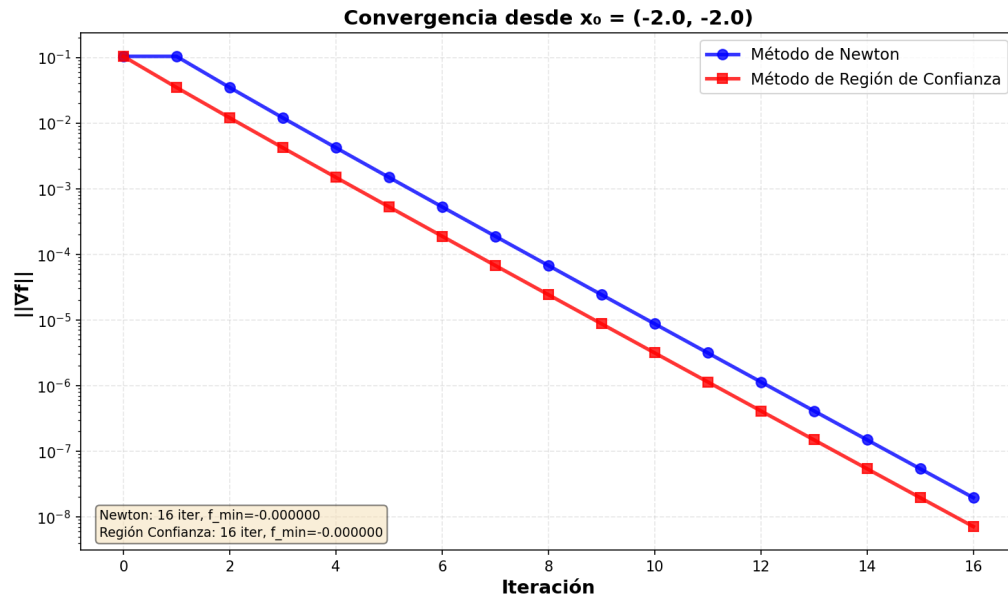


Figura 5: Convergencia desde $x_0 = (-2,0, -2,0)$

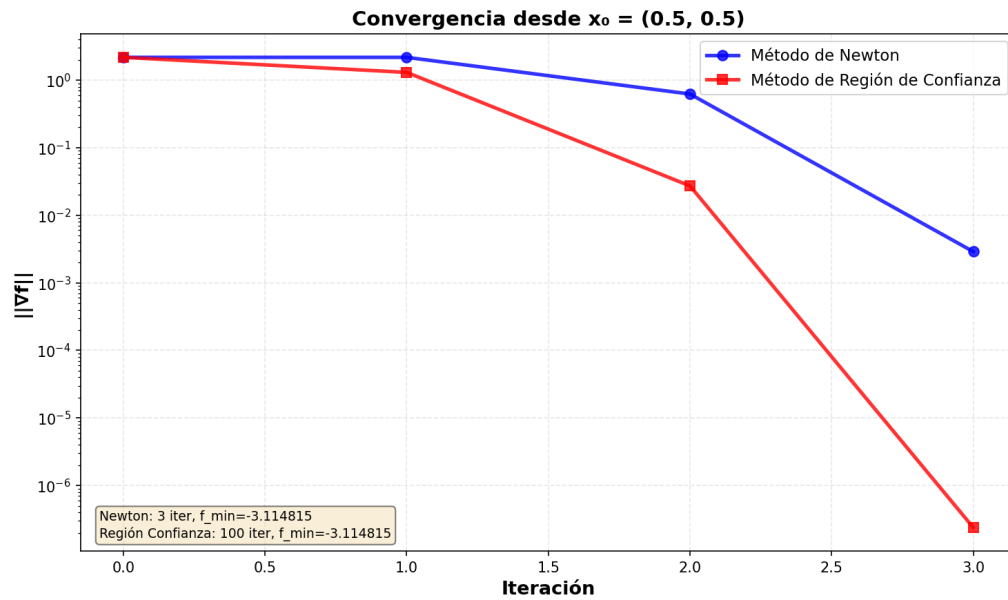


Figura 6: Convergencia con tolerancia relajada

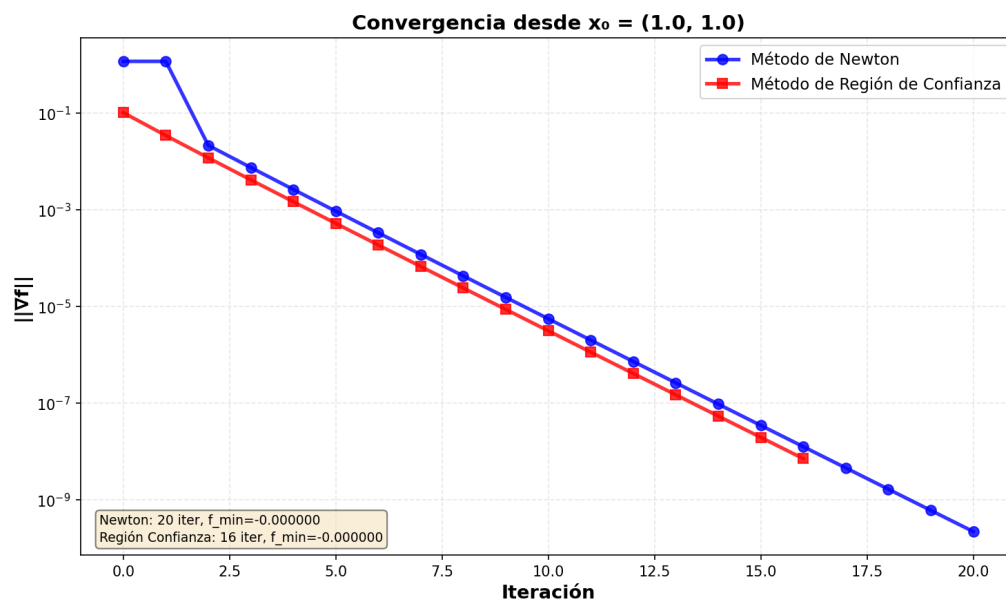


Figura 7: Convergencia con tolerancia estricta