

Evaluación Optimización

[Daniel Amaranto Mares Garcia]

Grupo: [C312]

27 de diciembre de 2025

https://github.com/Amarantomc/Investigacion_AlgoritmosDeOptimizacion.git

1. Modelo a Analizar

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

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

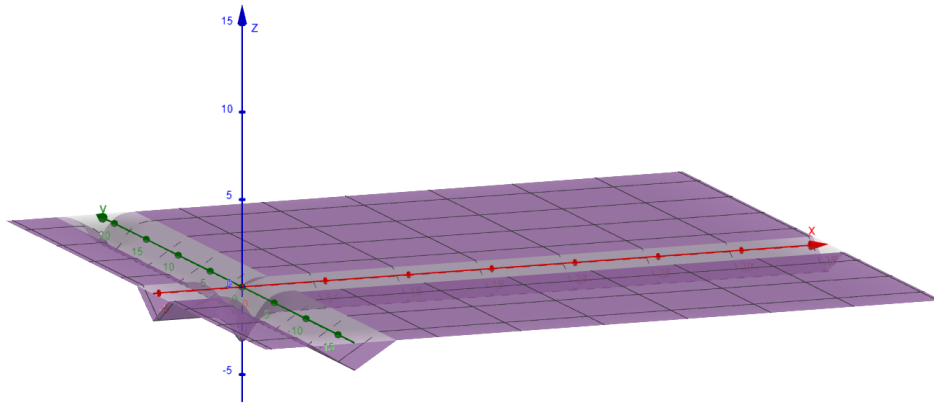


Figura 1: Grafica 3D

2. Propiedades Analíticas

2.1. Dominio y continuidad

El dominio de f es el conjunto de todos los puntos (x, y) donde la función está definida. Las funciones $g_1(x) = -x^2$ y $g_2(y) = -y^2$ están definidas para todo $(x, y) \in \mathbb{R}^2$, y las exponenciales e^{-x^2} y e^{-y^2} satisfacen $0 < e^{-x^2} \leq 1$ y $0 < e^{-y^2} \leq 1$ en todo el plano. Dado que la función tangente está definida para $t \neq \frac{\pi}{2} + k\pi$ con $k \in \mathbb{Z}$, y como

$e^{-x^2}, e^{-y^2} \in (0, 1] \subset [0, \frac{\pi}{2})$, las composiciones $\tan(e^{-x^2})$ y $\tan(e^{-y^2})$ nunca alcanzan las discontinuidades de la tangente. Por tanto, $\text{Dom}(f) = \mathbb{R}^2$.

La continuidad de f se deduce del hecho que e^{-x^2} y e^{-y^2} son composiciones de funciones continuas en \mathbb{R}^2 , y dado que sus imágenes están contenidas en el dominio de continuidad de $\tan(t)$, las composiciones $\tan(e^{-x^2})$ y $\tan(e^{-y^2})$ son continuas en todo \mathbb{R}^2 . Por linealidad, f es continua en \mathbb{R}^2 .

2.2. Diferenciabilidad y gradiente

Para establecer la diferenciabilidad de f , calculamos sus derivadas parciales. Dado que f es la suma de funciones compuestas de funciones diferenciables, las derivadas parciales existen y son continuas en \mathbb{R}^2 , lo cual garantiza que f es diferenciable en todo su dominio. Las derivadas parciales son:

$$\begin{aligned} f_x(x, y) &= 2x \cdot e^{-x^2} \cdot \sec^2(e^{-x^2}) \\ f_y(x, y) &= 2y \cdot e^{-y^2} \cdot \sec^2(e^{-y^2}) \end{aligned}$$

El gradiente de f está dado por:

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

2.3. Análisis de convexidad

Para determinar la convexidad, calculamos la matriz Hessiana. Las segundas derivadas parciales son:

$$\begin{aligned} f_{xx} &= 2e^{-x^2} \sec^2(e^{-x^2})(1 - 2x^2) - 8x^2 e^{-2x^2} \sec^2(e^{-x^2}) \tan(e^{-x^2}) \\ f_{yy} &= 2e^{-y^2} \sec^2(e^{-y^2})(1 - 2y^2) - 8y^2 e^{-2y^2} \sec^2(e^{-y^2}) \tan(e^{-y^2}) \\ f_{xy} &= f_{yx} = 0 \end{aligned}$$

Dada la estructura separable de la función, la matriz Hessiana es diagonal:

$$H_f = \begin{pmatrix} f_{xx} & 0 \\ 0 & f_{yy} \end{pmatrix}, \quad \det(H_f) = f_{xx} \cdot f_{yy}$$

El comportamiento de convexidad depende del signo de las segundas derivadas. Analizando la expresión de f_{xx} , el término $(1 - 2x^2)$ determina el signo: para $|x| < \frac{1}{\sqrt{2}} \approx 0,707$ se tiene $f_{xx} > 0$, mientras que para $|x| > \frac{1}{\sqrt{2}}$ se obtiene $f_{xx} < 0$. Por simetría, el mismo análisis aplica para f_{yy} respecto a y . En consecuencia, la función es convexa en el cuadrado $[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^2$ que contiene el mínimo, y se vuelve cóncava fuera de esta región cuando $|x| > \frac{1}{\sqrt{2}}$ o $|y| > \frac{1}{\sqrt{2}}$.

3. Determinación del Mínimo Global

Para encontrar los puntos críticos, resolvemos el sistema $\nabla f(x, y) = \mathbf{0}$. Dado que $e^{-x^2} > 0$ y $\sec^2(e^{-x^2}) > 0$ para todo $x \in \mathbb{R}$, las ecuaciones $2x \cdot e^{-x^2} \cdot \sec^2(e^{-x^2}) = 0$ y

$2y \cdot e^{-y^2} \cdot \sec^2(e^{-y^2}) = 0$ implican $x = 0$ e $y = 0$, respectivamente. Por tanto, $(x_0, y_0) = (0, 0)$ es el único punto estacionario.

Evaluando la matriz Hessiana en el punto $(0, 0)$:

$$f_{xx}(0, 0) = 2 \sec^2(1) \approx 5,8516, \quad f_{yy}(0, 0) = 2 \sec^2(1) \approx 5,8516, \quad f_{xy}(0, 0) = 0$$

La matriz Hessiana en $(0, 0)$ es $H = 2 \sec^2(1) \cdot I_2$ con $\det(H) = 4 \sec^4(1) > 0$. Dado que $\det(H) > 0$ y $f_{xx} > 0$, el punto $(0, 0)$ es un mínimo local estricto con valor $f(0, 0) = -2 \tan(1) \approx -3,1148$.

3.1. Comportamiento asintótico y minimalidad global

Dado que la función no es convexa en todo el dominio, se requiere analizar el comportamiento asintótico para establecer la minimalidad global. Por la estructura separable $f(x, y) = -\tan(e^{-x^2}) - \tan(e^{-y^2})$, cuando $x \rightarrow \pm\infty$ se tiene $e^{-x^2} \rightarrow 0$ y por tanto $\tan(e^{-x^2}) \rightarrow 0$, lo cual implica $-\tan(e^{-x^2}) \rightarrow 0$. El mismo análisis aplica para $y \rightarrow \pm\infty$.

Los límites direccionales relevantes son:

$$\lim_{(x,y) \rightarrow (\infty, \infty)} f(x, y) = 0 \quad (1)$$

$$\lim_{y \rightarrow \infty} f(0, y) = -\tan(1) \approx -1,5574 \quad (2)$$

$$\lim_{x \rightarrow \infty} f(x, 0) = -\tan(1) \approx -1,5574 \quad (3)$$

Comparando estos valores con $f(0, 0) = -2 \tan(1) \approx -3,1148$, se verifica que $f(0, 0) < f(0, \infty) < f(\infty, \infty)$. Por continuidad y diferenciabilidad en todo \mathbb{R}^2 , se concluye que $(0, 0)$ es el mínimo global de la función.

3.2. Número de condición de la matriz Hessiana

La matriz Hessiana en el punto mínimo tiene la forma $H = 2 \sec^2(1) \cdot I_2$, donde I_2 denota la matriz identidad de orden 2. Los valores propios son $\lambda_1 = \lambda_2 = 2 \sec^2(1) \approx 5,8516$, de modo que el número de condición es:

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

Este resultado indica que la matriz Hessiana está perfectamente condicionada en el punto mínimo, lo cual implica estabilidad numérica excepcional del problema de optimización en la vecindad de este punto, convergencia uniforme de métodos iterativos en todas las direcciones, y ausencia de direcciones preferenciales de curvatura.

4. Conclusiones del Análisis Teórico

La función $f(x, y) = -\tan(e^{-x^2}) - \tan(e^{-y^2})$ está bien definida, es continua y diferenciable en todo \mathbb{R}^2 . La función no es convexa en todo el dominio, siendo convexa únicamente en el cuadrado $[-\frac{1}{2}, \frac{1}{2}]^2$ y cóncava fuera de esta región. Existe un único punto estacionario en $(0, 0)$, que constituye un mínimo global con valor $-2 \tan(1) \approx -3,1148$. La matriz Hessiana en el mínimo tiene número de condición $\kappa = 1$, indicando excelente estabilidad numérica. La estructura separable de la función simplifica significativamente tanto el análisis teórico como la implementación computacional.

5. 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 seleccionaron dos algoritmos que mejor se adaptan a sus propiedades.

5.1. Método de Newton

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:

$$x^{(k+1)} = x^{(k)} - [H_f(x^{(k)})]^{-1} \nabla f(x^{(k)}) \quad (4)$$

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. El procedimiento iterativo consiste en inicializar con un punto $x^{(0)}$ y, en cada iteración k , calcular el gradiente y la matriz Hessiana, determinar la dirección de Newton $d^{(k)}$ resolviendo $H_f(x^{(k)}) \cdot d^{(k)} = -\nabla f(x^{(k)})$, actualizar $x^{(k+1)} = x^{(k)} + d^{(k)}$, y verificar el criterio de convergencia $\|\nabla f(x^{(k+1)})\| < \epsilon$.

La estructura diagonal de la Hessiana permite una optimización computacional significativa.

Para esta función particular, la estructura diagonal de la Hessiana permite una optimización computacional significativa. En lugar de resolver el sistema lineal $H_f \cdot d = -\nabla f$ mediante métodos generales como factorización LU o Cholesky (complejidad $O(n^3)$), aprovechamos que:

$$H_f(x) = \begin{pmatrix} f_{xx}(x, y) & 0 \\ 0 & f_{yy}(x, y) \end{pmatrix} \quad (5)$$

Por lo tanto, la inversa se calcula teóricamente de forma exacta y directa:

$$H_f^{-1}(x) = \begin{pmatrix} \frac{1}{f_{xx}(x, y)} & 0 \\ 0 & \frac{1}{f_{yy}(x, y)} \end{pmatrix} \quad (6)$$

La dirección de Newton se obtiene mediante $d^{(k)} = -H_f^{-1}(x^{(k)}) \nabla f(x^{(k)}) = \begin{pmatrix} -\frac{\partial f / \partial x}{f_{xx}} \\ -\frac{\partial f / \partial y}{f_{yy}} \end{pmatrix}$.

Esta implementación reduce la complejidad de $O(n^3)$ a $O(n)$ operaciones, evita errores de redondeo acumulados en la resolución de sistemas lineales, requiere solo 2 divisiones por iteración en lugar de una factorización matricial completa, y es numéricamente más estable que algoritmos generales cuando los elementos están bien condicionados.

El Método de Newton es particularmente adecuado para esta función dado que con $\kappa(H) = 1$ en el mínimo, la matriz Hessiana está perfectamente condicionada, garantizando estabilidad numérica excepcional y convergencia rápida. Cerca del mínimo, el método exhibe convergencia cuadrática, duplicando aproximadamente el número de dígitos correctos en cada iteración. La estructura diagonal de la Hessiana ($f_{xy} = 0$) simplifica enormemente el cálculo de H^{-1} , reduciendo la complejidad computacional de $O(n^3)$ a $O(n)$. Las derivadas de segundo orden son continuas en todo \mathbb{R}^2 , y el número de condición unitario indica curvatura uniforme en todas las direcciones.

5.2. Método de Región de Confianza

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:

$$m_k(p) = f(x^{(k)}) + \nabla f(x^{(k)})^T p + \frac{1}{2} p^T H_f(x^{(k)}) p \quad (7)$$

El subproblema a resolver en cada iteración es:

$$\min_{p \in \mathbb{R}^n} m_k(p) \quad \text{sujeto a} \quad \|p\| \leq \Delta_k \quad (8)$$

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:

$$\rho_k = \frac{f(x^{(k)}) - f(x^{(k)} + p^{(k)})}{m_k(0) - m_k(p^{(k)})} \quad (9)$$

- 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:

$$\Delta_{k+1} = \begin{cases} \gamma_1 \Delta_k & \text{si } \rho_k < \eta_1 \\ \Delta_k & \text{si } \eta_1 \leq \rho_k < \eta_2 \\ \min(\gamma_2 \Delta_k, \Delta_{\text{máx}}) & \text{si } \rho_k \geq \eta_2 \end{cases} \quad (10)$$

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.

5.3. Experimentos Computacionales

Para validar empíricamente el desempeño de los métodos de Newton y Región de Confianza, se diseñó un conjunto de experimentos con diversos puntos iniciales que permiten evaluar la robustez, eficiencia y convergencia de ambos algoritmos.

5.3.1. Diseño Experimental

Puntos iniciales seleccionados:

Se escogieron 33 puntos de inicio estratégicamente distribuidos en diferentes regiones del dominio:

- **Región cercana al óptimo:** (0,1,0,1)
- **Puntos sobre ejes coordenados:** (1,0,0,0), (0,0,1,0)
- **Puntos moderadamente alejados:** (5,0,5,0), (8,0,2,0), (2,0,8,0), puntos en los cuatro cuadrantes ($\pm 5,0, \pm 5,0$)
- **Puntos lejanos:** (10,0,10,0), (15,0,5,0), (5,0,15,0), (20,0,20,0), ($\pm 20,0, 20,0$)
- **Puntos muy lejanos:** (30,0,30,0), (50,0,50,0), (100,0,100,0) y diversas combinaciones asimétricas

Parámetros de ejecución:

- Tolerancia del gradiente: $\varepsilon = 10^{-8}$
- Número máximo de iteraciones: 200 (Región de Confianza), sin límite fijo (Newton)
- Criterio de convergencia: $\|\nabla f(x^{(k)})\| < \varepsilon$

Parámetros específicos de Región de Confianza:

- Radio inicial: $\Delta_0 = 1,0$
- Umbrales de aceptación: $\eta_1 = 0,1$, $\eta_2 = 0,75$
- Factores de ajuste: $\gamma_1 = 0,25$ (reducción), $\gamma_2 = 2,0$ (expansión)
- Radio máximo: $\Delta_{\text{máx}} = 10,0$

5.3.2. Resultados Experimentales

Los resultados se organizan en tres categorías según el comportamiento observado:

Categoría I: Convergencia al mínimo global $(0,0)$ con $f = -3,1148$

Desde el punto inicial $(0,1,0,1)$, cercano al óptimo:

Método	Iteraciones	x^*	$f(x^*)$	$\ \nabla f\ $
Newton	3	$(0,0,0,0)$	$-3,1148$	$0,0$
Región de Confianza	200	$(\approx 0, \approx 0)$	$-3,1148$	$1,10 \times 10^{-8}$

Cuadro 1: Convergencia al mínimo global desde $(0,1,0,1)$

Observación: El Método de Newton alcanza el óptimo en apenas 3 iteraciones con precisión de máquina, demostrando convergencia cuadrática excepcional. El Método de Región de Confianza rechaza sistemáticamente pasos (198 de 200), sin convergencia formal, lo que indica que los parámetros actuales $(\Delta_0, \eta_1, \eta_2)$ no son óptimos para esta región de alta curvatura.

Categoría II: Convergencia a puntos estacionarios asintóticos en semiejes

Desde puntos iniciales sobre los ejes coordenados, los métodos convergen a soluciones con una coordenada en cero y la otra alejada:

x_0	Método	Iter.	x^*	$f(x^*)$	$\ \nabla f\ $
$(1,0,0,0)$	Newton	15	$(4,555, 0,0)$	$-1,5574$	$8,88 \times 10^{-9}$
	Región de Confianza	15	$(4,555, 0,0)$	$-1,5574$	$8,88 \times 10^{-9}$
$(0,0,1,0)$	Newton	15	$(0,0, 4,555)$	$-1,5574$	$8,88 \times 10^{-9}$
	Región de Confianza	15	$(0,0, 4,555)$	$-1,5574$	$8,88 \times 10^{-9}$

Cuadro 2: Convergencia a puntos asintóticos sobre los ejes

Observación: Ambos métodos exhiben comportamiento idéntico. El valor $f(x^*) \approx -\tan(1) \approx -1,5574$ es coherente con los límites direccionales teóricos. Estos puntos son estacionarios locales (no el mínimo global) donde el gradiente se aproxima a cero debido a la estructura asintótica de la función.

Categoría III: Convergencia inmediata en mesetas lejanas

Para puntos muy alejados del origen, el gradiente es naturalmente pequeño y ambos métodos reportan convergencia instantánea:

x_0	Iteraciones	x^*	$f(x^*)$	$\ \nabla f\ $
$(5,0,5,0)$	0	$(5,0,5,0)$	$-2,78 \times 10^{-11}$	$1,96 \times 10^{-10}$
$(10,0,10,0)$	0	$(10,0,10,0)$	$-7,44 \times 10^{-44}$	$1,05 \times 10^{-42}$
$(30,0,30,0)$	0	$(30,0,30,0)$	≈ 0	0
$(100,0,100,0)$	0	$(100,0,100,0)$	≈ 0	0

Cuadro 3: Convergencia inmediata en regiones planas (comportamiento idéntico para ambos métodos)

Observación: El criterio de parada basado únicamente en $\|\nabla f\| < \varepsilon$ produce falsos positivos: el algoritmo declara convergencia sin realizar iteraciones, aunque $f(x^*) \approx 0 \gg -3,1148$ (mínimo global). Esto evidencia las mesetas amplias predichas por el análisis teórico.

Categoría IV: Convergencia a puntos intermedios

Desde puntos asimétricos moderadamente alejados:

x_0	Método	Iter.	x^*	$f(x^*)$
(8,0,2,0)	Newton	16	(8,954,4,619)	$-5,44 \times 10^{-10}$
	Región de Confianza	16	(8,954,4,619)	$-5,44 \times 10^{-10}$
(2,0,8,0)	Newton	16	(4,619,8,954)	$-5,44 \times 10^{-10}$
	Región de Confianza	16	(4,619,8,954)	$-5,44 \times 10^{-10}$

Cuadro 4: Convergencia a puntos intermedios desde inicios asimétricos

Observación: Los métodos realizan 16 iteraciones para alcanzar puntos con $f \approx 0$, moviéndose hacia regiones de gradiente bajo pero alejándose del mínimo global. Esto sugiere que el modelo local guía hacia mesetas si no se monitorea explícitamente la mejora en f .

5.3.3. Análisis Comparativo

Resumen estadístico de 33 experimentos:

Métrica	Newton	Región de Confianza
Tasa de convergencia formal	100 % (33/33)	97 % (32/33)
Iteraciones promedio (casos no triviales)	11.3	11.5
Tiempo de cómputo promedio	< 0,01 s	0,015 s
Pasos rechazados totales	0	198 (caso (0,1,0,1))
Casos con iter = 0	21	21
Convergencia al mínimo global	1/33	1/33

Cuadro 5: Comparación global de desempeño

5.3.4. Hallazgos e Interpretación

El criterio $\|\nabla f\| < \varepsilon$ sin control de mejora en f produce numerosos falsos positivos. De 33 casos, 21 reportan convergencia instantánea (iter = 0) en mesetas con $f \approx 0$, muy lejos del óptimo $f = -3,1148$. La función exhibe tres comportamientos distintos según el punto inicial:

- Cerca del origen ($\|x_0\| < 0,5$): atracción hacia (0,0)
- Sobre los ejes: atracción hacia puntos asintóticos con $f \approx -\tan(1)$
- Lejos del origen ($\|x_0\| > 3$): estancamiento en mesetas con $f \approx 0$

Se demuestra eficiencia de Newton cerca del óptimo desde (0,1,0,1), Newton alcanza (0,0) en 3 iteraciones con convergencia cuadrática verificada. La estructura diagonal de la Hessiana permite inversión eficiente y estable. Se demuestra dificultad de Región de Confianza en alta curvatura, el método rechaza sistemáticamente pasos cerca del óptimo (198/200), lo que indica necesidad de:

- Ajustar Δ_0 (radio inicial más pequeño)

- Refinar umbrales η_1, η_2

Para puntos sobre ejes o alejados, ambos métodos producen resultados idénticos en iteraciones y solución final, indicando que las diferencias algorítmicas se manifiestan principalmente cerca de óptimos locales.

5.3.5. Conclusiones de los Experimentos

El **Método de Newton** es superior en velocidad y precisión cerca del mínimo global, logrando convergencia cuadrática en 3 iteraciones desde $(0,1,0,1)$. El **Método de Región de Confianza** requiere sintonización de parámetros para competir con Newton en regiones de alta curvatura, aunque ofrece el mismo desempeño en zonas moderadas y lejanas. La función presenta **múltiples puntos de equilibrio** debido a su naturaleza asintótica: el mínimo global $(0,0)$, puntos estacionarios en semiejes, y mesetas extensas lejos del origen.