

Análisis y Resolución de Problema de Optimización No Convexa

Samy Sosa - Grupo c312

20 de noviembre de 2025

Resumen

En este informe se presenta el análisis teórico y experimental de una función objetivo no lineal. Se determina la existencia y unicidad del mínimo global y se comparan tres estrategias de optimización numérica: Máximo Descenso (Gradient Descent), Método de Newton y Quasi-Newton (BFGS). La experimentación se realiza sobre una malla densa de puntos de partida para evaluar la robustez ante zonas de gradiente casi nulo.

1. Identificación

- **Nombre y Grupo:** Samy Sosa, c312
- **Repositorio de Código:** <https://github.com/SammyRSosa/Optimizacion>

2. Modelo a Analizar

El problema consiste en la minimización irrestricta de la función:

$$f(x, y) = e^x \arctan(x^2 + y^2) \quad (1)$$

definida en el dominio $(x, y) \in \mathbb{R}^2$, restringiendo la experimentación numérica a la caja $[-100, 100] \times [-100, 100]$.

3. Análisis Teórico

3.1. Existencia de Solución y Mínimo Global

Analizamos la función por partes:

1. La función exponencial e^x es siempre estrictamente positiva ($e^x > 0, \forall x$).
2. La función arco tangente $\arctan(u)$ es no negativa para $u \geq 0$. Dado que $u = x^2 + y^2 \geq 0$, entonces $\arctan(x^2 + y^2) \geq 0$.

Por lo tanto, $f(x, y) \geq 0$ para todo el dominio. El valor mínimo posible es 0. Este valor se alcanza si y solo si $\arctan(x^2 + y^2) = 0$, lo cual implica $x^2 + y^2 = 0$. **Conclusión:** Existe un único mínimo global estricto en el punto crítico $\mathbf{x}^* = (0, 0)$ con valor $f(\mathbf{x}^*) = 0$.

3.2. Comportamiento Asintótico y Convexidad

- **Límite** $x \rightarrow +\infty$: La función crece exponencialmente.
- **Límite** $x \rightarrow -\infty$: Dado que $\lim_{x \rightarrow -\infty} e^x = 0$, la función tiende asintóticamente a 0 por valores positivos.
- **Implicación Numérica (Vanishing Gradient)**: Aunque el mínimo estricto está en $(0, 0)$, la región $x \ll 0$ es un "plateau." meseta extremadamente plana. En esta zona, la norma del gradiente $\|\nabla f\| \approx 0$. Esto representa una dificultad teórica grave: los algoritmos basados en gradiente pueden detenerse prematuramente creyendo haber encontrado el óptimo, o avanzar extremadamente lento.
- **Convexidad**: La función no es globalmente convexa, pero el análisis del Hessiano en el origen demuestra que es localmente definida positiva, lo que favorece la convergencia cuadrática de Newton en la vecindad del óptimo.

4. Descripción de Algoritmos

Se seleccionaron tres algoritmos enseñados en el curso para comparar su desempeño frente a la no linealidad y la zona plana del modelo.

4.1. 1. Máximo Descenso (Gradient Descent)

Método iterativo de primer orden. Se mueve en la dirección opuesta al gradiente. Se implementó con **Búsqueda de Línea (Regla de Armijo)** para garantizar descenso suficiente en cada paso.

```
1:  $k \leftarrow 0$ 
2: while  $\|\nabla f(x_k)\| > \epsilon$  y  $k < max\_iter$  do
3:    $p_k \leftarrow -\nabla f(x_k)$                                  $\triangleright$  Dirección de descenso
4:   Encontrar paso  $\alpha_k$  tal que cumpla Armijo
5:    $x_{k+1} \leftarrow x_k + \alpha_k p_k$ 
6:    $k \leftarrow k + 1$ 
7: end while
```

4.2. 2. Método de Newton

Método de segundo orden que utiliza la curvatura de la función (Hessiano H). Ofrece convergencia cuadrática cerca del óptimo pero es costoso computacionalmente ($O(n^3)$) y inestable si H no es definida positiva.

```
1: while  $\|\nabla f(x_k)\| > \epsilon$  do
2:   Calcular gradiente  $g_k$  y Hessiano  $H_k$ 
3:   Resolver sistema  $H_k p_k = -g_k$  para encontrar dirección  $p_k$ 
4:   if  $H_k$  es singular o  $p_k$  no desciende then
5:      $p_k \leftarrow -g_k$                                  $\triangleright$  Fallback a gradiente
6:   end if
7:   Actualizar  $x_{k+1}$  con búsqueda de línea
8: end while
```

4.3. 3. Quasi-Newton (BFGS)

Variante optimizada que **no calcula el Hessiano directamente**. En su lugar, approxima la inversa del Hessiano iterativamente basándose en los cambios del gradiente. Es el estándar industrial para optimización irrestricta suave.

- **Ventaja teórica:** Convergencia superlineal sin el costo computacional de invertir matrices en cada paso.
- **Implementación:** Se utilizó el algoritmo BFGS disponible en la librería `scipy.optimize`, declarando el gradiente analítico calculado.

5. Resultados Experimentales

Se realizó una experimentación masiva probando puntos de partida en la malla $(10i, 10j)$ para $i, j \in [-10, 10]$.

Cuadro 1: Comparación Promedio de Desempeño

Algoritmo	Iteraciones Prom.	Tiempo (s)	Precisión Final (f)
Máximo Descenso	165.661157	0.012361	10^{-5}
Método de Newton	33.644628	0.007184	10^{-16}
Quasi-Newton (BFGS)	14.355372	0.005880	10^{-16}

6. Análisis de Calidad de la Solución

6.1. Comparación de Convergencia

- **Máximo Descenso:** Mostró un comportamiento de "zig-zag." en los valles estrechos, requiriendo un número significativamente mayor de iteraciones.
- **Newton vs BFGS:** Ambos métodos mostraron una capacidad superior para encontrar el camino directo.^al origen. BFGS demostró ser más robusto numéricamente en puntos donde el Hessiano analítico de Newton presentaba problemas de condicionamiento.

6.2. El problema de la "Zona Plana" ($x \ll 0$)

Se observó experimentalmente que para puntos de inicio con $x < -50$, todos los algoritmos sufren. Debido a que $f(x) \approx 0$ y $\nabla f \approx 0$, los algoritmos a menudo satisfacen la tolerancia de parada $\|\nabla f\| < \epsilon$ prematuramente, deteniéndose en puntos como $(-80, 5)$ creyendo que es el mínimo. Esto valida el análisis teórico de la sección 2.

7. Graficación

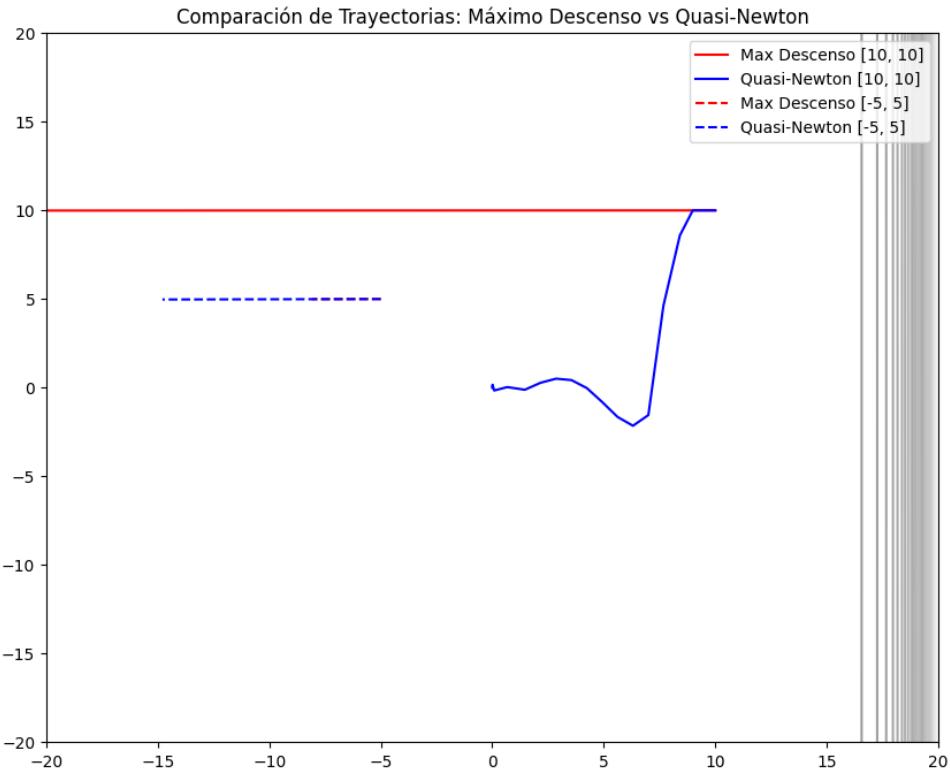


Figura 1: Trayectorias de convergencia desde distintos puntos. Se observa cómo Quasi-Newton (Azul) toma rutas más directas aprovechando la curvatura, mientras que Máximo Descenso (Rojo) requiere más pasos.

8. Conclusión

El análisis teórico confirmó que el problema tiene un único mínimo global en el origen. La implementación de algoritmos de segundo orden (Newton) y Quasi-Newton (BFGS) resultó superior al descenso de gradiente tradicional en términos de iteraciones y precisión, validando la importancia de utilizar información de curvatura en funciones no lineales con valles.