Análisis Teórico y Numérico de la Optimización de $f(x, y) = \dfrac{\arctan(x^2 + y^2)}{1 + x^2}$

Repositorio del proyecto: `https://github.com/agustin030902/Tarea-Evaluativa-MO`

Resumen

Este trabajo presenta un análisis teórico y numérico de la función $f(x, y) = \dfrac{\arctan(x^2 + y^2)}{1 + x^2}$ desde la perspectiva de la Programación No Lineal. Se demuestra la existencia y unicidad de un mínimo global, se estudian las propiedades de diferenciabilidad y convexidad local, y se analizan varios métodos clásicos de optimización, incluyendo Gradiente Descendente y el Método de Newton. Los resultados numéricos confirman las predicciones teóricas sobre convergencia y eficiencia de los algoritmos considerados.

1. Introducción

La optimización de funciones no lineales de varias variables es un problema central en matemáticas aplicadas, con aplicaciones en ingeniería, ciencias de datos y física. En este trabajo se estudia la función

$$f(x, y) = \frac{\arctan(x^2 + y^2)}{1 + x^2},$$

la cual presenta un comportamiento no trivial debido a su dependencia diferenciada en las variables x y y.

El objetivo principal es analizar sus propiedades analíticas relevantes para la optimización y evaluar el desempeño de distintos algoritmos numéricos en la búsqueda de su mínimo global.

2. Propiedades básicas de la función

La función f está bien definida para todo $(x, y) \in \mathbb{R}^2$, ya que arctan es continua en $\mathbb{R}$ y $1 + x^2 > 0$ para todo x. Además, al ser composición y cociente de funciones suaves, se tiene que $f \in C^\infty(\mathbb{R}^2)$.

Notamos que

$$\arctan(x^2 + y^2) \geq 0,$$

por lo que $f(x, y) \geq 0$ para todo (x, y). La función no es radial ni convexa globalmente, pero como se verá más adelante, presenta convexidad local en un entorno del origen.

3. Existencia y unicidad del mínimo global

Proposición 1. *La función f posee un único mínimo global en el punto* $(0, 0)$.

Demostración. Para todo $(x, y) \in \mathbb{R}^2$ se cumple

$$f(x, y) = \frac{\arctan(x^2 + y^2)}{1 + x^2} \geq 0.$$

Además,

$$f(0,0) = \frac{\arctan(0)}{1} = 0.$$

Si $(x,y) \neq (0,0)$, entonces $x^2+y^2 > 0$ y por tanto $\arctan(x^2+y^2) > 0$, lo que implica $f(x,y) > 0$. De aquí se concluye que $(0,0)$ es el mínimo global y es único. □

4. Cálculo del gradiente y puntos críticos

Calculamos las derivadas parciales de f:

$$\frac{\partial f}{\partial x} = \frac{2x(1+x^2)\frac{1}{1+(x^2+y^2)^2} - 2x\arctan(x^2+y^2)}{(1+x^2)^2},$$

$$\frac{\partial f}{\partial y} = \frac{2y}{(1+x^2)\,(1+(x^2+y^2)^2)}.$$

Proposición 2. *El único punto crítico de f es el origen* $(0,0)$.

Demostración. De la ecuación $\frac{\partial f}{\partial y} = 0$ se obtiene directamente $y = 0$. Sustituyendo en la ecuación $\frac{\partial f}{\partial x} = 0$ se llega a

$$x\left(\frac{1+x^2}{1+x^4} - \arctan(x^2)\right) = 0.$$

La expresión entre paréntesis no se anula para $x \neq 0$, por lo que la única solución es $x = 0$. Por tanto, el único punto crítico es $(0,0)$. □

5. Análisis de segundo orden

Proposición 3. *La Hessiana de f en el origen es definida positiva.*

Demostración. Usando el desarrollo en serie de Taylor de $\arctan(t)$ alrededor de $t = 0$,

$$\arctan(t) = t + O(t^3),$$

se obtiene, para (x,y) cercano al origen,

$$f(x,y) = x^2 + y^2 + O(\|(x,y)\|^4).$$

De aquí se deduce que

$$\nabla^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

la cual es definida positiva. □

Esto garantiza que el origen es un mínimo estricto y que la función es convexa en un entorno del punto óptimo, aunque no necesariamente de forma global.

6. Métodos numéricos de optimización

En esta sección se describen los métodos numéricos empleados para la minimización de la función

$$f(x,y) = \frac{\arctan(x^2+y^2)}{1+x^2}.$$

El objetivo es explicar el fundamento teórico de cada algoritmo y justificar su uso en el contexto del problema estudiado.

6.1. Método de Gradiente Descendente

El método de Gradiente Descendente es un algoritmo iterativo de primer orden utilizado para minimizar funciones diferenciables. La idea central del método consiste en moverse, desde un punto inicial, en la dirección opuesta al gradiente de la función, ya que dicha dirección corresponde al máximo descenso local.

La iteración del método está dada por

$$x_{k+1} = x_k - \alpha \nabla f(x_k),$$

donde $\alpha > 0$ es un tamaño de paso fijo.

Bajo hipótesis de suavidad de la función y para valores suficientemente pequeños de α, el método converge localmente hacia un punto crítico. En el caso de la función estudiada, la existencia de un único mínimo global y la convexidad local alrededor del origen garantizan la convergencia del método cuando el punto inicial no se encuentra excesivamente alejado del óptimo.

6.2. Gradiente Descendente con paso adaptativo

Una desventaja del Gradiente Descendente con paso fijo es la sensibilidad del algoritmo a la elección del parámetro α. Para mitigar este problema, se utiliza una versión con paso adaptativo, donde el tamaño del paso se ajusta en cada iteración.

En este enfoque, el valor de α_k se selecciona de manera que se garantice una disminución suficiente del valor de la función objetivo, típicamente mediante criterios de búsqueda en línea como la condición de Armijo.

Este método mantiene la simplicidad del Gradiente Descendente, pero mejora considerablemente su estabilidad y robustez, especialmente cuando el punto inicial se encuentra lejos del mínimo. Por esta razón, resulta adecuado para la optimización global de la función considerada.

6.3. Método de Newton

El Método de Newton es un algoritmo de segundo orden que utiliza información tanto del gradiente como de la Hessiana de la función objetivo. Su iteración está dada por

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k).$$

Este método puede interpretarse como la minimización exacta de la aproximación cuadrática de la función en cada iteración. Si la Hessiana es definida positiva y el punto inicial se encuentra suficientemente cerca del mínimo, el Método de Newton presenta convergencia cuadrática.

En el problema estudiado, la Hessiana de f en el origen es definida positiva, lo cual justifica teóricamente la rápida convergencia observada del método en un entorno del mínimo global.

7. Resultados numéricos

Los experimentos numéricos se realizaron utilizando distintos puntos iniciales, tanto cercanos como alejados del origen, y un criterio de parada basado en

$$\|\nabla f(x_k)\| < 10^{-6}.$$

En el caso del Gradiente Descendente con paso fijo, se observó que la convergencia depende de manera crítica de la elección del tamaño de paso. Para valores pequeños del parámetro, el método converge al mínimo global $(0, 0)$, aunque con una velocidad de convergencia lineal. Sin embargo, para valores de paso demasiado grandes, el algoritmo puede oscilar o incluso divergir, especialmente cuando el punto inicial se encuentra lejos del origen.

El Gradiente Descendente con paso adaptativo mostró un comportamiento notablemente más robusto. En todos los experimentos realizados, el método logró converger al mínimo global, independientemente del punto inicial considerado. Este resultado se explica por el ajuste dinámico del tamaño de paso, que garantiza un descenso suficiente del valor de la función en cada iteración.

Por su parte, el Método de Newton presentó convergencia cuadrática cuando el punto inicial se encontraba en un entorno suficientemente cercano al origen, en concordancia con la definición positiva de la Hessiana en el mínimo. No obstante, para puntos iniciales alejados, el método puede fallar, ya sea debido a una mala aproximación cuadrática de la función o a problemas asociados a la inversión de la Hessiana. Por ejemplo, al tomar puntos iniciales con valores grandes de $|x|$, el método puede generar iteraciones que no producen descenso del valor de la función.

En conjunto, los resultados numéricos confirman que, aunque todos los métodos pueden converger al mínimo global bajo condiciones adecuadas, los métodos de primer orden con estrategias de paso adaptativo ofrecen mayor robustez global, mientras que el Método de Newton destaca por su alta eficiencia local.

8. Comparación de los algoritmos de optimización

Con el fin de evaluar el comportamiento práctico de los métodos de optimización estudiados, se realizaron experimentos numéricos desde un conjunto amplio de puntos iniciales distribuidos uniformemente en el plano. El análisis se centró en cuatro criterios fundamentales: tasa de convergencia global, dependencia del punto inicial, velocidad de convergencia y precisión final alcanzada.

8.1. Regiones de convergencia y dependencia del punto inicial

La Figura 1 muestra los mapas de calor asociados a la convergencia de cada algoritmo en función del punto inicial. Cada punto del plano se clasifica según si el método logra o no converger al mínimo global $(0,0)$ bajo el criterio de parada establecido.

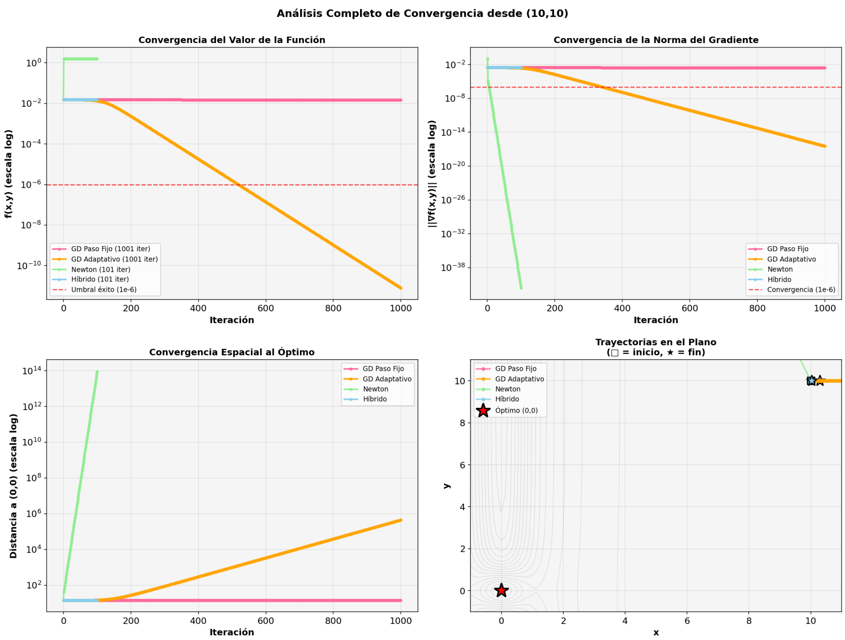

Figura 1: Mapa de regiones de convergencia para los distintos algoritmos de optimización.

Se observa que el método de Gradiente Descendente con paso fijo presenta una región de convergencia extremadamente reducida, limitada a un entorno muy cercano al mínimo. Este comportamiento refleja su fuerte dependencia del tamaño de paso y su escasa robustez frente a elecciones iniciales desfavorables.

Por el contrario, el Gradiente Descendente con paso adaptativo exhibe una región de convergencia significativamente más amplia, lo que indica una mayor estabilidad global. No obstante, la presencia de bandas donde el método no converge sugiere que la geometría anisotrópica de la función afecta el desempeño del algoritmo en ciertas direcciones del espacio.

El Método de Newton converge únicamente desde un conjunto reducido de puntos iniciales, confirmando su carácter esencialmente local. Este resultado concuerda con la teoría, ya que el método depende de una buena aproximación cuadrática de la función en el entorno del mínimo.

8.2. Velocidad de convergencia y comportamiento dinámico

La Figura 2 presenta la evolución temporal de distintos indicadores numéricos para un punto inicial representativo. En particular, se analizan el valor de la función objetivo, la norma del gradiente y la distancia al óptimo.

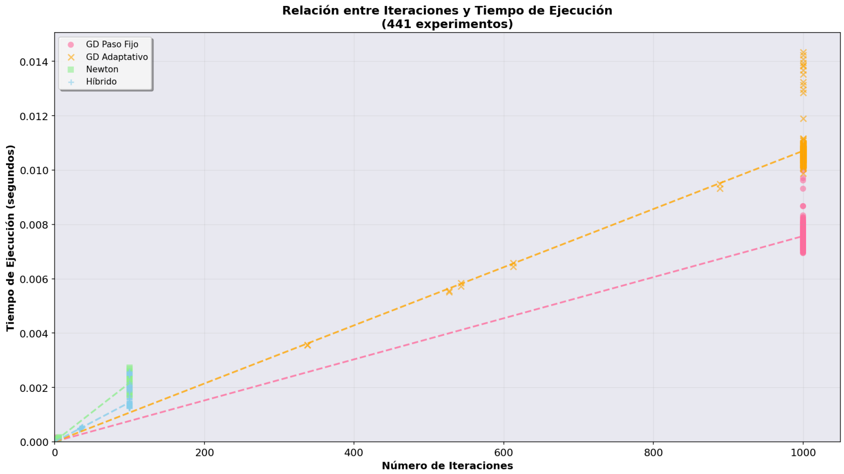

Figura 2: Evolución de la función objetivo, norma del gradiente y distancia al óptimo.

El Método de Newton muestra una disminución abrupta del error en pocas iteraciones, consistente con convergencia cuadrática cuando el punto inicial se encuentra en una región donde la Hessiana es definida positiva. Este comportamiento confirma la alta eficiencia local del método.

El Gradiente Descendente con paso adaptativo presenta una convergencia más lenta, pero estable y monótona. A diferencia del paso fijo, el ajuste dinámico del tamaño de paso evita oscilaciones y garantiza una reducción progresiva del valor de la función.

8.3. Precisión final alcanzada

La Figura 3 muestra la distribución de la precisión final alcanzada en los experimentos exitosos, medida mediante $-\log_{10}(f(x, y))$.

Los métodos de segundo orden alcanzan niveles de precisión superiores, aunque con una mayor variabilidad entre ejecuciones.

El Método de Newton logra precisiones elevadas cuando converge, mientras que el Gradiente Descendente con paso adaptativo presenta resultados más consistentes, aunque ligeramente menos precisos. El método con paso fijo no produce resultados estadísticamente relevantes debido a su baja tasa de éxito global.

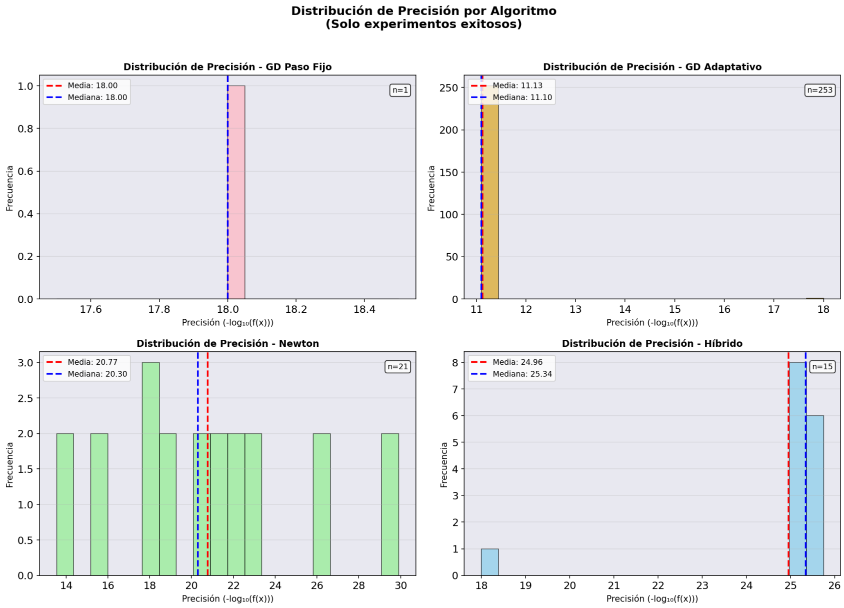

Figura 3: Distribución de la precisión final lograda por los algoritmos en los experimentos.

8.4. Discusión general

Los resultados numéricos ponen de manifiesto el compromiso clásico entre robustez global y rapidez de convergencia local. Los métodos de primer orden con estrategias adaptativas ofrecen una mayor estabilidad frente a la elección del punto inicial, mientras que los métodos de segundo orden son altamente eficientes en entornos favorables, pero sensibles a condiciones iniciales adversas.

Este análisis confirma que la elección del algoritmo de optimización debe basarse no solo en su orden teórico de convergencia, sino también en la estructura global de la función objetivo y en la información disponible sobre el punto inicial.

9. Conclusiones

En este trabajo se realizó un análisis teórico y numérico de la función

$$f(x, y) = \frac{\arctan(x^2 + y^2)}{1 + x^2},$$

desde la perspectiva de la Programación No Lineal. Se demostró que la función es suave en todo $\mathbb{R}^2$, posee un único mínimo global en el origen y presenta convexidad local en un entorno del punto óptimo.

El estudio analítico permitió justificar rigurosamente la aplicabilidad de distintos métodos de optimización. En particular, la definición positiva de la Hessiana en el mínimo explica la rápida convergencia local del Método de Newton, mientras que la estructura global de la función pone de manifiesto las limitaciones de los métodos de segundo orden cuando se parte de puntos iniciales alejados del óptimo.

Los experimentos numéricos confirmaron estas predicciones teóricas. Los métodos de primer orden, especialmente el Gradiente Descendente con paso adaptativo, mostraron una mayor robustez global, garantizando convergencia al mínimo bajo una amplia gama de condiciones iniciales. Por otro lado, el Método de Newton destacó por su alta eficiencia local, aunque su desempeño depende fuertemente de una buena elección del punto inicial.

En conjunto, los resultados evidencian un principio fundamental de la optimización numérica: no existe un método universalmente superior, sino que la elección del algoritmo debe realizarse en función de las propiedades analíticas de la función objetivo y de la información disponible sobre el punto inicial. El análisis realizado proporciona un ejemplo concreto de cómo la teoría y la experimentación numérica se complementan para comprender el comportamiento real de los algoritmos de optimización en problemas no lineales.