

Análisis teórico y experimental de un modelo de optimización no convexo

Guillermo Hughes Cardona – Grupo C311

11 de noviembre de 2025

Resumen

En este trabajo se analiza un problema de optimización no convexo propuesto en clase, cuya estructura permite estudiar el comportamiento de distintos métodos numéricos. Se demuestra teóricamente la existencia de un único mínimo global y se evalúan empíricamente dos algoritmos clásicos: el descenso por gradiente y el método de Newton regularizado. El estudio incluye más de 400 experimentos, combinando casos diseñados con intención analítica y puntos iniciales aleatorios. Se presentan los resultados, se discuten las trayectorias de convergencia y se comparan las características de cada enfoque.

1 Introducción

El objetivo de este estudio es examinar un problema de optimización sin restricciones que, aunque simple en apariencia, combina propiedades analíticas interesantes: es suave, acotado inferiormente y no convexo globalmente. Este tipo de problemas permite poner a prueba métodos iterativos de primer y segundo orden, y observar cómo responden en presencia de regiones planas o curvaturas cambiantes.

El trabajo se divide en dos partes. Primero, se desarrolla un análisis teórico completo de la función y se demuestra la existencia y unicidad del mínimo global. Luego, se implementan y comparan dos algoritmos de optimización: el descenso por gradiente, con reglas de paso fijo y adaptativo (Armijo), y un método de Newton regularizado. Ambos se aplican sobre una amplia gama de condiciones iniciales, permitiendo observar su desempeño y robustez.

2 Modelo y análisis teórico

La función a minimizar está dada por:

$$f(x, y) = -\tan\left(\frac{1}{(x+y)^2+1}\right) - \tan\left(\frac{1}{(x-y)^2+1}\right),$$

una expresión suave en todo \mathbb{R}^2 . Al definir $u = x + y$ y $v = x - y$, el problema se separa como $f(x, y) = g(u) + g(v)$ con $g(t) = -\tan(1/(t^2+1))$. Este cambio de variables simplifica el análisis, ya que el comportamiento global de f depende exclusivamente de las propiedades de g .

La función g es continua y derivable en todo el dominio. Como $1/(t^2+1) \in (0, 1]$ y \tan es creciente, se tiene $-\tan(1) \leq g(t) < 0$. La derivada

$$g'(t) = \frac{2t}{((t^2+1)^2)\cos^2(1/(t^2+1))}$$

es negativa para $t < 0$, nula en $t = 0$ y positiva para $t > 0$. En consecuencia, g decrece hasta el origen y luego crece, mostrando un mínimo global único en $t = 0$. El valor de dicho mínimo es $g(0) = -\tan(1)$.

A partir de esta simetría, el punto crítico de f es $(x^*, y^*) = (0, 0)$ y su valor mínimo $f^* = -2 \tan(1) \approx -3.1148$. La segunda derivada de g ,

$$g''(t) = \frac{-6t^4 - 4t^2 - 8t^2 \tan(1/(t^2 + 1)) + 2}{((t^2 + 1)^4) \cos^2(1/(t^2 + 1))},$$

resulta positiva cerca del origen pero negativa para $|t|$ grandes, lo cual indica que g y, por tanto, f , no son convexas globalmente. Sin embargo, la Hessiana de f en $(0, 0)$,

$$\nabla^2 f(0, 0) = 2g''(0)I = \begin{bmatrix} 13.702 & 0 \\ 0 & 13.702 \end{bmatrix},$$

es definida positiva, garantizando un mínimo local estricto. Dado el comportamiento monótono de g , ese mínimo local es también el mínimo global. El problema, aunque no convexo, es unimodal y estable.

3 Métodos de optimización

Se implementaron dos algoritmos clásicos para resolver el problema: el descenso por gradiente y el método de Newton regularizado, ambos con regla de paso basada en la condición de Armijo.

Descenso por gradiente

El método del gradiente avanza en la dirección de mayor descenso $d_k = -\nabla f(x_k)$ con una longitud de paso α_k :

$$x_{k+1} = x_k + \alpha_k d_k.$$

Se evaluaron dos estrategias de paso. La primera usa un α fijo (por ejemplo, 0.1), lo que permite observar la sensibilidad del método a la escala. La segunda emplea *backtracking* de Armijo, reduciendo α hasta que se cumple

$$f(x_k + \alpha d_k) \leq f(x_k) + c \alpha \nabla f_k^T d_k.$$

Esto asegura descenso controlado, aunque puede requerir más evaluaciones de f .

Newton regularizado

El método de Newton busca una dirección d_k que satisfaga $H_k d_k = -\nabla f_k$, donde H_k es la Hessiana. Para evitar problemas cuando H_k no es definida positiva, se añadió una regularización $(H_k + \mu I)d_k = -\nabla f_k$, con $\mu > 0$. Si la dirección calculada no es de descenso, se sustituye por $-\nabla f_k$. Este método aprovecha la información de curvatura y logra convergencia cuadrática cerca del óptimo.

Ambos algoritmos se detienen cuando la norma del gradiente es menor que 10^{-10} o tras un número máximo de iteraciones. Su implementación se realizó desde cero en **Python**, utilizando **NumPy** para el manejo de vectores y matrices.

4 Diseño experimental

Se ejecutaron más de 400 experimentos organizados en dos grupos. Los *principales* fueron diseñados con propósito analítico, mientras que los restantes se generaron aleatoriamente para evaluar la robustez global.

En el primer grupo se estudiaron inicios ubicados sobre los ejes coordenados $(\pm r, 0)$ y $(0, \pm r)$ y sobre las diagonales $(\pm r, \pm r)$ y $(\pm r, \mp r)$ con $r \in \{1, 2, 5, 10, 20\}$. Además, se exploró la influencia

del tamaño de paso en el método del gradiente, probando valores $\alpha = \{0.05, 0.1, 0.2, 0.5\}$ desde los puntos $(5, -5)$ y $(10, 0)$.

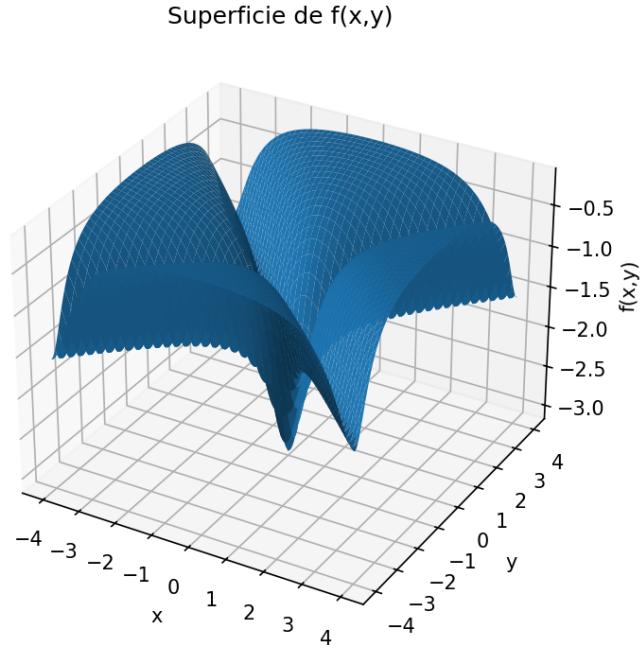
El segundo grupo incluyó 300 inicios aleatorios dentro del cuadrado $[-100, 100]^2$, la mayoría evaluados con Newton regularizado, y una fracción menor con gradiente y Armijo, para mostrar diferencias de velocidad y estabilidad. En cada experimento se registraron el número de iteraciones, el valor final de la función y la norma del gradiente final.

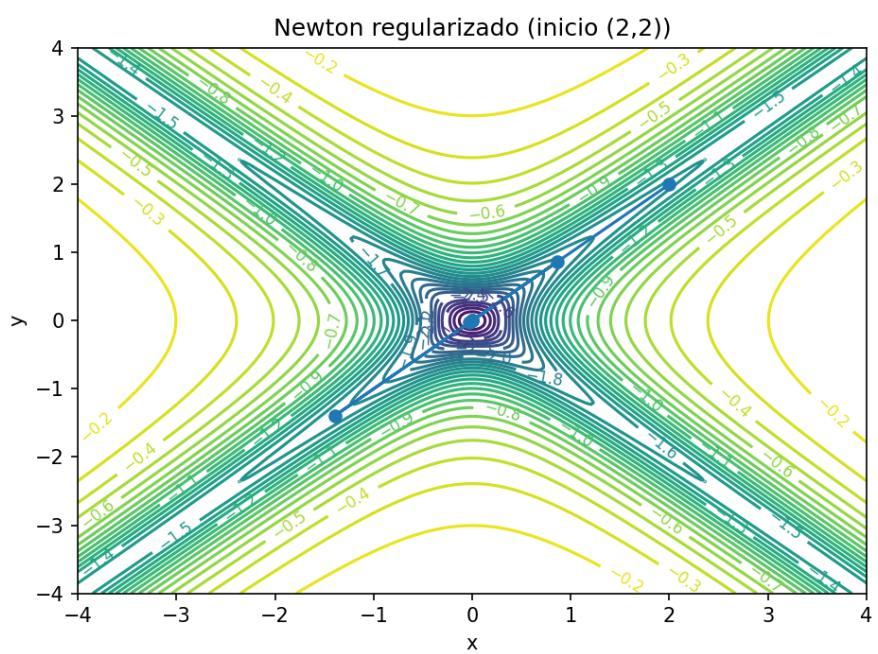
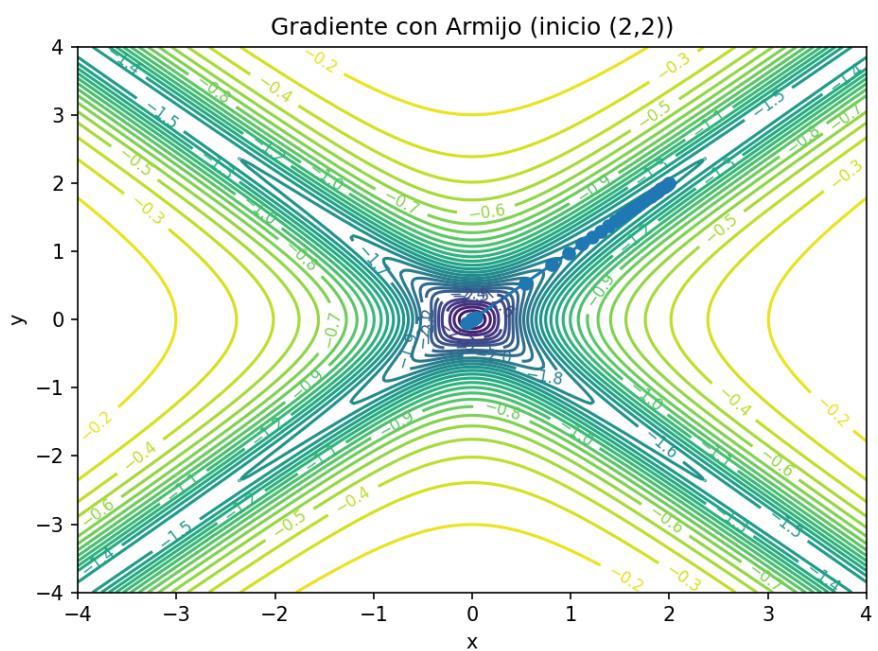
5 Resultados y discusión

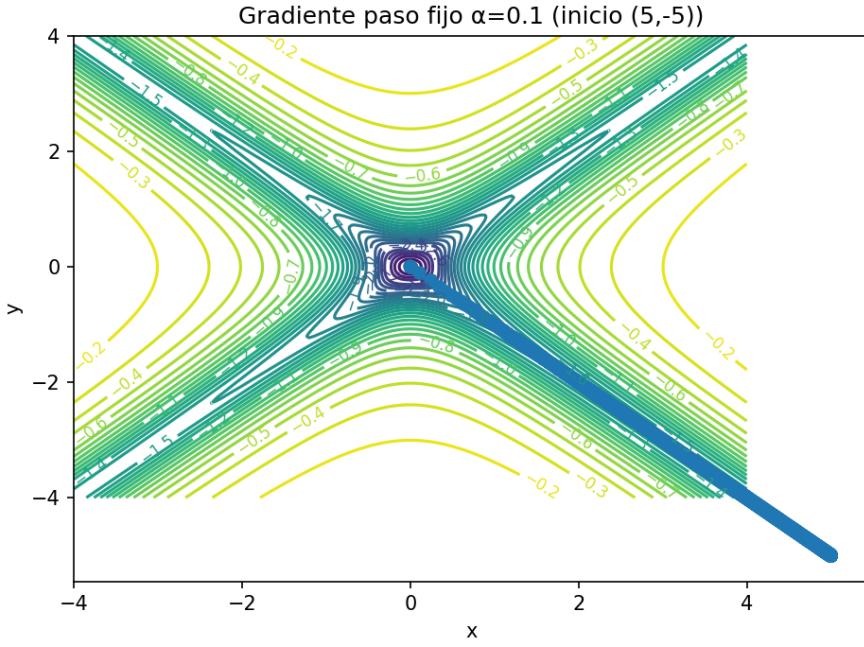
Las simulaciones confirman el análisis teórico. En todos los casos, los métodos convergen al mismo punto $(0, 0)$ con $f = -3.1148$. El gradiente descendente presenta convergencia lenta en regiones planas, especialmente cuando $|x|, |y|$ son grandes o α es pequeño. El ajuste de Armijo mejora la estabilidad, aunque con más evaluaciones.

El método de Newton regularizado se comporta de forma más eficiente: alcanza el mínimo en pocas iteraciones, incluso desde puntos lejanos, y no muestra divergencia. El uso de la regularización y de Armijo evita los problemas típicos del Newton puro en zonas no convexas.

Las figuras generadas muestran la superficie tridimensional de $f(x, y)$ y las trayectorias de ambos algoritmos sobre los contornos de nivel. En todas ellas, las curvas de convergencia se dirigen suavemente hacia el centro del valle.







Para mayor claridad, la tabla 1 resume algunos de los resultados más representativos.

Inicio x0	Inicio y0	Método	x*	y*	f(x*)	Iteraciones
2	2	Gradiente (Armijo)	-1.80784e-09	-1.80784e-09	-3.11482	10000
2	2	Gradiente ($\alpha=0.1$)	1.31608e-11	1.31608e-11	-3.11482	225
2	2	Newton (regularizado, Armijo)	1.07533e-22	1.09188e-22	-3.11482	6
5	-5	Gradiente (Armijo)	-3.31411e-09	3.31411e-09	-3.11482	10000
5	-5	Gradiente ($\alpha=0.1$)	1.13337e-11	-1.13337e-11	-3.11482	6526
5	-5	Newton (regularizado, Armijo)	-1.2643e-18	1.2643e-18	-3.11482	7
10	0	Gradiente (Armijo)	-6.62823e-09	0	-3.11482	10000
10	0	Gradiente ($\alpha=0.1$)	2.26674e-11	0	-3.11482	6526
10	0	Newton (regularizado, Armijo)	-2.52839e-18	0	-3.11482	7

Cuadro 1: Resumen de resultados para diferentes puntos iniciales y métodos.

6 Conclusiones

El problema analizado, pese a su falta de convexidad global, presenta un único mínimo global y un comportamiento unimodal. El método de Newton regularizado con Armijo demostró ser el más eficiente y estable. El descenso por gradiente, aunque más simple, confirmó la influencia decisiva del tamaño de paso y de las condiciones iniciales. Los experimentos aleatorios evidencian la robustez del método de Newton, que converge en todas las pruebas.

El trabajo permitió integrar análisis teórico, implementación numérica y evaluación experimental, reproduciendo de forma práctica los conceptos estudiados en el curso de Optimización.