

Análisis de un Modelo · Optimización

Jery Rodríguez Fernández · C312

Noviembre 2025

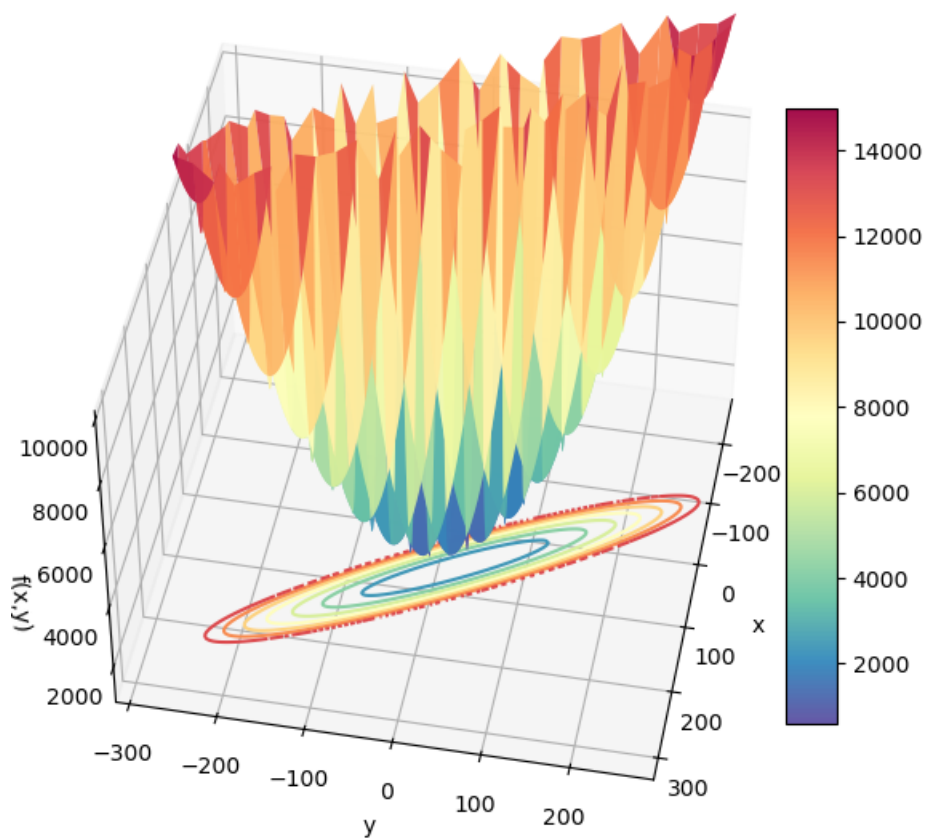


Figura 1: Superficie de la función $f(x,y) = (x+y-7)^2 + (2x+y-5)^2$.

1. Introducción del Modelo: ¿A qué nos enfrentamos?

En este trabajo se analizará el modelo matemático definido por la función

$$f(x, y) = (x + y - 7)^2 + (2x + y - 5)^2,$$

dicha función representa una superficie cuadrática en dos variables continuas y encierra una rica estructura matemática que nos permitirá aplicar y consolidar varios de los conceptos fundamentales vistos en el curso. Entonces... sin más preámbulos, comencemos.

2. Análisis del modelo: ¡Este es el terreno de batalla!

Vamos a analizar paso a paso la estructura de nuestro modelo $f(x, y)$. Para ello, recordemos que una función cuadrática puede expresarse en forma vectorial como:

$$f(x) = \frac{1}{2}x^T Qx + c^T x + r$$

donde:

- $\frac{1}{2}x^T Qx$: **Término cuadrático.** Q es una matriz simétrica que define la curvatura de la función. Si Q es *definida positiva*, la función es convexa.
- $c^T x$: **Término lineal.** c es un vector de coeficientes que desplaza el punto donde se alcanza el mínimo o máximo.
- r : **Término constante.** Es un escalar que representa un desplazamiento vertical de la función. No afecta la ubicación del óptimo, pero sí el valor de la función en ese punto.

2.1. ¿Será $f(x, y)$ un modelo cuadrático?

Queremos mostrar, paso a paso, que

$$f(x, y) = (x + y - 7)^2 + (2x + y - 5)^2$$

puede escribirse en la forma cuadrática vectorial

$$f(x) = \frac{1}{2}x^T Qx + c^T x + r, \quad x = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2,$$

determinando explícitamente Q , c y r . Para ello expandimos los cuadrados para obtener el polinomio en x, y :

$$\begin{aligned} f(x, y) &= (x + y - 7)^2 + (2x + y - 5)^2 \\ &= (x^2 + 2xy + y^2 - 14x - 14y + 49) \\ &\quad + (4x^2 + 4xy + y^2 - 20x - 10y + 25) \\ &= 5x^2 + 6xy + 2y^2 - 34x - 24y + 74. \end{aligned}$$

Luego, sea la matriz

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$$

Entonces,

$$x^\top Q x = q_{11}x^2 + (q_{12} + q_{21})xy + q_{22}y^2$$

y

$$\frac{1}{2}x^\top Q x = \frac{q_{11}}{2}x^2 + \frac{q_{12} + q_{21}}{2}xy + \frac{q_{22}}{2}y^2$$

Normalmente tomamos Q simétrica ($q_{12} = q_{21}$), porque cualquier componente anti-simétrica de Q no contribuye a $x^\top Q x$. Además, para la función cuadrática dada, la matriz Q coincide con la matriz Hessiana H_f (segunda derivada), porque:

$$\nabla f(x) = \frac{1}{2}(Q + Q^\top)x + c$$

y si $Q = Q^\top$, entonces:

$$\nabla f(x) = Qx + c, \quad \nabla^2 f(x) = Q$$

Comparamos los coeficientes de los términos cuadráticos entre:

$$\frac{1}{2}x^\top Q x \quad \text{y} \quad 5x^2 + 6xy + 2y^2$$

Suponiendo Q simétrica ($q_{12} = q_{21}$), se tiene:

$$\frac{q_{11}}{2} = 5 \Rightarrow q_{11} = 10 \quad \frac{q_{22}}{2} = 2 \Rightarrow q_{22} = 4 \quad \frac{2q_{12}}{2} = 6 \Rightarrow q_{12} = 6$$

Por tanto,

$$Q = \begin{pmatrix} 10 & 6 \\ 6 & 4 \end{pmatrix}$$

Ahora procedemos con los coeficientes lineales y el término independiente. La parte lineal $c^\top x = c_1x + c_2y$ debe coincidir con $-34x - 24y$. Entonces:

$$c = \begin{pmatrix} -34 \\ -24 \end{pmatrix}$$

Mientras que el término independiente es:

$$r = 74$$

Por tanto:

$$Q = \begin{pmatrix} 10 & 6 \\ 6 & 4 \end{pmatrix}, \quad c = \begin{pmatrix} -34 \\ -24 \end{pmatrix}, \quad r = 74$$

2.2. ¿Y si exploramos la existencia y unicidad de una solución?

De conferencias sabemos que un modelo de optimización busca un vector $x \in \mathbb{R}^n$ que minimice o maximice una función objetivo $f(x)$, sujeta a un conjunto de restricciones.

$$\begin{aligned} &\text{Minimizar (o maximizar)} \quad f(x) \\ &\text{sujeto a} \quad h_i(x) = 0, \quad i = 1, \dots, m \\ &\quad \quad \quad g_j(x) \leq 0, \quad j = 1, \dots, p \end{aligned}$$

donde:

- $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es la función objetivo (lo que queremos optimizar).
- $h_i(x) = 0$ son las restricciones de igualdad.
- $g_j(x) \leq 0$ son las restricciones de desigualdad.

y denotamos por M al conjunto de soluciones factibles, es decir:

$$M = \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} h_i(x) = 0, \quad i = 1, \dots, m \\ g_j(x) \leq 0, \quad j = 1, \dots, p \end{array} \right\}$$

Para analizar la existencia o no de solución en nuestro modelo vamos a introducir el término “coercividad”. En términos matemáticos, una función $f(x)$ es coerciva si:

$$\|x\| \rightarrow \infty \Rightarrow f(x) \rightarrow \infty$$

Es decir, cuando las variables se alejan mucho del origen, el valor de la función también crece sin límite. Esto asegura que no existan soluciones en el infinito, es decir: si una función no es coerciva, puede que no tenga un mínimo alcanzable, porque la solución “óptima” podría estar en el infinito. Para ello veamos si

$$\lim_{\|(x,y)\| \rightarrow \infty} f(x,y) = +\infty,$$

Para el siguiente análisis, utilizaremos un enfoque basado en coordenadas polares, contando con el valioso respaldo del análisis matemático impartido por la excelente profesora Idania Urrutia. Dicho lo anterior, sabemos que:

$$x = r \cos \theta, \quad y = r \sin \theta, \quad \|(x,y)\| = r$$

Tomando nuestra función en su forma explícita:

$$f(x,y) = 5x^2 + 6xy + 2y^2 - 34x - 24y + 74,$$

sustituimos en $f(x,y)$:

$$\begin{aligned} f(r,\theta) &= 5(r \cos \theta)^2 + 6(r \cos \theta)(r \sin \theta) + 2(r \sin \theta)^2 \\ &\quad - 34(r \cos \theta) - 24(r \sin \theta) + 74 \\ &= r^2(5 \cos^2 \theta + 6 \cos \theta \sin \theta + 2 \sin^2 \theta) \\ &\quad - r(34 \cos \theta + 24 \sin \theta) + 74 \end{aligned}$$

Observamos que:

$$f(r,\theta) = r^2 A(\theta) - r B(\theta) + 74$$

donde:

$$A(\theta) = 5 \cos^2 \theta + 6 \cos \theta \sin \theta + 2 \sin^2 \theta \quad \text{y} \quad B(\theta) = 34 \cos \theta + 24 \sin \theta$$

Aplicamos la factorización r^2 :

$$f(r,\theta) = r^2 \left[A(\theta) - \frac{B(\theta)}{r} + \frac{74}{r^2} \right]$$

Analizamos el límite cuando $r \rightarrow \infty$:

$$\Rightarrow \lim_{r \rightarrow \infty} f(r,\theta) = \lim_{r \rightarrow \infty} r^2 \left[A(\theta) - \frac{B(\theta)}{r} + \frac{74}{r^2} \right] = \lim_{r \rightarrow \infty} r^2 A(\theta)$$

Antes de continuar, necesitamos saber si $A(\theta)$ es positivo o negativo. Para ello recurrimos a Álgebra II. Pues, en su momento estudiamos que una forma cuadrática en dos variables reales x y y se puede

escribir como:

$$A(x, y) = ax^2 + 2bxy + cy^2$$

Esto matricialmente se ve algo así:

$$A(x, y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \vec{v}^T Q \vec{v}$$

donde Q es una matriz simétrica. Ahora... si vemos la expresión $A(\theta)$ como una forma cuadrática resultaría de la siguiente manera:

$$A(\theta) = \begin{bmatrix} \cos \theta & \sin \theta \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$$

donde:

$$A(\theta) = \vec{v}^T Q \vec{v} \quad \text{con} \quad \vec{v} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad \text{y} \quad Q = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$$

Lo interesante de esta vía es que el signo de $A(\theta)$ se reduce a analizar si Q es definida positiva, indefinida o negativa. Por tanto, calculamos los menores principales:

- Primer menor: $5 > 0$
- Determinante:

$$\det(Q) = 5 \cdot 2 - 3^2 = 10 - 9 = 1 > 0$$

Conclusión: Por el criterio de Sylvester, si todos los menores principales son positivos, la matriz Q es definida positiva y como Q es definida positiva, se cumple:

$$A(\theta) = \vec{v}^T Q \vec{v} > 0 \quad \text{para todo } \theta$$

Luego, como ya sabemos que $A(\theta) > 0$ para todo θ , el término $r^2 A(\theta)$ crece positivamente sin límite. Por lo que:

$$\lim_{r \rightarrow \infty} f(r, \theta) = +\infty$$

cumpléndose lo anterior podemos concluir que:

$$f(x, y) \rightarrow +\infty \quad \text{cuando} \quad \|(x, y)\| \rightarrow \infty.$$

Finalmente, la función $f(x, y)$ es coercitiva, lo cual garantiza que no existen soluciones en el infinito. En consecuencia, se asegura la existencia de un mínimo alcanzable, es decir, el mínimo de $f(x, y)$

en M se alcanza.

2.3. ¿No te intriga lo que revela el gradiente y la matriz Hessiana?

Para refrescar los conceptos de la matriz Hessiana y el gradiente, considere la siguiente narrativa que los conecta con la intuición geométrica:

La función $f(x, y)$ vivía tranquila... hasta que apareció **el gradiente**, anunciando con voz firme:

“¡Soy el vector de las derivadas parciales y vengo a señalar el camino más empinado!”

Pero desde las sombras, emergió **la matriz Hessiana**, con tono misterioso:

“Yo decido si esto es un valle, una montaña... o un punto de silla que te rompe el corazón.”

Así comienza el análisis multivariable:

- El gradiente nos dice hacia dónde cambia más la función.
- La Hessiana nos revela cómo cambia: si la curvatura nos lleva a un mínimo, un máximo... o a una incómoda indecisión.

2.3.1. El punto de partida será el gradiente

Sea

$$f(x, y) = 5x^2 + 6xy + 2y^2 - 34x - 24y + 74.$$

El gradiente $\nabla f(x, y)$ es el vector de derivadas parciales de primer orden:

$$\nabla f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} = \begin{pmatrix} 10x + 6y - 34 \\ 6x + 4y - 24 \end{pmatrix}.$$

Interpretación:

- En cada punto (x, y) , el vector $\nabla f(x, y)$ apunta hacia la dirección de mayor crecimiento de la función f .
- La magnitud $\|\nabla f(x, y)\|$ indica qué tan rápido crece f en esa dirección.
- La dirección opuesta al gradiente, $-\nabla f(x, y)$, corresponde al descenso más pronunciado de f .

2.3.2. Bien, sigamos con la matriz Hessiana

La matriz Hessiana de una función f en un punto $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ se define como la matriz cuadrada de segundas derivadas parciales:

$$H_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Sea la función

$$f(x, y) = 5x^2 + 6xy + 2y^2 - 34x - 24y + 74.$$

Calculamos las derivadas parciales de primer y segundo orden:

$$f_x(x, y) = \frac{\partial f}{\partial x} = 10x + 6y - 34, \quad f_y(x, y) = \frac{\partial f}{\partial y} = 6x + 4y - 24.$$

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} = 10, \quad f_{xy} = \frac{\partial^2 f}{\partial x \partial y} = 6, \quad f_{yx} = \frac{\partial^2 f}{\partial y \partial x} = 6, \quad f_{yy} = \frac{\partial^2 f}{\partial y^2} = 4.$$

Como $f \in C^2$, las derivadas mixtas coinciden ($f_{xy} = f_{yx}$), por lo tanto la matriz Hessiana es simétrica.

$$H_f(x, y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 10 & 6 \\ 6 & 4 \end{bmatrix}.$$

(Nota: es constante porque f es un polinomio cuadrático.)

Usamos el criterio de Sylvester (menores principales) para determinar si nuestra matriz simétrica es definida positiva:

- Primer menor principal: $10 > 0$
- Determinante de la Hessiana:

$$\det(H) = 10 \cdot 4 - 6 \cdot 6 = 40 - 36 = 4 > 0$$

Como todos los menores principales son positivos, H es definida positiva. Alternativamente, los autovalores son:

$$\lambda_1 = 7 + \sqrt{45} \approx 13,354, \quad \lambda_2 = 7 - \sqrt{45} \approx 0,646$$

Ambos positivos, lo que confirma la positividad.

Conclusión: La función f es estrictamente convexa, y cualquier punto estacionario será un mínimo global. De igual forma, dado que la Hessiana es definida positiva, el mínimo es estricto y único. No hay otros mínimos locales.

2.4. Localizando el punto estacionario

Para determinar los puntos estacionarios (esos donde, como bien sabemos, todo se detiene) se retoman los criterios previamente establecidos, identificando aquellos en los que la pendiente se anula mediante el gradiente previamente calculado:

$$\nabla f(x, y) = \begin{pmatrix} 10x + 6y - 34 \\ 6x + 4y - 24 \end{pmatrix}$$

Igualemos el gradiente a cero:

$$\begin{cases} 10x + 6y - 34 = 0 \\ 6x + 4y - 24 = 0 \end{cases}$$

Simplificamos las ecuaciones, dividiéndolas por 2, tenemos:

$$\begin{cases} 5x + 3y = 17 \\ 3x + 2y = 12 \end{cases}$$

Multiplicamos la segunda ecuación por $-\frac{5}{3}$ para eliminar x :

$$-\frac{5}{3}(3x + 2y = 12) \Rightarrow -5x - \frac{10}{3}y = -20$$

Ahora sumamos con la primera ecuación:

$$(5x + 3y) + (-5x - \frac{10}{3}y) = 17 - 20$$

Simplificamos:

$$3y - \frac{10}{3}y = -3$$

Llevamos a común denominador:

$$\frac{9y - 10y}{3} = -3 \Rightarrow \frac{-y}{3} = -3$$

Despejamos y :

$$y = 9$$

Sustituimos en una ecuación para hallar x (usamos la segunda):

$$3x + 2(9) = 12 \Rightarrow 3x + 18 = 12 \Rightarrow 3x = -6 \Rightarrow x = -2$$

Por lo tanto, el punto estacionario es:

$$(x^*, y^*) = (-2, 9)$$

Para verificar que es mínimo utilizamos la matriz Hessiana:

$$H = \begin{bmatrix} 10 & 6 \\ 6 & 4 \end{bmatrix}$$

Determinante:

$$\det(H) > 0$$

Por tanto, H es definida positiva, y el punto estacionario (extremo local) es un mínimo global.

Calculando el valor mínimo:

$$f(-2, 9) = ((-2) + 9 - 7)^2 + (2(-2) + 9 - 5)^2 = (0)^2 + (0)^2 = 0$$

Resultado final:

$$(x^*, y^*) = (-2, 9), \quad \boxed{f_{\min} = 0}$$

3. Del papel al procesador: ¡Es la hora de los algoritmos!

En esta sección se presentan los pseudocódigos correspondientes a los algoritmos implementados, acompañados de una descripción breve de su funcionamiento. Para una revisión más exhaustiva, consultar el archivo complementario en formato .ipynb adjunto a este documento el cual contiene el código fuente en Python junto con una explicación detallada y robusta de cada método.

3.1. Método de Máximo Descenso

El Método de Máximo Descenso es uno de los algoritmos fundamentales en la optimización numérica. Su base es puramente geométrica y analítica: moverse en la dirección de mayor descenso de una función en cada paso, guiado por el gradiente. A continuación, veremos los basamentos matemáticos necesarios para codificar el algoritmo. Queremos minimizar una función diferenciable:

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

en nuestro caso $f(x, y)$, sin restricciones, es decir:

$$\min_{x \in \mathbb{R}^n} f(x)$$

La idea geométrica parte del gradiente, pues, el gradiente en un punto x_k es:

$$\nabla f(x_k) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_k) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_k) \end{bmatrix}$$

como bien sabemos este apunta en la dirección de máximo incremento de f .

Por tanto, si queremos disminuir f , debemos movernos en la dirección opuesta:

$$d_k = -\nabla f(x_k)$$

Esta es la dirección de descenso más pronunciado. Luego, la actualización del método se hace por iteraciones:

$$x_{k+1} = x_k + \alpha_k d_k \quad \text{con } \alpha_k \geq 0$$

Aquí, α_k es el tamaño de paso. El gran problema es: la elección del tamaño de paso. Un valor excesivamente grande de α_k puede provocar que el algoritmo sobrepase el mínimo de la función, mientras que un valor demasiado pequeño puede ralentizar significativamente la convergencia del método. Entonces, tenemos dos enfoques:

- **Fijo (heurístico):** Escogemos un valor pequeño constante, como si fuese una regla empírica, por ejemplo:

$$\alpha_k = 0,01$$

- **Óptimo (analítico):** Como bien describe la conferencia 5, escogemos α_k que minimiza f sobre la recta, esto garantiza el mayor descenso posible en esa dirección, esto se expresa

como:

$$\alpha_k = \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$$

o lo mismo:

$$\alpha_k = \min_{\alpha \geq 0} f(x_k + \alpha d_k)$$

lo cual equivale a derivar:

$$\frac{d}{d\alpha} f(x_k - \alpha \nabla f(x_k))$$

Pero... ¿y si logramos encontrar una expresión explícita para α que podamos insertar directamente en el algoritmo? De ser así, podríamos beneficiar el algoritmo de una implementación más precisa y parametrizada. Con esto en mente, avancemos hacia la formulación explícita de α .

Como vimos anteriormente, nuestro modelo corresponde a una función cuadrática estrictamente convexa:

$$\begin{aligned} f(x) &= \frac{1}{2} x^T Q x + c^T x + r \\ &= \frac{1}{2} x^T \begin{pmatrix} 10 & 6 \\ 6 & 4 \end{pmatrix} x + \begin{pmatrix} -34 & -24 \end{pmatrix} x + 74 \end{aligned}$$

del cual sabemos que el gradiente es:

$$\nabla f(x) = Qx + c$$

entonces... sustituimos $x_k + \alpha d_k$ en $f(x)$:

$$f(x_k + \alpha d_k) = \frac{1}{2} (x_k + \alpha d_k)^T Q (x_k + \alpha d_k) + c^T (x_k + \alpha d_k) + r$$

expandimos los términos:

$$\begin{aligned} (x_k + \alpha d_k)^T Q (x_k + \alpha d_k) &= x_k^T Q x_k + \alpha d_k^T Q x_k + \alpha x_k^T Q d_k + \alpha^2 d_k^T Q d_k \\ &= x_k^T Q x_k + 2\alpha x_k^T Q d_k + \alpha^2 d_k^T Q d_k \quad (\text{ya que } Q \text{ es simétrica}) \end{aligned}$$

incorporamos lo anterior:

$$\begin{aligned} f(x_k + \alpha d_k) &= \frac{1}{2} (x_k^T Q x_k + 2\alpha x_k^T Q d_k + \alpha^2 d_k^T Q d_k) + c^T (x_k + \alpha d_k) + r \\ &= \frac{1}{2} x_k^T Q x_k + \alpha x_k^T Q d_k + \frac{1}{2} \alpha^2 d_k^T Q d_k + c^T x_k + \alpha c^T d_k + r \end{aligned}$$

Ahora, agrupamos estratégicamente (en potencia de α):

$$f(x_k + \alpha d_k) = \left(\frac{1}{2}x_k^\top Q x_k + c^\top x_k + r\right) + \alpha(x_k^\top Q d_k + c^\top d_k) + \frac{1}{2}\alpha^2 d_k^\top Q d_k$$

Note que:

$$f(x_k) = \frac{1}{2}x_k^\top Q x_k + c^\top x_k + r$$

$$\nabla f(x_k) = Q x_k + c$$

$$\nabla f(x_k)^\top d_k = (Q x_k + c)^\top d_k = x_k^\top Q d_k + c^\top d_k$$

Por tanto, podemos verlo como:

$$f(x_k + \alpha d_k) = f(x_k) + \alpha \nabla f(x_k)^\top d_k + \frac{1}{2}\alpha^2 d_k^\top Q d_k$$

Procedemos a derivar respecto a α e igualar a cero:

$$\frac{d}{d\alpha} f(x_k + \alpha d_k) = \nabla f(x_k)^\top d_k + \alpha d_k^\top Q d_k = 0$$

despejamos α_k :

$$\alpha_k = -\frac{\nabla f(x_k)^\top d_k}{d_k^\top Q d_k}$$

finalmente, sustituimos $d_k = -\nabla f(x_k)$, entonces:

$$\alpha_k = \frac{\nabla f(x_k)^\top \nabla f(x_k)}{\nabla f(x_k)^\top Q \nabla f(x_k)}$$

3.1.1. El Camino hacia el Óptimo: Construyendo el Algoritmo

Bien... el algoritmo es bastante sencillo e intuitivo:

Listing 1: Pseudocódigo del Algoritmo de Máximo Descenso Óptimo

```

1 Función MaximoDescensoOptimo(Q, c, x0, tol, max_iter, verbose):
2
3   Convertir Q, c, x0 a tipo flotante
4   x <- x0
5
6   n <- dimensión de x
7   Asegurar que Q es de tamaño (n, n)
8   Asegurar que c es de tamaño (n)
9
10  Si Q no es simétrica (con tolerancia 1e-12):
```

```

11     Q <- (Q + Transpuesta(Q)) / 2
12
13     Inicializar listas: iters, fvals, alphas, norms
14
15     Definir función f_val(x_vec):
16         Retornar 0.5 * x_vec * Q * x_vec + c * x_vec
17
18     Para k desde 1 hasta max_iter:
19         g <- Q * x + c
20         gn <- norma de g
21
22         Añadir k a iters
23         Añadir f_val(x) a fvals
24         Añadir gn a norms
25
26         Si verbose es verdadero:
27             Imprimir "Iteración k: ||g|| = gn, f = f_val(x)"
28
29         Si gn < tol:
30             Si verbose:
31                 Imprimir "Criterio de parada alcanzado"
32             Salir del bucle
33
34         Qg <- Q * g
35         denom <- g * Qg
36
37         Si denom < 0:
38             Lanzar error "Denominador no positivo o Q no definida positiva"
39
40         alpha <- (g * g) / denom
41         Añadir alpha a alphas
42
43         x <- x - alpha * g
44
45     info <- Diccionario con iters, fvals, alphas, norms
46
47     Redondear cada elemento de x a 2 decimales
48
49     Retornar x, info

```

Pasos del algoritmo:

1. Se reciben los parámetros: matriz Q , vector c , punto inicial x_0 , tolerancia tol , máximo de iteraciones max_iter , y bandera `verbose`.
2. Se validan los datos: se convierten a tipo `float`, se verifica la dimensión y simetría de Q .

3. Se inicializan listas para registrar el historial: iteraciones, valores de función, pasos α_k , y normas del gradiente.
4. Se define la función objetivo cuadrática para su uso en desarrollos posteriores del algoritmo.
5. Se ejecuta un ciclo **for** que:
 - Calcula el gradiente y su norma.
 - Evalúa la función y guarda el historial.
 - Verifica el criterio de parada: $\|g_k\| < \text{tol}$.
 - Calcula el paso óptimo α_k y actualiza x .
6. Al finalizar, se redondea cada componente del vector solución y se retorna junto con el historial de información (**info**).

3.1.2. Rastro de iteraciones del método de Máximo Descenso (trayectoria):

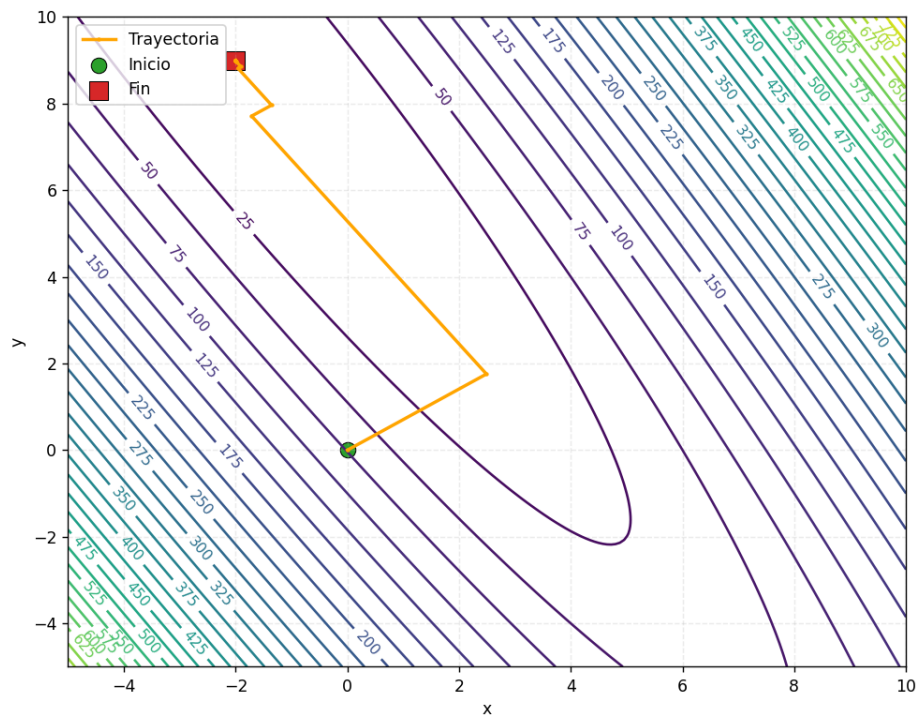


Figura 2: Ruta de convergencia del método de máximo descenso desde el punto de inicio (0; 0) hasta el óptimo (-2; 9)

3.2. Método de Newton

Método de Newton es un método iterativo para encontrar puntos donde el gradiente de una función se anula, este propone una forma rápida y cuadráticamente convergente de aproximarlos. El método se basa en aproximar localmente la función $f(x)$ mediante su expansión en serie de Taylor de segundo orden alrededor del punto actual x_k :

$$f(x) \approx f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T H(x_k) (x - x_k)$$

donde:

- $\nabla f(x_k)$ es el gradiente
- $H(x_k)$ es la matriz Hessiana

Partiendo de la expansión de Taylor de segundo orden de $f(x)$ alrededor del punto x_k , se obtiene una aproximación cuadrática local de la función. Para hallar el punto x que minimiza esa aproximación, se deriva respecto a x y se iguala el gradiente a cero, lo que da la ecuación:

$$\nabla f(x_k) + H(x_k)(x - x_k) = 0$$

si resolverla para x , se obtiene la expresión:

$$x = x_k - H(x_k)^{-1} \nabla f(x_k)$$

Este resultado define la regla iterativa del método de Newton, que actualiza el punto actual x_k restando el producto de la inversa de la Hessiana por el gradiente evaluado en x_k , avanzando así hacia el punto donde el gradiente se anula.

Ahora... recordando que nuestro modelo es cuadrático:

$$f(x) = \frac{1}{2} x^T Q x + c^T x + r, \quad x \in \mathbb{R}^n,$$

entonces, es posible interpretar el método de Newton desde otra perspectiva, como mismo hicimos en el método de máximo descenso. De secciones anteriores sabemos que:

$$\nabla f(x) = Qx + c \quad H(x) = \nabla^2 f(x) = Q$$

así como: Q simétrica y el término r no afecta ni al gradiente ni a la Hessiana. Seguimos, la iteración de Newton estándar es:

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

Sustituyendo para la cuadrática $H(x_k) = Q$ y $\nabla f(x_k) = Qx_k + c$:

$$x_{k+1} = x_k - Q^{-1}(Qx_k + c) = x_k - x_k - Q^{-1}c = -Q^{-1}c.$$

Por tanto, independientemente de x_k , obtenemos:

$x_{k+1} = -Q^{-1}c$

3.2.1. Construyendo el algoritmo

A partir del resultado obtenido, se ha logrado desarrollar un algoritmo caracterizado por su simplicidad estructural:

Listing 2: Pseudocódigo del Método de Newton para problemas de naturaleza cuadrática, como nuestro modelo.

```

1  Función Newton(Q, c, x0 )
2
3      # Q: matriz simétrica de tamaño (n x n)
4      # c: vector de tamaño (n)
5      # x0: vector inicial
6      # Devuelve x1 tal que x1 = - Q      * c si Q es invertible
7
8      n <- número de filas de Q
9
10     Si x0 no está definido entonces
11         x0 <- vector de ceros de tamaño n
12     Fin Si
13
14     b <- - (Q * x0 + c)      # Gradiente negativo en x0
15
16     # Resolver el sistema lineal Q * p = b
17     p <- resolver_Sistema_Lineal(Q, b)
18
19     x1 <- x0 + p              # Actualización de Newton
20
21     Devolver x1
22
23 Fin Función

```

Pasos del algoritmo:

1. Se determina la dimensión del problema:

$$n \leftarrow \text{número de filas de } Q$$

2. Si el vector inicial x_0 no está definido, se inicializa como el vector nulo:

$$x_0 \leftarrow \mathbf{0} \in \mathbb{R}^n$$

3. Se calcula el gradiente negativo de $f(x)$ en el punto x_0 :

$$b \leftarrow -(Qx_0 + c)$$

4. Se resuelve el sistema lineal:

$$Qp = b$$

para obtener el vector de dirección p . Nótese que si Q es invertible, entonces:

$$p = Q^{-1}b = -Q^{-1}(Qx_0 + c)$$

lo que implica que:

$$x_1 = x_0 + p = x_0 - Q^{-1}(Qx_0 + c) = -Q^{-1}c$$

5. Se actualiza la solución usando el paso de Newton:

$$x_1 \leftarrow x_0 + p$$

6. Se devuelve el nuevo punto x_1 , que corresponde a:

$$x_1 = -Q^{-1}c$$

3.2.2. Rastro de iteraciones del método de Newton (trayectoria):

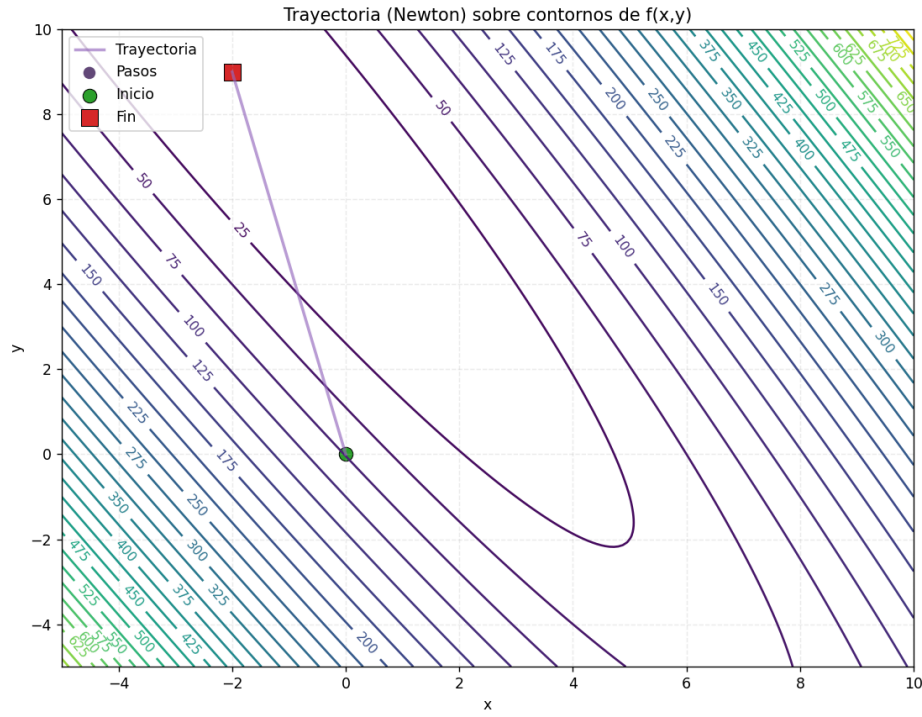


Figura 3: Ruta de convergencia del método de Newton desde el punto de inicio (0; 0) hasta el óptimo (-2; 9)

3.2.3. Análisis de Resultados:

Tras realizar múltiples experimentaciones con los algoritmos implementados de Newton y de Máximo Descenso, en su mayoría dentro del dominio definido por el intervalo $[-100, 100]$ para ambas variables x y y , se procederá a presentar los resultados obtenidos a partir de estas pruebas empíricas. A continuación, se analizará el comportamiento de ambos métodos bajo tres escenarios representativos, con el fin de ilustrar sus respectivas dinámicas de convergencia, eficiencia computacional y sensibilidad a las condiciones iniciales.

Prueba 1

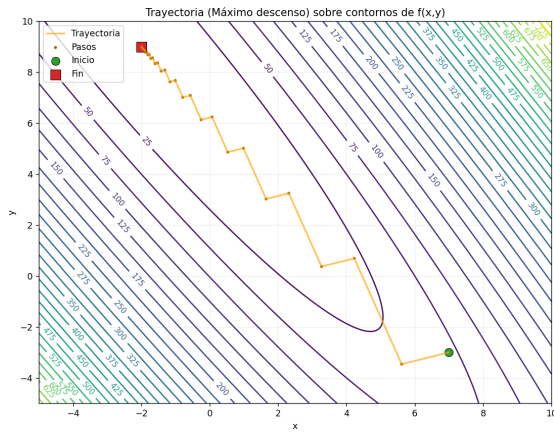


Figura 4: Método de Máximo Descenso con punto de inicio $(7; -3)$ · 100 iteraciones

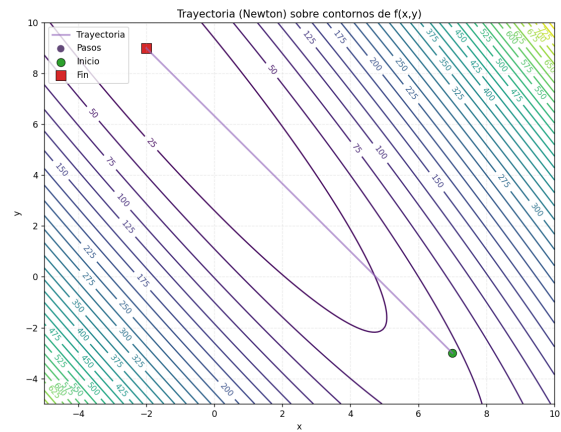


Figura 5: Método de Newton con punto de inicio $(7; -3)$

Análisis: Acá se puede observar como el método de máximo descenso muestra una trayectoria más gradual y “zigzagueante”, por así decirlo, siguiendo la dirección del gradiente en cada paso, lo que lo hace más conservador pero también más estable frente a irregularidades del terreno (relieve de la superficie). En contraste, el método de Newton avanza con pasos más amplios y directos, aprovechando la curvatura de la función, para ser preciso con una sola iteración. Este último converge más rápido, como se observa en la menor cantidad de iteraciones, aunque puede ser más sensible a la elección del punto inicial y a zonas donde la Hessiana (Q) no esté bien condicionada. Ambos métodos buscan el mínimo, pero Newton lo hace mejor y, por tanto, con mayor eficiencia en esta prueba. Note que el punto de inicio está relativamente cerca del óptimo y el algoritmo tiene un carácter bastante “zigzagueante”.

Prueba 2

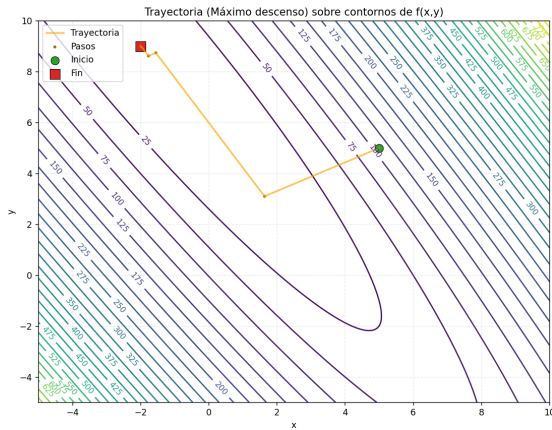


Figura 6: Método de Máximo Descenso con punto de inicio $(5; 5)$ · 19 iteraciones

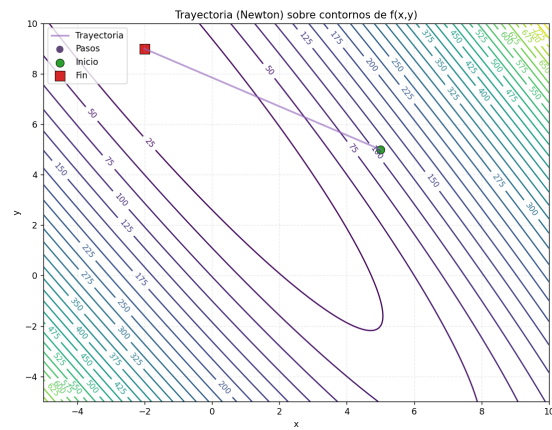


Figura 7: Método de Newton con punto de inicio $(5; 5)$ · 1 iteración

Análisis: En este segundo caso, Newton vuelve a mostrar su carácter decidido: la trayectoria es más recta y rápida, con menos pasos para alcanzar el mínimo. Se nota que aprovecha la curvatura de la función, lo que le permite “leer” mejor el terreno y ajustar su dirección con precisión. En cambio, el método de máximo descenso sigue una ruta más cautelosa, con pasos pequeños al igual que la prueba 1, además esta prueba en conjunto con la anterior muestran como el método de máximo descenso tiene una tendencia a “zigzagueante”, especialmente cuando las curvas de nivel se estrechan, es decir cuanto más cerca está del óptimo, sin embargo cuanto más lejos se encuentra el punto de inicio del óptimo más directos y grandes son los tamaños de pasos, y de esta forma, acercándose rápidamente a una vecindad relativamente reducida. Aunque ambos métodos llegan al mismo destino, Newton lo hace con mayor eficiencia, siempre que la función sea bien comportada y la Hessiana sea invertible. Este ejemplo reafirma que, cuando se puede aplicar, Newton (siempre que Q sea invertible) ofrece una ventaja clara en velocidad de convergencia.

Prueba 3

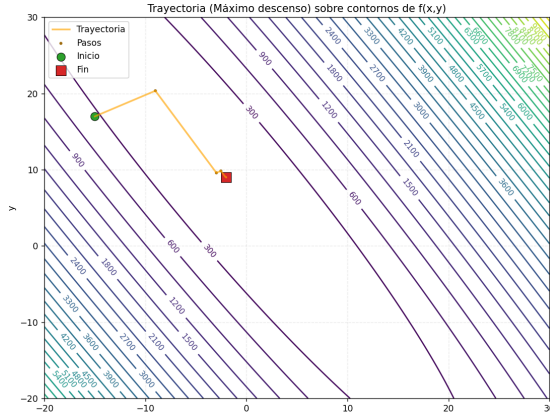


Figura 8: Método de Máximo Descenso con punto de inicio $(-15; 17)$ · 1 iteración

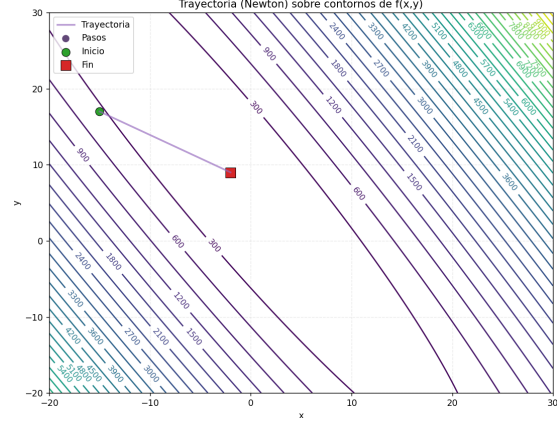


Figura 9: Método de Newton con punto de inicio $(-15; 17)$ · 1 iteración

Análisis: Para esta tercera prueba, se reafirman las conclusiones precipitadas de las pruebas anteriores. Una vez más, el método de máximo descenso sigue una trayectoria fluctuosa, con pasos pequeños que se ajustan localmente al gradiente (al acercar la imagen en ejecución del programa se nota mejor lo antes dicho). Esto lo hace más robusto pero también más lento, especialmente en regiones donde las curvas de nivel se alargan o se estrechan. En cambio, el método de Newton avanza con pasos más largos y directos, aprovechando la curvatura de la función y la naturaleza cuadrática de nuestro modelo.

3.3. Conclusiones: confrontación y comparación de los métodos usados:

En esta sección se realizará un análisis comparativo entre los algoritmos de máximo descenso y Newton. Se evaluarán diferentes aspectos con el objetivo de identificar las fortalezas y limitaciones de ambos enfoques, resaltando en qué situaciones uno puede ser más conveniente que el otro.

3.3.1. Naturaleza de los métodos:

Máximo Descenso con paso óptimo

- Iterativo: cada iteración ajusta $x_{k+1} = x_k - \alpha_k g_k$.
- El paso $\alpha_k = \frac{g^T g}{g^T Q g}$ se calcula de forma analítica.
- Requiere varias iteraciones hasta que $\|g_k\| < \text{tol}$.
- Sensible al punto de partida x_0 , aunque el paso óptimo reduce la cantidad de iteraciones comparado con un paso fijo.

- Permite ver cómo el valor de la función y el gradiente evolucionan.

Newton para cuadráticas

- Directo: en una sola iteración se obtiene el óptimo $x^* = -Q^{-1}c$.
- No depende del punto inicial x_0 para funciones cuadráticas.
- El “tamaño de paso” está implícito en la resolución del sistema lineal $Qp = -\nabla f(x_0)$.
- Perfecto para cuadráticas: converge en una iteración exactamente.

3.3.2. Comparación de número de pasos:

Método	Número de pasos típicos	Observaciones
Máximo descenso (paso óptimo)	Depende de x_0 , típicamente 2–5 iteraciones para problemas simples de 2 variables	Con paso óptimo, converge mucho más rápido que con paso fijo. El número de iteraciones aumenta si x_0 está lejos del mínimo.
Newton cuadrático	1	Garantizado para funciones cuadráticas. Independiente de x_0 .

3.3.3. Comparación de sensibilidad al punto de partida x_0 :

Máximo descenso:

- Si x_0 está lejos del mínimo, el gradiente inicial es grande y α_0 grande, pero puede requerir más pasos.

Newton:

- x_0 no importa. Siempre se obtiene $x^* = -Q^{-1}c$ en una iteración.

3.3.4. Comparación de “efecto” de tamaño de paso:

Máximo descenso paso óptimo:

- Cada paso está calculado específicamente para minimizar la función a lo largo de la dirección del gradiente.
- Esto evita elegir manualmente un α fijo, lo que podría hacer que se requieran decenas de pasos.

Newton:

- El tamaño de paso no se ajusta: se resuelve directamente la ecuación $Qp = -g$, equivalente a “paso óptimo global”, por así decirlo, en la dirección del gradiente para cuadráticas.

4. Conclusiones Generales

En este trabajo se logró realizar un estudio completo y coherente del modelo propuesto, abordando tanto su análisis teórico como la aplicación práctica de métodos numéricos de optimización. A través del desarrollo se comprobó que la función $f(x, y) = (x + y - 7)^2 + (2x + y - 5)^2$ es estrictamente convexa, ya que su matriz Hessiana resultó ser definida positiva, lo que garantiza la existencia y unicidad del mínimo global. Además, se verificó la coercividad del modelo, asegurando que no existen soluciones en el infinito y que el valor mínimo es alcanzable en un punto finito del plano. El análisis permitió identificar el punto óptimo exacto $(x^*, y^*) = (-2, 9)$, con un valor mínimo $f(x^*, y^*) = 0$. Este resultado no solo confirma la validez de los cálculos analíticos, sino también la consistencia del modelo con las propiedades de convexidad y continuidad demostradas.

En cuanto a los métodos implementados, el método del Máximo Descenso con paso óptimo mostró un comportamiento iterativo más gradual, donde la convergencia depende en parte del punto inicial, aunque el uso del paso analítico mejoró notablemente la eficiencia frente a un paso fijo. En contraste, el método de Newton ofreció una convergencia inmediata en una sola iteración, tal como se esperaba para funciones cuadráticas estrictamente convexas, reflejando su superioridad teórica y práctica en este tipo de problemas.

El método de Newton presenta una alta eficiencia computacional en la resolución de problemas de optimización, siempre que se cumpla la condición de invertibilidad de la matriz Hessiana Q , es decir, que su determinante sea distinto de cero. En escenarios donde esta condición no se satisface, resulta más adecuado emplear el método de descenso máximo como alternativa.

Con el objetivo de mejorar el desempeño de este último, se propone ajustar el criterio de parada mediante una reducción en la tolerancia de la norma del gradiente, lo cual puede mitigar el comportamiento oscilatorio o zigzagueante en las proximidades relativas del punto óptimo. No obstante, esta modificación podría comprometer la precisión de la solución, generando resultados potencialmente alejados del valor óptimo real, en algunos casos.

En síntesis, este trabajo permitió consolidar el vínculo entre la teoría y la práctica: la comprensión del gradiente y la Hessiana no solo facilita el análisis geométrico de la función, sino que guía la elección del método de optimización adecuado. Se evidenció cómo la estructura matemática del modelo define el comportamiento numérico de los algoritmos y cómo, en modelos cuadráticos bien condicionados, la precisión y rapidez del método de Newton es inigualable.

Link del repositorio: <https://github.com/jery04/Optimization-Analysis-of-a-Quadratic-Model.git>