

Optimización de la función $f(x, y) = x^4 - 4x^3 + 4x + y^2$

Nombre: Ruben Martinez Rojas
Grupo: C311

November 26, 2025

1 Modelos a analizar

La función a analizar es:

$$f(x, y) = x^4 - 4x^3 + 4x + y^2$$

2 Análisis de los modelos

2.1 Existencia de solución

Para estudiar la existencia de puntos óptimos, se calcularán los ceros del vector gradiente de la función, dado que todo mínimo (o máximo) local diferenciable debe satisfacer

$$\nabla f(x, y) = 0$$

. El gradiente está dado por el vector de derivadas parciales:

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 4x^3 - 12x^2 + 4 \\ 2y \end{bmatrix}$$

Para encontrarlos, igualamos cada componente del gradiente a cero:

$$\frac{\partial f}{\partial x} = 4x^3 - 12x^2 + 4 = 0, \quad \frac{\partial f}{\partial y} = 2y = 0.$$

De la ecuación en y se obtiene directamente:

$$y = 0.$$

Para la ecuación en x , simplificamos dividiendo entre 4:

$$x^3 - 3x^2 + 1 = 0.$$

Esta ecuación cónica puede tener una, dos o tres raíces reales.

2.2 Convexidad

La matriz Hessiana de la función

$$f(x, y) = x^4 - 4x^3 + 4x + y^2$$

es:

$$H(x, y) = \begin{bmatrix} 12x^2 - 24x & 0 \\ 0 & 2 \end{bmatrix}.$$

Para que la función sea convexa, $H(x, y)$ debe ser definida positiva. Vemos que no lo es para todos los valores de x , por lo tanto la función no es convexa globalmente. Sin embargo, se puede analizar localmente para regiones donde $12x^2 - 24x > 0$, es decir, para $x > 2$ o $x < 0$. Pues en estas regiones donde la Hessiana es definida positiva, la función tiene curvatura positiva en todas las direcciones, lo cual implica que alrededor del punto es similar a un tazón. En estas regiones la curvatura de la función no cambia de signo y las direcciones de descenso son estables. Por ello, los métodos de Gradiente, Newton y Quasi-Newton tienden a converger rápidamente hacia un mínimo local.

En cambio, cuando la Hessiana posee valores negativos (como ocurre en la región $0 < x < 2$), la función presenta curvatura negativa en alguna dirección, generando zonas tipo silla de montar. Allí los métodos de optimización tienen comportamientos menos predecibles: el método de Gradiente puede oscilar o avanzar muy lentamente debido a cambios abruptos en la pendiente, mientras que el método de Newton puede producir pasos en direcciones incorrectas, dirigirse hacia un máximo local o incluso divergir por completo. Esto se debe a que la inversión del Hessiano introduce direcciones inestables cuando existen eigenvalores negativos.

Por este motivo, en regiones no convexas es más probable que los métodos queden atrapados en mínimos locales o se comporten de manera errática, y se vuelve necesario usar técnicas globales como *Multi-start* para explorar distintas zonas del dominio.

Algoritmos recomendados

- **Métodos locales (Gradiente, Newton, Quasi-Newton):** funcionan bien en regiones convexas locales y para refinar soluciones, pero pueden quedarse atrapados en mínimos locales.
- **Métodos de búsqueda global o estrategias híbridas:** como Multi-start, Recocido Simulado, Algoritmos Genéticos, permiten explorar distintas regiones del espacio y aumentan la probabilidad de encontrar el mínimo global.

Uso de Multi-start

La estrategia Multi-start consiste en ejecutar un algoritmo de búsqueda de mínimo local desde muchos puntos iniciales distintos. Cada ejecución puede converger a un mínimo local diferente, y al comparar los valores finales se elige la mejor solución como candidato al mínimo global. Este enfoque es especialmente útil para funciones no convexas como la presente, ya que ayuda a evitar quedar atrapado en mínimos locales subóptimos. La efectividad del método depende tanto de la distribución de los puntos iniciales como de la capacidad del algoritmo local para converger rápidamente hacia un mínimo en la región donde se inicia el proceso.

2.3 Algoritmos de Optimización

Dado que la función

$$f(x, y) = x^4 - 4x^3 + 4x + y^2$$

no es convexa globalmente, es importante elegir algoritmos adecuados y considerar estrategias de múltiples puntos iniciales para buscar un mínimo global.

1. Método de Descenso por Gradiente

Descripción: El método de Gradiente es un algoritmo local que actualiza iterativamente la variable en la dirección opuesta al gradiente:

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

Recomendaciones para este problema:

- Funciona bien en regiones donde la Hessiana es definida positiva (curvatura positiva), es decir, $x > 2$ o $x < 0$.
- Puede quedar atrapado en mínimos locales si se inicia en la región $0 < x < 2$.
- Es recomendable usar **Multi-start** para explorar distintas regiones.

Pseudocódigo:

```
Elegir x0, tolerancia epsilon, paso alpha  
Para k = 0,1,2,...:  
    Calcular gradiente gk = grad(f)(xk)  
    Si ||gk|| < epsilon: terminar  
    Actualizar x_{k+1} = xk - alpha * gk  
Fin
```

2. Método de Newton

Descripción: El método de Newton utiliza información de segunda derivada (Hessiano) para actualizar la variable:

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

donde $H_k = \nabla^2 f(x_k)$.

Recomendaciones para este problema:

- Converge rápidamente en regiones donde la Hessiana es positiva definida.
- Puede fallar o divergir en regiones donde la Hessiana tiene eigenvalores negativos ($0 < x < 2$).
- Requiere calcular el Hessiano e invertirlo, por lo que es más costoso computacionalmente.

Pseudocódigo:

```

Elegir x0, tolerancia epsilon
Para k = 0,1,2,...:
    Calcular gradiente gk = grad(f)(xk)
    Si ||gk|| < epsilon: terminar
    Calcular Hessiano Hk = Hess(f)(xk)
    Resolver Hk * pk = gk
    Actualizar x_{k+1} = xk - pk
Fin

```

3. Método Quasi-Newton (BFGS)

Descripción: Quasi-Newton aproxima el Hessiano inverso iterativamente sin calcularlo explícitamente:

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

donde $W_k \approx H_k^{-1}$ se actualiza en cada iteración usando la fórmula BFGS.

Recomendaciones para este problema:

- Combina la velocidad de Newton con la simplicidad del Gradiente.
- Funciona bien localmente en regiones de curvatura positiva y es más estable que Newton si se parte de un punto alejado.
- También se recomienda usar Multi-start para aumentar la probabilidad de encontrar mínimos globales.

Pseudocódigo:

```

Elegir x0, tolerancia epsilon
Inicializar W0 = identidad
Para k = 0,1,2,...:
    Calcular gradiente gk = grad(f)(xk)
    Si ||gk|| < epsilon: terminar
    Calcular pk = Wk * gk
    Actualizar x_{k+1} = xk - pk
    Actualizar Wk+1 usando fórmula BFGS
Fin

```

3 Comparación de Resultados y Graficación

3.1 Experimentos

Se realizaron experimentos utilizando los tres algoritmos: Descenso por Gradiente, Newton y Quasi-Newton (BFGS) sobre la función

$$f(x, y) = x^4 - 4x^3 + 4x + y^2$$

partiendo de múltiples puntos iniciales (**Multi-start**) para explorar diferentes regiones de la función y evaluar la convergencia local y global.

3.2 Criterios de Comparación

Los resultados se comparan considerando:

- Número de iteraciones hasta convergencia.
- Tiempo de cómputo (si se dispone).
- Valor final de la función $f(x, y)$.
- Sensibilidad al punto inicial.
- Trayectorias de iteración en el plano (x, y) .

3.3 Observaciones

- **Descenso por Gradiente:** converge más lentamente, especialmente si el punto inicial está cerca de regiones de curvatura plana o mínima local no global.
- **Método de Newton:** converge muy rápido cerca de mínimos locales con Hessiana positiva definida, pero puede fallar si el Hessiano es indefinido.
- **Quasi-Newton (BFGS):** balancea velocidad y estabilidad, evitando el cálculo directo del Hessiano, con convergencia superlineal en regiones locales.
- **Multi-start:** aumenta significativamente la probabilidad de encontrar el mínimo global, explorando varias regiones de la función no convexa.

3.4 Graficación de la Función y Trayectorias

Para visualizar la convergencia de los algoritmos se pueden utilizar gráficos de contorno de la función y superponer la trayectoria de iteraciones de cada algoritmo:

Pseudocódigo para graficar trayectorias:

Crear una malla de puntos (X, Y) para evaluar $f(X, Y)$

Calcular $Z = f(X, Y)$

Dibujar contornos de Z

Para cada algoritmo:

 Obtener historial de puntos iterativos $[x_k, y_k]$

 Trazar línea conectando los puntos del historial

Añadir leyenda y etiquetas

3.5 Resultados Esperados

- Las trayectorias muestran cómo cada algoritmo se aproxima a diferentes mínimos locales.
- Multi-start permite que al menos uno de los algoritmos alcance el mínimo global.
- Descenso por Gradiente puede requerir más iteraciones que Newton o BFGS.

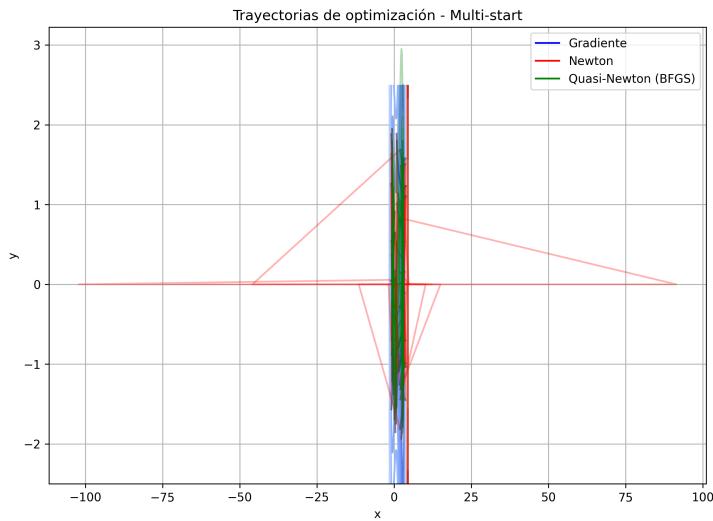


Figure 1: Contorno de $f(x, y)$ con trayectorias de los algoritmos desde múltiples puntos iniciales.

3.6 Figura Ejemplo

4 Conclusiones

- El método de Newton converge más rápido localmente, pero requiere cálculo del Hessiano.
- El descenso por gradiente es más simple, pero puede necesitar más pasos.
- Quasi-Newton (BFGS) combina velocidad y menor cálculo de derivadas.