

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ingeniería



Ascenso de máxima inclinación y Ascenso óptimo de máxima inclinación

Presentado por:

Edgar Samuel Chavarría Castañón

22055

Guatemala, octubre de 2025

Índice

1. Marco teórico	2
2. Metodología	2
3. Experimentos	4
4. Resultados	5
5. Discusión	7
6. Conclusiones	8

1 Marco teórico

El **gradiente** de una función escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es el vector que reúne sus derivadas parciales:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}.$$

Este vector apunta en la dirección en la que f crece con mayor rapidez. La razón es que la **derivada direccional** de f en una dirección unitaria u se define como

$$D_u f(x) = \nabla f(x)^\top u,$$

y su valor máximo se alcanza cuando u es paralelo a $\nabla f(x)$. Por tanto, el gradiente señala la dirección de **máxima inclinación ascendente**, mientras que $-\nabla f(x)$ indica el sentido de **descenso más pronunciado**. En optimización, esta propiedad se aprovecha para mover el punto de búsqueda hacia los máximos o mínimos de la función objetivo.

El método de **ascenso de máxima inclinación** se basa en actualizar iterativamente:

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

donde α_k es el *tamaño de paso*. Existen dos variantes principales según cómo se elija α_k :

- **Ascenso con paso fijo:** utiliza un valor constante $\alpha_k = \alpha$ en todas las iteraciones. Es simple y rápido, pero su desempeño depende del valor elegido. Si α es muy pequeño, el avance es lento; si es muy grande, el método puede divergir.
- **Ascenso con paso óptimo:** calcula un α_k diferente en cada iteración buscando maximizar $f(x_k + \alpha \nabla f(x_k))$. Esta búsqueda puede hacerse por *rejilla*, *búsqueda dorada* o mediante el método de *Newton-Raphson unidimensional*. Aunque es más costoso computacionalmente, permite un avance más estable y eficiente al adaptar el paso al comportamiento local de la función.

En resumen, el ascenso con paso fijo es más sencillo pero sensible al valor de α , mientras que el paso óptimo mejora la convergencia al costo de un mayor esfuerzo de cálculo por iteración.

2 Metodología

Función objetivo y gradiente

Se trabaja con

$$f(x, y) = \sin(x) \cos(y) + \frac{x^2 + y^2}{10}, \quad \nabla f(x, y) = \begin{bmatrix} \cos(x) \cos(y) + 0.2x \\ -\sin(x) \sin(y) + 0.2y \end{bmatrix}.$$

La actualización general en ambas variantes es

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \nabla f(\mathbf{x}_k),$$

con paro por $\|\nabla f(\mathbf{x}_k)\| < 10^{-4}$ o $k \geq 200$.

Ascenso con paso fijo (ascenso_paso_fijo)

- **Idea:** usar un tamaño de paso constante $\alpha > 0$ en todas las iteraciones: $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \nabla f(\mathbf{x}_k)$.
- **Programación:** en cada iteración se evalúa $g_k = \nabla f(\mathbf{x}_k)$, se actualiza \mathbf{x}_{k+1} y se registra la trayectoria $\{\mathbf{x}_k\}$ y los valores $\{f(\mathbf{x}_k)\}$.
- **Costo/propiedades:** muy barato por iteración, pero sensible a la elección de α (pequeño \Rightarrow lento, grande \Rightarrow posible inestabilidad).

Ascenso de máxima inclinación con paso óptimo (ascenso_paso_optimo)

En cada iteración se elige α_k que *maximiza* la función unidimensional

$$\phi(\alpha) = f(\mathbf{x}_k + \alpha \nabla f(\mathbf{x}_k)), \quad \alpha \in [0, \alpha_{\text{máx}}],$$

mediante **una** de las siguientes búsquedas lineales (seleccionable con `mode`):

- (a) **Rejilla fina** (`mode="grid"`): muestrea n puntos equiespaciados en $[0, \alpha_{\text{máx}}]$ y toma $\alpha_k \in \arg \max_{\alpha \in \{\text{rejilla}\}} \phi(\alpha)$. Es el método “a mano” del curso; robusto, con costo $\mathcal{O}(n)$ por iteración.
- (b) **Búsqueda dorada** (`mode="golden"`): reduce el intervalo manteniendo la razón áurea; no usa derivadas de ϕ y alcanza la precisión deseada con costo logarítmico en el ancho final del intervalo.
- (c) **Newton–Raphson 1D** (`mode="newton"`): usa derivadas numéricas centralizadas $\phi'(\alpha)$ y $\phi''(\alpha)$ y actualiza $\alpha^{(t+1)} = \alpha^{(t)} - \phi'(\alpha^{(t)})/\phi''(\alpha^{(t)})$, con proyección a $[0, \alpha_{\text{máx}}]$ y respaldo automático a rejilla si ϕ'' es muy pequeña. Suele requerir pocas evaluaciones cuando ϕ es bien comportada.

La función unificada registra la trayectoria, los valores de f y la secuencia $\{\alpha_k\}$, además de la norma del gradiente y la causa de paro.

Exportación y visualización. Se generan dos CSV: (i) `iteraciones_todos.csv` con el historial por iteración y (ii) `resumen_experimentos.csv` con el estado final de cada corrida. Las figuras incluyen curvas de nivel con trayectorias y superficies 3D con las rutas.

3 Experimentos

Diseño experimental

El objetivo es comparar el comportamiento entre **ascenso con paso fijo** y **ascenso con paso óptimo** (con diferentes búsquedas lineales) manteniendo la misma función objetivo y criterios de paro.

Puntos iniciales

Se evaluaron cuatro condiciones de arranque para explorar distintas regiones del paisaje de f :

$$(-3, 3), \quad (-1, 4), \quad (0, 0), \quad (2.5, -0.5).$$

Parámetros comunes

- Tolerancia de gradiente: $\varepsilon = 10^{-4}$.
- Iteraciones máximas: $k_{\text{máx}} = 200$.

Tratamientos

1. **Paso fijo**: se corrieron $\alpha \in \{0.01, 0.05, 0.10\}$ para cada punto inicial.
2. **Paso óptimo (una variante a la vez)**: se seleccionó un único esquema de búsqueda lineal por corrida:
 - **Rejilla fina**: $\alpha_{\text{máx}} = 1.0$ con $n = 200$ muestras uniformes.
 - **Búsqueda dorada**: intervalo $[0, 1.0]$, tolerancia 10^{-5} .
 - **Newton 1D**: $\alpha_0 = 0.1$, $\alpha_{\text{máx}} = 1.0$, tolerancia 10^{-6} , con respaldo a rejilla si falla la curvatura.

Métricas recogidas

Para cada combinación (método, punto inicial y parámetros) se registró:

- trayectoria $\{\mathbf{x}_k\}$, valores $\{f(\mathbf{x}_k)\}$,
- (cuando aplica) α_k seleccionado en cada iteración,
- estado final (x^y) , valor $f(x)$, $\|\nabla f(x)\|$ y motivo de paro.

Archivos y figuras

- **CSV de iteraciones**: `iteraciones_todos.csv`.
- **CSV de resumen**: `resumen_experimentos.csv`.

- **Figuras:** curvas de nivel con trayectorias comparadas y superficies 3D con la ruta de cada método (paso fijo vs. paso óptimo).

Notas sobre la función evaluada

El término cuadrático positivo de f induce crecimiento al alejarse del origen; por ello, con un intervalo de búsqueda amplio es frecuente que la maximización de $\phi(\alpha)$ seleccione el extremo permitido ($\alpha_{\text{máx}}$). Para una comparación justa, se fija el mismo $\alpha_{\text{máx}}$ en las variantes de paso óptimo y se reportan las trayectorias y estados finales bajo ese límite.

4 Resultados

Las Figuras 1, 2a y 2b muestran, respectivamente, las curvas de nivel con las trayectorias y las superficies 3D con las rutas del *ascenso con paso fijo* y del *ascenso con paso óptimo* (búsqueda dorada).

A partir de los CSV generados, se resumen los hallazgos por punto inicial (x_0, y_0) :

- $(-3, 3)$ y $(-1, 4)$: ambos métodos incrementan f sin alcanzar condición de paro por tolerancia del gradiente dentro de 200 iteraciones. Con paso fijo, el mejor de los tres pasos evaluados ($\alpha = 0.01$) llega a $f \approx 2.62$ y $f \approx 2.55$ con normas de gradiente $\|\nabla f\| \approx 0.35$ y 0.25 , respectivamente (paro por `max_iter`). Con paso óptimo (dorada), al maximizar la función lineal $\phi(\alpha)$ dentro de $[0, 1]$ el método tiende a seleccionar el extremo y la trayectoria se aleja del origen, por lo que f crece fuertemente y también termina por `max_iter`.
- $(0, 0)$: el paso fijo con $\alpha = 0.10$ alcanza la tolerancia en 128 iteraciones ($f^* \approx 1.30948$), mientras que el paso óptimo (dorada) lo hace en **8** iteraciones con el mismo valor final $f^* \approx 1.30948$ y $\|\nabla f^*\| \approx 6.9 \times 10^{-5}$ (paro por `tol_grad`).
- $(2.5, -0.5)$: el paso fijo con $\alpha = 0.05$ reduce fuertemente la norma del gradiente pero no llega a la tolerancia dentro de 200 iteraciones (paro por `max_iter`); el paso óptimo (dorada) **sí** alcanza `tol_grad` en **9** iteraciones, convergiendo también a $f^* \approx 1.30948$.

Estas cifras se reflejan en las trayectorias: cerca del origen ambas estrategias convergen hacia un punto estacionario alrededor de $(x^*, y^*) \approx (1.30948, 0)$, mientras que, partiendo más lejos, las rutas se orientan hacia regiones de mayor amplitud donde f crece sin cota y no se satisface la tolerancia antes del límite de iteraciones.

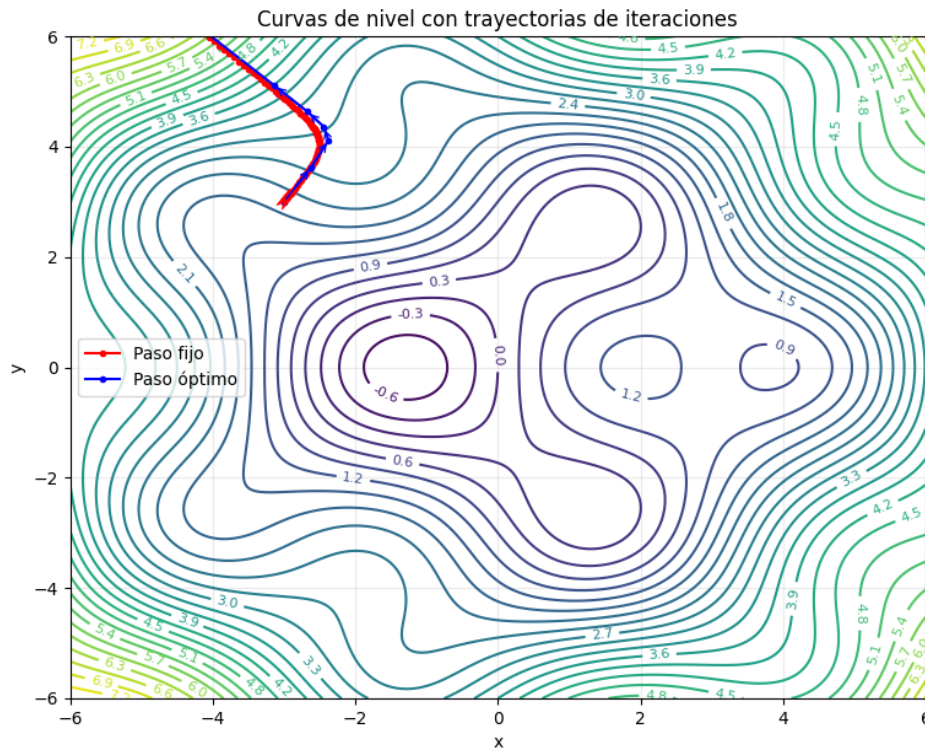


Figura 1: *Curvas de nivel de f con trayectorias (paso fijo vs. paso óptimo).*

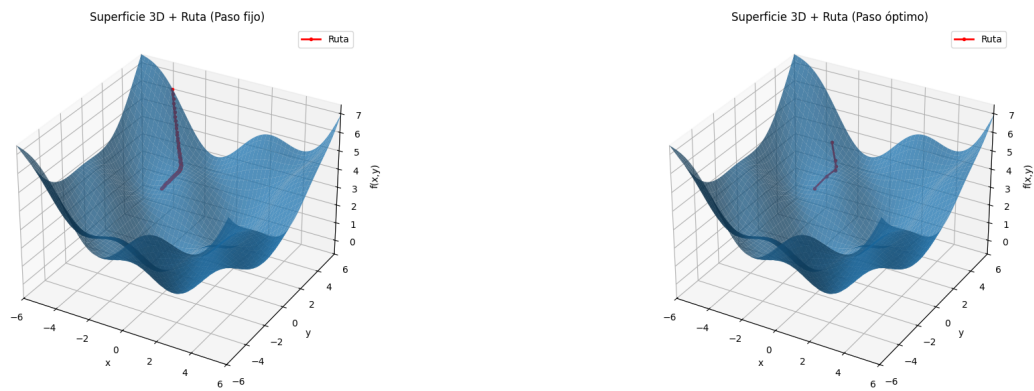


Figura 2: *Superficie 3D de f con la ruta de cada método.*

5 Discusión

Efecto del tamaño de paso. Con **paso fijo**, el desempeño depende fuertemente de α : valores pequeños ($\alpha = 0.01$) avanzan de forma estable pero lenta; valores grandes ($\alpha = 0.10$) pueden acelerar el progreso cerca de un máximo local (caso $(0,0)$), aunque no garantizan éxito desde estados alejados. Esta sensibilidad explica por qué en $(-3, 3)$ y $(-1, 4)$ ningún α alcanzó `tol_grad` en el presupuesto de 200 iteraciones.

Ventaja del paso óptimo cercano a máximos locales. El **paso óptimo** recalcula α_k en cada iteración mediante búsqueda lineal y, cuando la topología local de f es “benigna”, reduce drásticamente el número de iteraciones. En $(0,0)$ y $(2.5, -0.5)$ la búsqueda dorada alcanza la tolerancia en **8–9** iteraciones, contra más de cien del paso fijo, convergiendo a un mismo punto estacionario con $f^* \approx 1.30948$. Esto confirma la ganancia de eficiencia de la selección adaptativa de paso.

Saturación en el extremo del intervalo. Para estados iniciales *lejanos*, el término cuadrático positivo de f ($0.1(x^2 + y^2)$) hace que la función unidimensional $\phi(\alpha) = f(x_k + \alpha \nabla f(x_k))$ sea frecuentemente monótona creciente en $[0, \alpha_{\text{máx}}]$. En consecuencia, la búsqueda lineal tiende a escoger $\alpha_k = \alpha_{\text{máx}}$, empujando el iterado fuera de la zona representada en las curvas de nivel y llevando a paros por `max_iter`. Este comportamiento no es un “fallo” del algoritmo sino una propiedad de la función: sin cotas o regularización, el ascenso seguirá aumentando f .

Implicaciones prácticas.

- Cuando se busca *comparar métodos*, el paso óptimo evidencia su fortaleza cerca de regiones con curvatura moderada: menos iteraciones para lograr la misma tolerancia.
- Si el objetivo es controlar la magnitud de los iterados, conviene acotar el intervalo de búsqueda ($\alpha_{\text{máx}}$) o incorporar criterios tipo Armijo/Wolfe (backtracking) para evitar que α_k se pegue sistemáticamente al extremo.
- Para funciones con crecimiento no acotado, la comparación debe hacerse bajo un mismo $\alpha_{\text{máx}}$ y con métricas homogéneas (norma del gradiente, motivo de paro, valor final de f), como se reporta.

En síntesis, el **paso fijo** es sencillo pero requiere calibración del parámetro; el **paso óptimo** reduce iteraciones y dependencia de hiperparámetros cuando la búsqueda lineal percibe una curvatura favorable, aunque puede acelerar la salida del dominio de interés si f crece sin cota y $\alpha_{\text{máx}}$ es amplio. Las trayectorias y los CSV confirman estas conclusiones en los cuatro puntos iniciales evaluados.

6 Conclusiones

- El **ascenso con paso fijo** es simple y barato por iteración, pero su desempeño depende fuertemente del valor de α : valores pequeños producen avances estables pero lentos, y valores grandes pueden acelerar el progreso sólo en zonas favorables, sin garantizar alcanzar la tolerancia antes del límite de iteraciones.
- El **ascenso con paso óptimo reduce drásticamente** el número de iteraciones cuando se parte cerca de una región con curvatura moderada. En los experimentos, desde $(0, 0)$ alcanzó `tol_grad` en **8–9** iteraciones, mientras que el paso fijo requirió ≥ 128 o no llegó dentro de 200 iteraciones, convergiendo ambos a $f^* \approx 1.30948$.
- Para puntos iniciales **lejanos**, la función usada crece sin cota por el término cuadrático positivo; por ello la función unidimensional $\phi(\alpha)$ suele ser monótona en $[0, \alpha_{\text{máx}}]$ y la búsqueda lineal elige el $\alpha_{\text{máx}}$. En estas condiciones, ambos métodos tienden a terminar por `max_iter` si no se restringe el intervalo de búsqueda.
- La **selección del método de búsqueda lineal** impacta el costo por iteración: **grid** es robusto y predecible (costo $\mathcal{O}(n)$), **golden** no requiere derivadas y es eficiente en reducción de intervalo, y **newton** puede ser el más rápido en evaluaciones cuando ϕ es bien comportada, aunque es más sensible al punto inicial y requiere respaldo.
- Para una **comparación justa** entre variantes de paso óptimo, es clave fijar el mismo $\alpha_{\text{máx}}$ y reportar métricas homogéneas (norma del gradiente, motivo de paro y valor final de f); esto se reflejó coherentemente en las trayectorias y en los CSV obtenidos.

Referencias

1. Nocedal, J., & Wright, S. J. (2006). *Numerical Optimization* (2nd ed.). Springer. <https://www.math.uci.edu/~qnie/Publications/NumericalOptimization.pdf>
2. Fletcher, R. (1987). *Practical Methods of Optimization* (2nd ed.). John Wiley & Sons.
3. Burden, R. L., & Faires, J. D. (2011). *Análisis numérico* (9ª ed.). Cengage Learning. (Capítulos sobre métodos de búsqueda unidimensional y gradiente).
4. Chong, E. K. P., & Żak, S. H. (2013). *An Introduction to Optimization* (4th ed.). Wiley.