Trabajo Investigativo de Optimización

Amanda Medina Solis
Grupo: C312

1. Modelo a Analizar

El modelo a estudiar es la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por:

$$f(x, y) = x^2 + y^2 - 0{,}12 \cos(3\pi x) \cos(4\pi y) + 0{,}3$$

Se puede encontrar la graficación del modelo en graph.py.

2. Dominio

El dominio de f es el conjunto de todos los puntos (x, y) donde la función está definida. Dado que los términos x^2 e y^2 son polinomios definidos para todo $\mathbb{R}$, las funciones $\cos(3\pi x)$ y $\cos(4\pi y)$ están definidas para todo $\mathbb{R}$, y las operaciones de suma y multiplicación son cerradas en $\mathbb{R}$.

Llegamos a que:

$$D(f) = \mathbb{R}^2.$$

3. Continuidad

Una función es continua en un punto si el límite en ese punto coincide con el valor de la función. Formalmente, f es continua en (a, b) si:

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b).$$

Dado que las funciones x^2 e y^2 son continuas en $\mathbb{R}^2$, las funciones $\cos(3\pi x)$ y $\cos(4\pi y)$ son continuas en $\mathbb{R}^2$, y la suma y producto de funciones continuas es continua.

Se puede decir que: f es continua en todo $\mathbb{R}^2$.

4. Diferenciabilidad

Una función es diferenciable en un punto si existen sus derivadas parciales y son continuas en un entorno de ese punto.

Calculamos las derivadas parciales de primer orden:

$$\frac{\partial f}{\partial x} = 2x + 0{,}36\pi \sin(3\pi x) \cos(4\pi y)$$

$$\frac{\partial f}{\partial y} = 2y + 0{,}48\pi \cos(3\pi x) \sin(4\pi y)$$

Ambas derivadas son combinaciones de polinomios y funciones trigonométricas, las cuales son continuas en $\mathbb{R}^2$.

Conclusión: f es diferenciable en todo $\mathbb{R}^2$.

5. Análisis de Convexidad con Hessiana

Una función es convexa si su matriz Hessiana es semidefinida positiva en todo su dominio. Para una función $f : \mathbb{R}^2 \to \mathbb{R}$ de clase C^2, esto equivale a que todos los valores propios de la matriz Hessiana sean no negativos en todo punto.

Calculamos las derivadas parciales de segundo orden:

$$\frac{\partial^2 f}{\partial x^2} = 2 + 1{,}08\pi^2 \cos(3\pi x)\cos(4\pi y)$$
$$\frac{\partial^2 f}{\partial y^2} = 2 + 1{,}92\pi^2 \cos(3\pi x)\cos(4\pi y)$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 1{,}44\pi^2 \sin(3\pi x)\sin(4\pi y)$$

La matriz Hessiana es:

$$H(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 + 1{,}08\pi^2 \cos(3\pi x)\cos(4\pi y) & 1{,}44\pi^2 \sin(3\pi x)\sin(4\pi y) \\ 1{,}44\pi^2 \sin(3\pi x)\sin(4\pi y) & 2 + 1{,}92\pi^2 \cos(3\pi x)\cos(4\pi y) \end{bmatrix}$$

Criterio de Valores Propios para Convexidad: Para que H sea definida positiva (y por tanto f sea estrictamente convexa) en un punto, todos sus valores propios deben ser positivos. Para que sea semidefinida positiva (y f sea convexa), todos sus valores propios deben ser no negativos.

Los valores propios λ de H son:

$$\lambda = 2 + \frac{3\pi^2}{50}\left(25\cos(3\pi x)\cos(4\pi y) \pm \sqrt{49\cos^2(3\pi x)\cos^2(4\pi y) + 576\sin^2(3\pi x)\sin^2(4\pi y)}\right)$$

Luego, para que H sea definida positiva, se debe cumplir la condición de que ambos valores propios de H sean positivos.

Sin embargo, existe al menos un valor propio negativo en dichos puntos, ya que la traza de H (suma de valores propios) sería negativa en esas regiones.

Conclusión: La matriz Hessiana no es definida positiva en todo $\mathbb{R}^2$, ya que existen puntos donde tiene valores propios negativos. Por tanto, f no es convexa globalmente.

6. Determinación del Mínimo Global

Puntos Estacionarios:

Resolvemos el sistema $\nabla f(x,y) = 0$:

$$\frac{\partial f}{\partial x} = 2x + 0{,}36\pi \sin(3\pi x)\cos(4\pi y) = 0$$
$$\frac{\partial f}{\partial y} = 2y + 0{,}48\pi \cos(3\pi x)\sin(4\pi y) = 0$$

Debido a la periodicidad de las funciones trigonométricas $\cos(3\pi x)$ y $\cos(4\pi y)$, existen **múltiples puntos estacionarios** en el plano. Sin embargo, demostraremos que solo $(0,0)$ es el mínimo global.

Verificación del Mínimo Global

Establecer una cota inferior. Partimos de la función:

$$f(x,y) = x^2 + y^2 - 0{,}12\cos(3\pi x)\cos(4\pi y) + 0{,}3$$

Sabemos que el coseno está acotado: $-1 \leq \cos(3\pi x)\cos(4\pi y) \leq 1$

Por lo tanto:

$$-0{,}12 \leq -0{,}12\cos(3\pi x)\cos(4\pi y) \leq 0{,}12$$

Sumando los términos:

$$f(x,y) \geq x^2 + y^2 - 0{,}12 + 0{,}3 = x^2 + y^2 + 0{,}18$$

Analizar el caso de igualdad: Queremos encontrar cuándo se alcanza el valor mínimo posible. De la desigualdad anterior:

$$f(x,y) \geq 0{,}18 \quad \text{para todo } (x,y) \in \mathbb{R}^2$$

La igualdad $f(x,y) = 0{,}18$ ocurre si y solo si:

1. $x^2 + y^2 = 0$ (es decir, $x = 0$ e $y = 0$)
2. $-0{,}12\cos(3\pi x)\cos(4\pi y) = -0{,}12$ (es decir, $\cos(3\pi x)\cos(4\pi y) = 1$)

Verificar que solo $(0,0)$ cumple ambas condiciones

- **Condición 1:** $x = 0$ e $y = 0$
- **Condición 2:** $\cos(3\pi \cdot 0)\cos(4\pi \cdot 0) = \cos(0)\cos(0) = 1 \cdot 1 = 1$

El punto $(0,0)$ satisface ambas condiciones simultáneamente, por lo tanto:

$$f(0,0) = 0^2 + 0^2 - 0{,}12(1) + 0{,}3 = 0{,}18$$

Además tenemos que ningún otro punto puede alcanzar este valor
Para cualquier otro punto $(x,y) \neq (0,0)$, tenemos dos casos:

- Caso A: Si $x^2 + y^2 > 0$, entonces:

$$f(x,y) \geq x^2 + y^2 + 0{,}18 > 0{,}18$$

- Caso B: Si $x = 0$ e $y = 0$ no se cumplen simultáneamente, entonces $x^2 + y^2 > 0$, volviendo al Caso A.

Comportamiento asintótico:
Cuando $\|(x,y)\| \to \infty$, el término $x^2 + y^2$ domina:

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

Esto garantiza que no existen mínimos globales en el infinito.

Conclusión: El punto $(0,0)$ es el **único mínimo global** de la función con valor $f(0,0) = 0{,}18$. Cualquier otro punto estacionario que exista debido a la periodicidad será un mínimo local, punto de silla o máximo local, pero ninguno podrá tener un valor menor que 0,18.

7. Aproximación de solución mediante algoritmos

7.1. Método de Máximo Descenso

El Método de Máximo Descenso es un algoritmo de optimización iterativo de primer orden que busca el mínimo de una función multidimensional. Para una función $f : \mathbb{R}^n \to \mathbb{R}$, el algoritmo se fundamenta en la propiedad de que el gradiente $\nabla f(\mathbf{x})$ apunta en la dirección de máximo crecimiento local, por lo que moverse en dirección opuesta $(-\nabla f(\mathbf{x}))$ conduce hacia un mínimo local.

Fundamentos Matemáticos

- Dirección de descenso: $\mathbf{d}_k = -\nabla f(\mathbf{x}_k)$
- Actualización: $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$
- Condición de parada: $\|\nabla f(\mathbf{x}_k)\| < \varepsilon$

Algoritmo Implementado
El algoritmo se implementó en el archivo `gradientDescent.py`. Salida Automática:

- Tablas numeradas en consola
- Gráficos guardados como PNG
- Análisis estadístico final

Experimentos:
Se diseñaron dos conjuntos principales de experimentos:
Prueba 1: Diferentes Tamaños de Paso
Análisis de la Prueba 1:

Tamaño Paso (α)	Iteraciones	$f(x, y)$ final	Error	Estado
0.01	150	0.328860	$1,5 \times 10^{-1}$	Lento, inestable
0.05	29	0.328860	$1,5 \times 10^{-1}$	Óptimo
0.10	58	0.328860	$1,5 \times 10^{-1}$	Convergencia lenta
0.15	-	-	-	Divergencia
0.20	-	-	-	Divergencia
0.30	-	-	-	Divergencia
0.50	2	0.180000	$0,0 \times 10^{0}$	Muy rápido

Cuadro 1: Resultados para diferentes tamaños de paso (punto inicial: (1,1))

- Rango óptimo: $\alpha \in [0{,}05, 0{,}5]$ proporciona la convergencia más rápida
- Límite inferior: $\alpha = 0{,}01$ es ineficiente (150 iteraciones)
- Límite superior: $\alpha = 0{,}15, 0{,}20, 0{,}30$ divergen
- Zona de estabilidad: $\alpha \in [0{,}05, 0{,}10]$ garantiza convergencia
- Mejor caso: $\alpha = 0{,}5$ converge en 2 iteraciones al mínimo global

Prueba 2: Diferentes Puntos Iniciales

Punto Inicial	Distancia	Iteraciones	Error	Evaluación
(1.0, 1.0)	1.41	58	$1,5 \times 10^{-1}$	Aceptable
(2.0, 2.0)	2.83	43	$2,3 \times 10^{-1}$	Aceptable
(-1.0, 1.0)	1.41	58	$1,5 \times 10^{-1}$	Aceptable
(0.5, -0.5)	0.71	54	$1,5 \times 10^{-1}$	Aceptable
(3.0, -2.0)	3.61	61	$1,5 \times 10^{-1}$	Aceptable
(-2.0, -2.0)	2.83	43	$2,3 \times 10^{-1}$	Aceptable
(1.5, -1.5)	2.12	42	$2,3 \times 10^{-1}$	Aceptable
(-1.5, 2.0)	2.50	43	$2,3 \times 10^{-1}$	Aceptable
(2.5, 0.5)	2.55	18	$3,7 \times 10^{-1}$	Excelente

Cuadro 2: Resultados para diferentes puntos iniciales ($\alpha = 0{,}1$)

Análisis de la Prueba 2:

- Robustez: 100 % de éxito para todos los puntos iniciales
- Relación distancia-iteraciones: Los puntos más cercanos no necesariamente convergen más rápido
- Eficiencia: Punto (2.5, 0.5) muestra la mejor eficiencia (18 iteraciones)
- Consistencia: Error significativo en todos los casos, indicando convergencia a mínimos locales
- Comportamiento inesperado: Puntos con mayor distancia pueden converger más rápido

Conclusiones de los Experimentos:

1. Eficiencia variable: El método muestra alta eficiencia solo con $\alpha = 0{,}5$ (2 iteraciones)
2. Robustez limitada: Aunque todos los puntos convergen, la mayoría lo hace a mínimos locales
3. Sensibilidad alta: Alta sensibilidad al tamaño de paso, con 3 de 7 valores divergiendo
4. Confiabilidad moderada: 81.2 % de tasa de éxito, pero con errores significativos

Métrica	**Valor**
Total de pruebas ejecutadas	16
Pruebas exitosas	13
Tasa de éxito global	81.2 %
Iteraciones promedio	50.7
Iteraciones mínimas	2
Iteraciones máximas	150
Error promedio	$1,78 \times 10^{-1}$
Desviación estándar	33.2

Cuadro 3: Resumen estadístico del análisis

7.2. Método de Región de Confianza

El método de región de confianza es un algoritmo iterativo de optimización no lineal que utiliza un modelo local (generalmente cuadrático) para aproximar la función objetivo dentro de una región donde el modelo es considerado *confiable*.

Algoritmo Implementado

El algoritmo se implementó en el archivo `trustRegion.py`.

Explicación del método:

El modelo cuadrático local alrededor del punto x_k se define como:

$$\hat{f}(x_k + h) = f(x_k) + \nabla f(x_k)^T h + \frac{1}{2} h^T M h$$

donde:

- $f(x_k)$ es el valor de la función objetivo en el punto actual
- $\nabla f(x_k)$ es el vector gradiente en x_k
- M es una matriz simétrica que aproxima la Hessiana
- h es el paso a determinar dentro de la región de confianza

El subproblema:

$$\min_h \hat{f}(x_k + h) \quad \text{sujeto a} \quad \|h\|^2 \leq \Delta_k^2$$

puede resolverse mediante varios métodos:

- **Método del Paso de Cauchy**: Dirección de máximo descenso dentro de la región
- **Método Dogleg**: Combinación de dirección de Newton y gradiente
- **Método de Solución Exacta**: Usando descomposición espectral

Ratio de Calidad R_k

El ratio R_k mide la precisión del modelo:

$$R_k = \frac{\text{Reducción real}}{\text{Reducción predicha}} = \frac{f(x_k) - f(x_{k+1})}{\hat{f}(x_k) - \hat{f}(x_{k+1})}$$

La interpretación es:

- $R_k \approx 1$: Modelo excelente, puede expandirse la región
- $R_k > m_1$ (típicamente $m_1 = 0{,}75$): Modelo bueno
- $m_2 < R_k < m_1$ (típicamente $m_2 = 0{,}25$): Modelo aceptable
- $R_k < m_2$: Modelo pobre, contraer región
- $R_k \leq 0$: Paso no mejora la función, debe rechazarse

Tamaño Región (Δ)	**Iteraciones**	**$f(x, y)$ final**	**Error**	**Estado**
0.1	16	0.180000	$0,0 \times 10^0$	Bueno
0.3	16	0.328860	$1,5 \times 10^{-1}$	Bueno
0.5	15	0.328860	$1,5 \times 10^{-1}$	Excelente
1.0	9	0.328860	$1,5 \times 10^{-1}$	Excelente
1.5	10	0.180000	$0,0 \times 10^0$	Excelente
2.0	10	0.328860	$1,5 \times 10^{-1}$	Excelente
3.0	11	0.180000	$0,0 \times 10^0$	Excelente

Cuadro 4: Resultados para diferentes tamaños de región de confianza (punto inicial: (1,1))

Experimentos
Prueba 1: Diferentes Tamaños de Región de Confianza Inicial

Análisis de la Prueba 1:

- Rango óptimo: $\Delta \in [0{,}5, 3{,}0]$ proporciona convergencia eficiente (9-15 iteraciones)
- Comportamiento mixto: Algunos tamaños convergen al mínimo global ($f = 0{,}18$), otros a mínimos locales ($f = 0{,}32886$)
- Eficiencia variable: $\Delta = 1{,}0$ es el más rápido (9 iteraciones) pero converge a mínimo local
- Convergencia global: $\Delta = 0{,}1; 1{,}5; 3{,}0$ alcanzan el mínimo global con 10-16 iteraciones
- Mejor equilibrio: $\Delta = 1{,}5$ combina velocidad (10 iteraciones) y convergencia global

Prueba 2: Diferentes Puntos Iniciales

Punto Inicial	**Distancia**	**Iteraciones**	**Error**	**Evaluación**
(1.0, 1.0)	1.41	9	$1,5 \times 10^{-1}$	Muy rápido
(2.0, 2.0)	2.83	15	$0,0 \times 10^0$	Excelente
(-1.0, 1.0)	1.41	10	$1,5 \times 10^{-1}$	Muy rápido
(0.5, -0.5)	0.71	5	$3,7 \times 10^{-1}$	Muy rápido
(3.0, -2.0)	3.61	17	$1,5 \times 10^{-1}$	Bueno
(-2.0, -2.0)	2.83	16	$1,5 \times 10^{-1}$	Excelente
(1.5, -1.5)	2.12	12	$1,5 \times 10^{-1}$	Excelente
(-1.5, 2.0)	2.50	14	$2,3 \times 10^{-1}$	Excelente
(2.5, 0.5)	2.55	16	$1,5 \times 10^{-1}$	Excelente

Cuadro 5: Resultados para diferentes puntos iniciales ($\Delta = 1{,}0$)

Análisis de la Prueba 2:

- Robustez: 100 % de convergencia para todos los puntos iniciales
- Convergencia mixta: Solo (2.0, 2.0) alcanza el mínimo global; otros convergen a mínimos locales
- Eficiencia variable: Rango de 5-17 iteraciones, con promedio de 12.7
- Comportamiento inesperado: Puntos cercanos no necesariamente convergen más rápido
- Mejor caso: (0.5, -0.5) con solo 5 iteraciones (pero a mínimo local)
- Caso óptimo: (2.0, 2.0) alcanza mínimo global en 15 iteraciones

Conclusiones de los Experimentos

- El método de región de confianza demuestra superioridad en eficiencia (Promedio de 12.6 iteraciones vs 50.7 del gradiente descendente) y robustez (100 % de tasa de éxito vs 81.2 % del gradiente descendente)
- El método ajusta automáticamente el tamaño de región

Métrica	**Valor**
Total de pruebas ejecutadas	16
Pruebas exitosas	16
Tasa de éxito global	100.0 %
Iteraciones promedio	12.6
Iteraciones mínimas	5
Iteraciones máximas	17
Error promedio	$1,30 \times 10^{-1}$
Desviación estándar	3.4
Precisión alcanzada	27.7 %
Casos altamente eficientes (≤15 iteraciones)	11/16

Cuadro 6: Resumen estadístico del análisis - Región de Confianza