Tarea 11 - Optimización

Erick Salvador Alvarez Valencia

CIMAT A.C.,
erick.alvarez@cimat.mx

Resumen En el presente reporte se mostrará la resolución para los ejercicios descritos en la tarea 11 de la materia de optimización los cuales son referentes a problemas sobre multiplicadores de Langrande así como un ejercicio de mínimos cuadrados no lineales donde se usa la implementación de la tarea 10 pero esta vez se usa una aproximación de la matriz Jacobiana mediante diferencias finitas. Para este último ejercicio se mostrarán los resultados obtenidos al usar diferentes valores del parámetro h.

Keywords: Multiplicadores de Lagrange, Mínimos cuadrados no lineales.

1. Ejercicio 1

La temperatura sobre el elipsoide $4x^2+y^2+4z^2 = 16$ está dada por $T(x, y, z) = 8x^2 + 4yz - 16z + 600$ Encuentre el punto más caliente sobre el elipsoide.

Solución : Primero tenemos que reescribir el problema de optimización de la siguiente manera:

$$\begin{aligned} &\min_{(x,y,z)} T(x, y, z) = 8x^2 + 4yz - 16z + 600 \\ &s.a\ 4x^2 + y^2 + 4z^2 - 16 = 0 \end{aligned} \tag{1}$$

Por lo que empezamos a definir la función Lagrangiana.

$$\mathcal{L}(x, y, z; \lambda) = 8x^2 + 4yz - 16z + 600 - \lambda(4x^2 + y^2 + 4z^2 - 16) \tag{2}$$

Obtenemos las derivadas parciales de dicha función (gradiente) y las igualamos a cero para tener el sistema de ecuaciones.

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \lambda} = -(4x^2+y^2+4z^2-16) = 0 \rightarrow 4x^2+y^2+4z^2-16 = 0 \ 1) \\
\frac{\partial \mathcal{L}}{\partial x} = 16x - \lambda 8x = 0 \ 2) \\
\frac{\partial \mathcal{L}}{\partial y} = 4z - \lambda 2y = 0 \ 3) \\
\frac{\partial \mathcal{L}}{\partial z} = 4y - 16 - \lambda 8z = 0 \ 4)
\end{aligned} \tag{3}
$$

Para un mejor entendimiento, se asignó un número a las ecuaciones y se hará referencias a ellas por medio de los mismos.
Primero trabajamos con la ecuación 2).

$$
16x - \lambda 8x = 0 \rightarrow (16 - \lambda 8)x = 0 \tag{4}
$$

De aquí tenemos dos casos. El primero es $16 - \lambda 8 = 0$ y el segundo es $x = 0$. Empezamos trabajando con el primero.

$$
16 - \lambda 8 = 0 \rightarrow \lambda = 2 \tag{5}
$$

Sust. λ en 3).

$$
4z = 4y \rightarrow z = y \tag{6}
$$

Sust. y en 4).

$$
\begin{aligned}
4z - 16 - 16z = 0 \\
-12z = 16 \rightarrow z = \frac{-4}{3},\ y = \frac{-4}{3}
\end{aligned} \tag{7}
$$

Sust. x, y, λ en 1).

$$
\begin{aligned}
4x^2 + \frac{16}{9} + \frac{64}{9} - 16 = 0 \\
x = \pm\frac{4}{3}
\end{aligned} \tag{8}
$$

Para el presente caso se encontraron los siguientes valores: $x = \pm\frac{4}{3}$, $y = z = \frac{-4}{3}$, $\lambda = 2$. Ahora se trabajará el segundo caso, donde $x = 0$.
De la ecuación 3) despejamos λ.

$$
\begin{aligned}
4z = \lambda 2y \\
\lambda = \frac{2z}{y}
\end{aligned} \tag{9}
$$

Ahora sustituimos λ en 4)

$$\begin{aligned} 4y-16-8(\frac{2z}{y})z &= 0 \\ y-4-\frac{4z^2}{y} &= 0 \\ \frac{-4z^2}{y} &= 4-y \\ z^2 &= \frac{y^2-4y}{4} \\ z &= \pm\sqrt{\frac{y^2-4y}{4}} \end{aligned} \tag{10}$$

Se ha encontrado el valor de z^2 y por ende el valor de z. Ahora sustituimos z^2 en la ecuación 1). Hay que recordar que $x=0$.

$$\begin{aligned} y^2+4(\frac{y^2-4y}{4}) &= 16 \\ y^2-2y &= 8 \\ y^2-2y+1 &= 9 \\ (y-1)^2 &= 9 \\ y &= \pm3+1 \\ y_1=4,\ y_2 &= -2 \end{aligned} \tag{11}$$

Hemos encontrado 2 valores para y, estos mismos los sustituimos en la ecuación encontrada para z. Primero con $y=4$.

$$z=\pm\sqrt{\frac{4^2-4*4}{4}}=0 \tag{12}$$

Tenemos que para $y=4$, $z=0$, y por lo tanto $\lambda=0$.
Ahora sustituimos para $y=-2$.

$$z=\pm\sqrt{\frac{(-2)^2-4*(-2)}{4}}=\pm3 \tag{13}$$

Tenemos que para $y=-2$, $z=\pm3$, y por lo tanto $\lambda=\pm3$. Finalmente tenemos los siguientes valores encontrados en el sistema de ecuaciones, de igual manera se mostrará el valor de la función $T(x,y,z)$ evaluada en dichos puntos.

1. $x=\frac{4}{3}$, $y=z=\frac{-4}{3}$, $\lambda=2$, $T(x,y,z)=\frac{1928}{3}\approx 642{,}66$

2. $x = \frac{-4}{3}$, $y = z = \frac{-4}{3}$, $\lambda = 2$, $T(x, y, z) = \frac{1928}{3} \approx 642{,}66$

3. $x = 0$, $y = 4$, $z = 0$, $\lambda = 0$, $T(x, y, z) = 600$

4. $x = 0$, $y = -2$, $z = \sqrt{3}$, $\lambda = -\sqrt{3}$, $T(x, y, z) = 600 - 24\sqrt{3} \approx 558{,}43$

5. $x = 0$, $y = -2$, $z = -\sqrt{3}$, $\lambda = \sqrt{3}$, $T(x, y, z) = 600 + 24\sqrt{3} \approx 641{,}569$

Una vez vistas las evaluaciones de la función con todos los puntos encontrados en el sistema de ecuaciones, se ha determinado que los puntos que maximizan la misma son: $x = \frac{\pm 4}{3}$, $y = \frac{-4}{3}$, $z = \frac{-4}{3}$. Probablemente los otros puntos sean mínimos locales o puntos silla, se tendrían que verificar condiciones de optimalidad para determinar esto.

2. Ejercicio 2

Encuentre el valor máximo de la función $f(x, y, z) = xyz$ cuando el punto (x, y, z) está sobre los planos $x + y + z = 40$ y $x + y = z$

Solución : Primero tenemos que reescribir el problema de optimización de la siguiente manera:

$$\begin{aligned} \min_{(x,y,z)} f(x,y,z) &= xyz \\ s.a\ x + y + z - 40 &= 0 \\ s.a\ x + y - z &= 0 \end{aligned} \tag{14}$$

Por lo que empezamos a definir la función Lagrangiana.

$$\mathcal{L}(x, y, z; \lambda) = xyz - \lambda_1(x + y + z - 40) - \lambda_2(x + y - z) \tag{15}$$

Obtenemos las derivadas parciales de dicha función (gradiente) y las igualamos a cero para tener el sistema de ecuaciones.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \lambda_1} = -(x + y + z - 40) = 0 \rightarrow x + y + z - 40 = 0\ 1) \\ \frac{\partial \mathcal{L}}{\partial \lambda_2} = -(x + y - z) = 0 \rightarrow x + y - z = 0\ 2) \\ \frac{\partial \mathcal{L}}{\partial x} = yz - \lambda_1 - \lambda_2 = 0\ 3) \\ \frac{\partial \mathcal{L}}{\partial y} = xz - \lambda_1 - \lambda_2 = 0\ 4) \\ \frac{\partial \mathcal{L}}{\partial z} = xy - \lambda_1 + \lambda_2 = 0\ 5) \end{aligned} \tag{16}$$

Al igual que el ejercicio pasado, las ecuaciones fueron numeradas para un fácil referenciado.
Lo que se hará primero es tomar la ecuación 4) y restarle la 3)

$$\begin{aligned} yz - \lambda_1 - \lambda_2 &= 0\ - \\ xz - \lambda_1 - \lambda_2 &= 0 \\ \rightarrow (y-x)z &= 0 \end{aligned} \tag{17}$$

De lo anterior nos queda $(y-x)z = 0$ por lo que se generan dos casos. Primero vamos a tratar el caso donde $y - x = 0$.

$$\begin{aligned} y - x &= 0 \\ y &= x \end{aligned} \tag{18}$$

De la ecuación 2) sustituimos x.

$$\begin{aligned} x + x - z &= 0 \\ z &= 2x \end{aligned} \tag{19}$$

Sustituimos todo lo anterior en la ecuación 1).

$$\begin{aligned} x + x + 2x &= 40 \\ x = 10,\ y = 10,\ z &= 20 \end{aligned} \tag{20}$$

Ahora sumamos las ecuaciones 3) y 5) para eliminar λ_2 y posteriormente sustituimos los valores de x, y, z.

$$\begin{aligned} 200 - \lambda_1 - \lambda_2 &= 0\ - \\ 100 - \lambda_1 + \lambda_2 &= 0 \\ \rightarrow 300 - 2\lambda_1 = 0 \rightarrow \lambda_1 &= 150 \end{aligned} \tag{21}$$

Sustituimos el valor de λ_1 en la ecuación 3)

$$\begin{aligned} 200 - 150 &= \lambda_2 \\ \lambda_2 &= 50 \end{aligned} \tag{22}$$

Hemos llegado a una solición del sistema de ecuaciones, esta para el caso donde $x - y = 0$. Ahora veremos el otro caso donde $z = 0$. Así que de la ecuación 2) sustituimos z.

$$\begin{aligned} x + y &= 0 \\ x &= -y \end{aligned} \tag{23}$$

Sustituimos y, z en la ecuación 1)

$$\begin{aligned} x - x &= 40 \\ 0 &= 40 \end{aligned} \tag{24}$$

Hemos llegado a una contradicción, donde 0 no puede ser igual a 40, por lo cual z no puede valer 0.
Finalmente hemos llegado a que el punto de solución del sistema de ecuaciones y que maximiza la función es: $x = y = 10$, $z = 20$, $\lambda_1 = 150$ y $\lambda_2 = 50$. El valor de la función en los puntos encontrados es $f(x, y, z) = 2000$. Para comprobar que estos valores son máximos locales tendríamos que comprobar con las condiciones de optimalidad.

3. Ejercicio 3

Encuentre el valor máximo de la función $f(x, y, z) = x^2 + 2y - z^2$ sujeto a las restricciones $2x - y = 0$ y $y + z = 0$.

Solución : Primero tenemos que reescribir el problema de optimización de la siguiente manera:

$$\begin{aligned} \min_{(x,y,z)} f(x, y, z) &= x^2 + 2y - z^2 \\ s.a\ 2x - y &= 0 \\ s.a\ y + z &= 0 \end{aligned} \tag{25}$$

Por lo que empezamos a definir la función Lagrangiana.

$$\mathcal{L}(x, y, z; \lambda) = x^2 + 2y - z^2 - \lambda_1(2x - y) - \lambda_2(y + z) \tag{26}$$

Obtenemos las derivadas parciales de dicha función (gradiente) y las igualamos a cero para tener el sistema de ecuaciones.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \lambda_1} = -(2x - y) = 0 \to 2x - y = 0 \to y = 2x\ 1) \\ \frac{\partial \mathcal{L}}{\partial \lambda_2} = -(y + z) = 0 \to y + z = 0\ 2) \\ \frac{\partial \mathcal{L}}{\partial x} = 2x - 2\lambda_1 = 0 \to x = \lambda_1\ 3) \\ \frac{\partial \mathcal{L}}{\partial y} = 2 + \lambda_1 - \lambda_2 = 0\ 4) \\ \frac{\partial \mathcal{L}}{\partial z} = -2z - \lambda_2 = 0\ \to z = \frac{-\lambda_2}{2}\ 5) \end{aligned} \tag{27}$$

Al igual que el ejercicio pasado, las ecuaciones fueron numeradas para un fácil referenciado.
Lo primero que podemos ver es que de la ecuación 1) sustituimos x por λ_1 y sustituir lo obtenido en conjunto con la ecuación 5) en la ecuación 2)

$$\begin{aligned} 2\lambda_1 - \frac{\lambda_2}{2} &= 0 \\ \lambda_2 &= 4\lambda_1 \end{aligned} \tag{28}$$

Con esto anterior ya tenemos una expresión para λ_2 en términos de λ_1. Sustituimos lo anterior en la ecuación 4)

$$\begin{aligned} 2 + \lambda_1 - 4\lambda_1 &= 0 \\ -3\lambda_1 &= -2 \\ \lambda_1 = \frac{2}{3};\ \lambda_2 &= \frac{8}{3} \end{aligned} \tag{29}$$

De esta forma ya hemos deducido los valores de λ_1 y λ_2 por lo cual podemos deducir los valores de x, y, z ya que $x = \lambda_1$, $y = 2x$ y $z = \frac{-\lambda_2}{2}$. Al final, hemos llegado a que el punto de solución del sistema de ecuaciones y que maximiza la función es: $x = \frac{2}{3}$, $y = \frac{4}{3}$, $z = \frac{-4}{3}$. Evaluando estos valores en la función nos resulta $f(x, y, z) = \frac{4}{3}$. Para comprobar que estos valores son máximos locales tendríamos que comprobar con las condiciones de optimalidad.

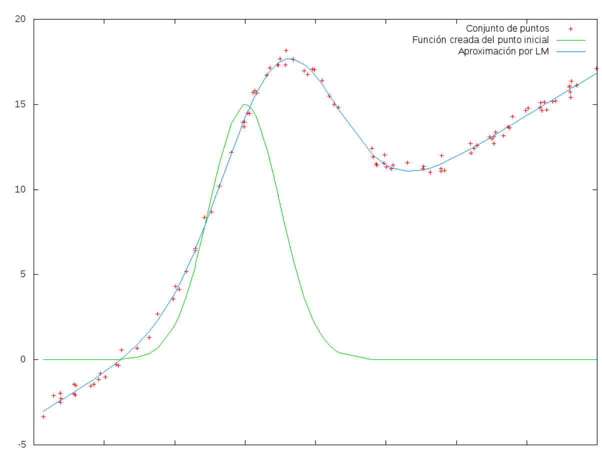
4. Ejercicio 4

Para este ejercicio se retomará la implementación del algoritmo LM de la tarea pasada y se hizo una modificación en el cálculo de la matriz Jacobiana, ahora se usó el método de diferencias finitas para aproximar la misma y usarla en el cáculo de los mínimos cuadrados no lineales.
Se volvió a repetir el ejercicio 3 de la tarea pasada en donde generábamos un modelo no lineal para aproximar a un conjunto de puntos previamente dados. Pero como se comentó en el párrafo anterior, en esta versión del algoritmo se usaron diferencias finitas hacia adelante para aproximar la matriz Jacobiana, las cuales tienen la siguiente fórmula $f'(x) = \frac{f(x+h)-f(x)}{h}$. A continuación se muestra una tabla donde se muestran los resultados obtenidos de aplicar el algoritmo con diferentes valores de h y un v fijo.
Para las ejecuciones que se mostrarán a continuación se usó como punto inicial el mismo de la tarea pasada, el cual es $p_0 = (0, 0, 15, -2, 1)^T$, una tolerancia de $\tau = 10^{-2}$ y un máximo número de 1000 iteraciones. Los valores que se reportarán son el h y la v usados, el número de la última iteración realizada por el algoritmo, las dos primeras componentes del vector resultante, el valor de la función en el vector resultante y el valor del gradiente en ese punto.

Cuadro 1: Resultados de las pruebas realizadas al método LM con diferentes valores de v y de h

h	v	**Iteración final** k	$(x_k)_1$	$(x_k)_2$	$f(x_k)$	$\|\|\nabla f(x_k)\|\|$
0.1	**1.5**	22	2.527718	1.700838	3.95375	0.00214488
0.01	**1.5**	22	2.527766	1.700371	3.95373	0.00229271
0.001	**1.5**	22	2.527769	1.700335	3.95373	0.00232093
0.1	**1.25**	34	2.527714	1.700873	3.95375	0.00425483
0.01	**1.25**	34	2.527770	1.700336	3.95373	0.00517923
0.001	**1.25**	34	2.527774	1.700299	3.95373	0.00525999

En la Tabla anterior podemos ver que se usaron dos valores de v y para cada uno de ellos se emplearon tres valores de h. Lo que podemos concluir es que pese a que el valor de h se hacía más pequeño, los resultados no fueron tan diferentes ya que el algoritmo tardó el mismo número de iteraciones para lograr convergencia (en ambos casos de v). En lo que se puede notar un pequeño cambio es en los valores del vector resultante así como en la evaluación de la función en dicho vector y en la norma del gradiente. Esto anterior puede indicar que la matriz Jacobiana no se malcondicionó al variar el valor de h, aunque hay que denotar que los valores tomados para este ejemplo no resultaron ser tan grandes ni tan pequeños para que ocurriera un cambio drástico.
A continuación se mostrará la gráfica generada con el modelo optimizado por el algoritmo LM usando un valor de $h = 0{,}001$.



(a) Figura 1. Modelo de la función encontrado con el algoritmo LM (azul). Modelo de la función con el punto inicial.

Como se puede apreciar el modelo generado con el punto optimizado por el algoritmo usando diferencias finitas ajusta muy bien al conjunto de puntos dado.

No se muestran las gráficas generadas con distintos valores de h ya que estan son casi idéndicas a la adjunta.