Tarea 12 - 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 12 de la materia de optimización los cuales son referentes a problemas sobre optimización con restricciones, y se usan los métodos de penalización cuadrática y Lagrangiano aumentado para resolverlos. Para el último ejercicio se hablará sobre la implementación del algoritmo así como los resultados obtenidos.

Keywords: Penalización cuadrática, Lagrangiano aumentado.

1. Ejercicio 1

Resuelva el problema

$$\begin{aligned} &\min_x \; x_1^2 + 2x_2^2 \\ &s.a. \; 1 - x_1 + x_2 \leq 0 \end{aligned} \tag{1}$$

usando el método de penalización cuadrática, es decir, calcule la expresión de los puntos x_μ que optimizan a las funciones de la forma $f(x)+\mu p(x)$, donde $p(x)$ donde p(x) es la función de penalización cuadrática, y calcule su límite cuando $\mu \to \infty$. A partir de esto describa que para con el número de condición de la matriz Hessiana.

Solución : Primero escribimos la función de penalización cuadrática

$$q(x;\mu) = x_1^2 + 2x_2^2 + \mu(max\{0, 1 - x_1 + x_2\})^2 \tag{2}$$

Ahora obtenemos las derivadas parciales de la función con respecto a x_1 y x_2

$$\begin{aligned} \frac{\partial q}{\partial x_1} &= 2x_1 - 2\mu(1 - x_1 + x_2), \; si \; 1 - x_1 + x_2 > 0 \\ \frac{\partial q}{\partial x_1} &= 2x_1, \; si \; 1 - x_1 + x_2 \leq 0 \\ \frac{\partial q}{\partial x_2} &= 4x_2 + 2\mu(1 - x_1 + x_2), \; si \; 1 - x_1 + x_2 > 0 \\ \frac{\partial q}{\partial x_2} &= 4x_4, \; si \; 1 - x_1 + x_2 \leq 0 \end{aligned} \tag{3}$$

Lo siguiente es trabajar con las ecuaciones que no pertenecen al caso trivial, es decir, cuando $1 - x_1 + x_2 > 0$. Igualamos a cero el gradiente y despejamos los valores de x que son puntos críticos de la función

$$\begin{aligned} \frac{\partial q}{\partial x_1} &= 2x_1 - 2\mu(1 - x_1 + x_2) = 0 \\ \frac{\partial q}{\partial x_2} &= 4x_2 + 2\mu(1 - x_1 + x_2) = 0 \end{aligned} \tag{4}$$

Hacemos un poco de simplificación algebraica

$$\begin{aligned} \frac{\partial q}{\partial x_1} &= x_1 - \mu + \mu x_1 - \mu x_2 = 0 \\ \frac{\partial q}{\partial x_2} &= 2x_2 + \mu - \mu x_1 + \mu x_2 = 0 \end{aligned} \tag{5}$$

Sumamos las ecuaciones anteriores por ambos lados de la igualdad

$$\begin{aligned} \frac{\partial q}{\partial x_1} + \frac{\partial q}{\partial x_1} \rightarrow x_1 + 2x_2 &= 0 \\ x_1 &= -2x_2 \end{aligned} \tag{6}$$

Ahora sustituimos x_1 en $\frac{\partial q}{\partial x_2}$

$$\begin{aligned} 2x_2 + \mu - \mu(-2x_2) + \mu x_2 &= 0 \\ 2x_2 + 3\mu x_2 &= -\mu \\ (2 + 3\mu)x_2 &= -\mu \\ x_2 &= \frac{-\mu}{2 + 3\mu} \\ x_1 &= \frac{2\mu}{2 + 3\mu} \end{aligned} \tag{7}$$

Ahora que encontramos los valores de x que son puntos críticos de la función vamos a ver qué pasa cuando $\mu \rightarrow \infty$. Primero para x_1

$$\lim_{\mu \to \infty} \frac{2\mu}{2 + 3\mu} = \frac{\infty}{\infty} \tag{8}$$

Este límite queda de la forma $\frac{\infty}{\infty}$ es decir, una forma indeterminada. Por lo cual aplicamos regla de L'Hopital

$$\lim_{\mu \to \infty} \frac{2\mu}{2 + 3\mu} = \lim_{\mu \to \infty} \frac{2}{3} = \frac{2}{3} \tag{9}$$

Hemos encontrado el valor al que llega dicho límite. Ahora trabajamos con x_2

$$\lim_{\mu\to\infty} \frac{-\mu}{2+3\mu} = \frac{-\infty}{\infty} \tag{10}$$

Este límite queda de la forma $\frac{-\infty}{\infty}$ es decir, una forma indeterminada. Por lo cual aplicamos regla de L'Hopital

$$\lim_{\mu\to\infty} \frac{-\mu}{2+3\mu} = \lim_{\mu\to\infty} \frac{-1}{3} = \frac{-1}{3} \tag{11}$$

Pasamos a la segunda parte del problema, la cual consiste en calcular la matriz Hessiana de la función q, encontrar sus eigenvalores y ver qué pasa con el número de condición de la matriz cuando $\mu \to \infty$, para esta parte vamos a trabajar de nuevo con el caso no trivial, esto porque en el otro caso μ se hace cero ya que no se cumple ninguna desigualdad. Primero encontramos la matriz Hessiana.

$$\nabla q = \begin{pmatrix} 2x_1 - 2\mu(1-x_1+x_2) & 4x_2 + 2\mu(1-x_1+x_2) \end{pmatrix}^T \tag{12}$$

Obtenemos las segundas derivadas y las derivadas cruzadas para formar la matriz Hessiana

$$\nabla^2 q = \begin{pmatrix} 2+2\mu & -2\mu \\ -2\mu & 4+2\mu \end{pmatrix} \tag{13}$$

Encontramos los eigenvalores de la matriz Hessiana y para ello se debe cumplir la siguiente ecuación $det(H-\lambda I) = 0$ donde λ es un eigenvalor de H

$$\nabla^2 q - \lambda I = \begin{pmatrix} 2+2\mu-\lambda & -2\mu \\ -2\mu & 4+2\mu-\lambda \end{pmatrix} \tag{14}$$

$$\begin{aligned} det(H-\lambda I) &= (2+2\mu-\lambda)(4+2\mu-\lambda) - 4\mu^2 \\ det(H-\lambda I) &= 8+4\mu-2\lambda+8\mu+4\mu^2-2\mu\lambda-4\lambda-2\mu\lambda+\lambda^2-4\mu^2 \\ det(H-\lambda I) &= \lambda^2-(6+4\mu)\lambda+12\mu+8 = 0 \end{aligned} \tag{15}$$

Para encontrar las raíces del polinomio de grado dos obtenido usamos la fórmula general

$$\begin{aligned}
\lambda &= \frac{4\mu + 6 \pm \sqrt{(-(4\mu+6))^2 - 4(12\mu+8)}}{2} \\
\lambda &= \frac{4\mu + 6 \pm \sqrt{16\mu^2 + 48\mu + 36 - 48\mu - 32}}{2} \\
\lambda &= \frac{4\mu + 6 \pm \sqrt{16\mu^2 + 4}}{2} \\
\lambda &= \frac{4\mu + 6 \pm 2\sqrt{4\mu^2 + 1}}{2} \\
\lambda_1 &= \frac{4\mu + 6 + 2\sqrt{4\mu^2 + 1}}{2} \\
\lambda_2 &= \frac{4\mu + 6 - 2\sqrt{4\mu^2 + 1}}{2}
\end{aligned} \tag{16}$$

Ahora que encontramos los valores de λ_1 y λ_2 podemos ver que pasa cuando $\mu \to \infty$. Primero vemos el caso con λ_1

$$\lim_{\mu \to \infty} 2\mu + 3 + \sqrt{4\mu^2 + 1} = \infty \tag{17}$$

Para este caso es muy fácil deducir que el límite va a infinito ya que tenemos una suma de términos positivos que dependen de μ, además de que la raíz de un infinito tiende a infinito. Ahora vamos con el caso de λ_2

$$\lim_{\mu \to \infty} 2\mu + 3 - \sqrt{4\mu^2 + 1} \tag{18}$$

En este caso no se puede determinar el límite a primera instancia ya que tendríamos una resta de infinitos lo cual no está definido. Entonces podemos racionalizar

$$\begin{aligned}
&\lim_{\mu \to \infty} 2\mu + 3 - \sqrt{4\mu^2 + 1}\frac{(2\mu + 3 + \sqrt{4\mu^2 + 1})}{(2\mu + 3 + \sqrt{4\mu^2 + 1})} \\
&\lim_{\mu \to \infty} \frac{(2\mu + 3)^2 - (\sqrt{4\mu^2 + 1})^2}{2\mu + 3 + \sqrt{4\mu^2 + 1}} \\
&\lim_{\mu \to \infty} \frac{4\mu^2 + 12\mu + 9 - 4\mu^2 - 1}{2\mu + 3 + \sqrt{4\mu^2 + 1}} \\
&\lim_{\mu \to \infty} \frac{12\mu + 8}{2\mu + 3 + \sqrt{4\mu^2 + 1}}
\end{aligned} \tag{19}$$

Ahora dividimos ambas partes de la fracción por μ

$$\begin{aligned}
&\lim_{\mu\to\infty} \frac{12+\frac{8}{\mu}}{2+\frac{3}{\mu}+\sqrt{4+\frac{1}{\mu^2}}} \\
&= \frac{12}{2+\sqrt{4}} = 3
\end{aligned} \tag{20}$$

Hemos llegado a un valor finito para el límite. Ahora podemos construir el número de condición de la matriz y para ello como la matriz Hessiana es normal, este número se define como $k_2 = \frac{|\lambda_{max}|}{|\lambda_{min}|}$ donde λ_i es un eigenvalor de la matriz. Para nuestro caso podemos ver fácilmente que λ_1 es el eigenvalor más grande de la Hessiana y λ_2 es el más pequeño (esto asumiendo $\mu > 0$). Sustituimos los valores de los eigenvalores con los límites aplicados para ver el número de condición

$$k_2 = \frac{\lambda_1}{\lambda_2} = \frac{\infty}{3} = \infty \tag{21}$$

Se ha demostrado que cuando $\mu \to \infty$ el número de condición también lo hace por lo cual la matriz Hessiana se malcondiciona al aplicar el algoritmo de penalización cuadrática e ir aumentando μ.

2. Ejercicio 2

Considere el problema

$$\begin{aligned}
&\min_{x} \ 3x_1^2 + x_2^2 \\
&s.a. \ x_1 + 2x_2 - 1 = 0
\end{aligned} \tag{22}$$

1. Escriba la función $L_A(x_1, x_2; \lambda, \mu)$ que se construye al aplicar el método de Lagrangiano aumentado.

$$L_A(x_1, x_2; \lambda, \mu) = 3x_1^2 + x_2^2 - \lambda(x_1 + 2x_2 - 1) + \frac{\mu}{2}(x_1 + 2x_2 - 1)^2 \tag{23}$$

2. Calcule la expresión del mínimo x_μ de la función $L_A(x_1, x_2; \lambda, \mu)$ para cada valor de μ dado.

Primero obtenemos las derivadas parciales de L_A con respecto a $x_1,\ x_2$ y las igualamos a cero. Primero mostramos la que es con respecto a x_1

$$\begin{aligned}
&\frac{\partial L_A}{\partial x_1} = 6x_1 - \lambda + \mu(x_1 + 2x_2 - 1) \\
1)\ &\frac{\partial L_A}{\partial x_1} = 6x_1 - \lambda + \mu x_1 + 2\mu x_2 - \mu = 0
\end{aligned} \tag{24}$$

Ahora la que es con respecto a x_2

$$\begin{aligned} \frac{\partial L_A}{\partial x_2} &= 2x_2 - 2\lambda + 2\mu(x_1 + 2x_2 - 1) \\ 2)\ \frac{\partial L_A}{\partial x_2} &= x_2 - \lambda + \mu x_1 + 2\mu x_2 - \mu = 0 \end{aligned} \tag{25}$$

Se han puesto números en las ecuaciones anteriores para referenciarlas más fácilmente. Ahora restamos la ecuación 2) de la 1) por ambos lados de la igualdad

$$\begin{aligned} 6x_1 - \lambda + \mu x_1 + 2\mu x_2 - \mu = 0\ - \\ x_2 - \lambda + \mu x_1 + 2\mu x_2 - \mu = 0 \\ \rightarrow 6x_1 - x_2 = 0 \\ x_2 = 6x_1 \end{aligned} \tag{26}$$

Hemos llegado a una expresión para x_2 en la que sólo depende de x_1 por lo cual sustituimos x_2 en la ecuación 2) y despejamos

$$\begin{aligned} 6x_1 - \lambda + \mu x_1 + 2\mu(6x_1) - \mu &= 0 \\ 6x_1 + 13\mu x_1 &= \mu + \lambda \\ (6 + 13\mu)x_1 &= \mu + \lambda \\ x_1 &= \frac{\mu + \lambda}{6 + 13\mu} \\ x_2 &= \frac{6(\mu + \lambda)}{6 + 13\mu} \end{aligned} \tag{27}$$

Se han encontrado los valores de x_1 y x_2 que fungen como puntos críticos de la función.

3. Usando lo anterior, encuentre el punto al que debe converger la secuencia x_μ cuando $\mu \rightarrow \infty$, y una vez hecho esto, obtenga el valor de λ.

   Buscamos los límites de x_1 y x_2 cuando $\mu \rightarrow \infty$. Primero trabajamos con x_1

$$\lim_{\mu\rightarrow\infty} x_1 = \frac{\mu + \lambda}{6 + 13\mu} = \frac{\infty}{\infty} \tag{28}$$

   Llegamos a una forma indeterminada de $\frac{\infty}{\infty}$ por lo cual aplicamos la regla de L'Hopital

$$\lim_{\mu\rightarrow\infty} \frac{\mu + \lambda}{6 + 13\mu} = \lim_{\mu\rightarrow\infty} \frac{1}{13} = \frac{1}{13} \tag{29}$$

Ahora trabajamos con x_2

$$\lim_{\mu \to \infty} x_2 = \frac{6(\mu + \lambda)}{6 + 13\mu} = \frac{\infty}{\infty} \tag{30}$$

Llegamos a una forma indeterminada de $\frac{\infty}{\infty}$ por lo cual aplicamos la regla de L'Hopital

$$\lim_{\mu \to \infty} \frac{6(\mu + \lambda)}{6 + 13\mu} = \lim_{\mu \to \infty} \frac{6}{13} = \frac{6}{13} \tag{31}$$

Una vez encontrados los límites sustituimos x_1 y x_2 en la ecuación 1)

$$\begin{aligned} \frac{6}{13} - \lambda + \frac{\mu}{13} + \frac{12\mu}{13} - \mu = 0 \\ \lambda = \frac{6}{13} \end{aligned} \tag{32}$$

3. Ejercicio 3

Para este ejercicio se implementó el algoritmo de penalización cuadrática, el cual es utilizado para problemas de optimización con restricciones tanto de igualdad como de desigualdad. Para la parte de optimización del mismo se utilizó el algoritmo de gradiente conjugado no lineal con la Beta de Fletcher - Reeves en donde se aplica el enfoque de diferencias finitas con $dx = 0{,}0010$ para obtener el gradiente de la función. Para ir modificando el parámetro de penalización μ se va multiplicando por un factor constante para que esta vaya creciendo poco a poco. El parámetro τ_k se da inicial para la primer iteración y en cada nueva iteración se actualiza de la forma $\tau_k = \frac{1}{i}$ donde i indica el número de iteración.

El problema a minimizar es el siguiente:

$$\begin{aligned} \min_{x_1, x_2} \ & (x_1 - 6)^2 + (x_2 - 7)^2 \\ s.a. \ & -3x_1 - 2x_2 + 6 \leq 0 \\ & -x_1 + x_2 - 3 \leq 0 \\ & x_1 + x_2 - 7 \leq 0 \\ & \frac{2}{3}x_1 - x_2 - \frac{4}{3} \leq 0 \end{aligned} \tag{33}$$

Para este problema se proponen dos puntos iniciales, además de una tolerancia para la función $p(x)$ de $\tau = 0{,}005$, $\mu = 1$ y el primer valor de la secuencia $\tau_0 = 1{,}0$ para la tolerancia del algoritmo GCNL. A continuación se muestra una tabla con los resultados obtenidos para cada punto inicial dado.

Cuadro 1: Resultados obtenidos por el algoritmo de penalización cuadrática

Punto inicial	**Punto final**	**Número de iteraciones**	$f(x)$	$h_1(x)$	$h_2(x)$	$h_3(x)$	$h_4(x)$
(-15, 25)ˆT	(3.012, 4.012)ˆT	6	17.8541	-11.0611	-2.00034	0.0243683	-3.33711
(-15, -10)ˆT	(3.027, 3.996)ˆT	6	17.8546	-11.0762	-2.03064	0.0243655	-3.31186

Podemos ver que para los dos puntos inciales dados y la tolerancia final dada, el algoritmo demoró 6 iteraciones para converger, en ambas se llegó al punto $(3,4)^T$ aproximadamente, además de que se cumple casi todas las restricciones propuestas, la tres no la cumple ya que no encontró los valores exactos del punto final y por lo tanto se sale un poco de la zona factible.