Métodos de Penalización y Lagrangiano Aumentado

Oscar S. Dalmau Cedeño
dalmau@cimat.mx

22 de noviembre de 2018

- El metodo del Lagrangiano aumentado este relacionado con los metodos de penalizacion.
- Este metodo intenta resolver el problema del mal condicionamiento que se presenta en los algoritmos de penalizacion: cuadraticos y de barrera

Motivacion: En el problema de optimizacion con restricciones de igualdad, la penalizacion cuadratica penaliza las restricciones que son violadas, ie, dado el problema de optimizacion, con restricciones de igualdad

$$\begin{aligned} & \min f(\boldsymbol{x}) \\ s.a. \qquad & c_i(\boldsymbol{x}) = 0,\ i \in \mathcal{E} \end{aligned}$$

La funcion de penalizacion cuadratica

$$Q(\boldsymbol{x}, \mu_k) = f(\boldsymbol{x}) + \frac{\mu_k}{2} \sum_{i \in \mathcal{E}} (c_i(\boldsymbol{x}))^2$$

Segun el teorema visto en clases, sobre la convergencia del Framework o algoritmo, las restricciones $c_i(\boldsymbol{x}_k) = 0$ no necesariamente se satisfacen para cada subproblema resuelto, y solo se obtiene una perturbacion, es decir, segun el teorema

$$\begin{aligned} \lambda_i^* &\approx -\mu_k c_i(\boldsymbol{x}_k) \\ c_i(\boldsymbol{x}_k) &\approx -\frac{\lambda_i}{\mu_k} \propto \frac{1}{\mu_k} \end{aligned}$$

- Es decir, $c_i(\boldsymbol{x}_k)$ es proporcional a $\frac{1}{\mu_k}$ y solo se garantiza que $c_i(\boldsymbol{x}_k) \to 0$ cuando $\mu_k \to \infty$.
- Lo anterior no es conveniente en alguno, pues en ocasiones nos gustaria que se satisficieran las restricciones de igualdad o al menos que se obtuviera una mejor aproximacion en un numero 'moderado' de iteraciones.

Lo anterior se puede lograr mediante la introduccion de una nueva funcion, conocida como **Lagrangiano aumentado** que considera o es una combinacion del Lagrangiano y el termino de penalizacion, y se define como sigue:

$$\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}; \mu) \stackrel{def}{=} f(\boldsymbol{x}) - \sum_{i \in \mathcal{E}} \lambda_i c_i(\boldsymbol{x}) + \frac{\mu}{2} \sum_{i \in \mathcal{E}} (c_i(\boldsymbol{x}))^2$$

que es muy parecido al Lagrangiano, y solo se diferencia en el termino de penalizacion.

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{i \in \mathcal{E}} \lambda_i c_i(\boldsymbol{x})$$

La idea ahora es desarrollar una algoritmo, similar a los de penalizacion en los que se resuelve una secuencia de subproblemas

$$\boldsymbol{x}_k = \arg\min_{\boldsymbol{x}} \mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k)$$

donde los parametros $\boldsymbol{\lambda} = \boldsymbol{\lambda}^k$ y $\mu = \mu_k$ son fijos en cada iteracion. En la practica, $\boldsymbol{x}_k$ es solo un minimizador aproximado de $\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k)$.

- Para la actualizacion del parametro de penalizacion μ_k se sigue la misma idea usada en penalizacion cuadratica, es decir, se selecciona una secuencia creciente y positiva, $0 < \mu_k < \mu_{k+1}$.
- Para el parámetro $\boldsymbol{\lambda}^k$ se comparan las condiciones de optimalidad del Lagrangiano y del Lagrangiano aumentado..

$$
\begin{aligned}
0 &= \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \nabla f(\boldsymbol{x}^*) - \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*) \\
0 &\approx \nabla_{\boldsymbol{x}} \mathcal{L}_A(\boldsymbol{x}_k, \boldsymbol{\lambda}^k; \mu_k) = \nabla f(\boldsymbol{x}_k) - \sum_{i \in \mathcal{E}} (\lambda_i^k - \mu_k c_i(\boldsymbol{x}_k)) \nabla c_i(\boldsymbol{x}_k)
\end{aligned}
$$

Luego, comparando se tiene que:

$$\lambda_i^* \approx \lambda_i^k - \mu_k c_i(\boldsymbol{x}_k) \tag{1}$$

$$c_i(\boldsymbol{x}_k) \approx \frac{\lambda_i^k - \lambda_i^*}{\mu_k} \tag{2}$$

y si λ_i^k es lo suficientemente cercano a λ_i^*, entonces la no factibilidad de $\boldsymbol{x}_k$ es menor a $\frac{1}{\mu_k}$, en lugar de ser proporcional a $\frac{1}{\mu_k}$ como en el caso de la penalizacion cuadratica.

Por otro lado, la relacion 1 sugiere una forma de actualizacion de $\boldsymbol{\lambda}^k$ es decir:

$$\lambda_i^{k+1} \quad = \quad \lambda_i^k - \mu_k c_i(\boldsymbol{x}_k) \tag{3}$$

Basado en lo anterior se tiene el siguiente algoritmo o esquema

Algorithm 1 Algoritmo de Lagrangiano Aumentado (restricciones de igualdad)

Dado un punto inicial x_0^s, $\boldsymbol{\lambda}^0$, $\tau_0, \mu_0 > 0$
for $k = 0, 1, 2, \cdots$ **do**
  Encontrar un minimizador aproximado $\boldsymbol{x}_k$ de $\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k)$ iniciando en x_k^s y se termina cuando $\|\nabla_{\boldsymbol{x}} \mathcal{L}_A(\boldsymbol{x}_k, \boldsymbol{\lambda}; \mu_k)\| \leq \tau_k$
  if converge **then**
    Parar el algoritmo con solucion $\boldsymbol{x}_k$
  end if
  Actulizar $\lambda_i^{k+1} = \lambda_i^k - \mu_k c_i(\boldsymbol{x}_k)$
  Seleccionar $\mu_{k+1} > \mu_k$ y τ_{k+1}
  Seleccionar un nuevo punto inicial $\boldsymbol{x}_{k+1}^s$
end for

- El metodo converge sin necesidad de incrementar μ indefinidamente
- Debido a lo anterior se reduce el problema del mal condicionamiento, presente en los algoritmos de penalizacion
- El metodo es menos sensible a la selección del punto inicial x_{k+1}^s

Teorema

Sea $\boldsymbol{x}^*$ un minimo local del problema

$$\begin{aligned} & \min f(\boldsymbol{x}) \\ s.a. \qquad & c_i(\boldsymbol{x}) = 0, \; i \in \mathcal{E} \end{aligned}$$

Supongamos que gradientes $\nabla c_i(\boldsymbol{x}^*), i \in \mathcal{E}$ son linealmente independientes (condicion LICQ). Si cumplen se las condiciones suficientes de segundo orden (Teorema 12.6), es decir, si se satisfacen las condiciones KKT del problema anterior para $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ y si ademas $\nabla^2_{\boldsymbol{xx}}\mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ es positivo definido en para $\omega \neq 0 \in \mathcal{C}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$. Entonces existe un umbral $\hat{\mu}$ tal que para todo $\mu \geq \hat{\mu}$, $\boldsymbol{x}^*$ es un minimizador local de $\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^*; \mu)$

Nota:

$$\mathcal{F}(x) = \{d|\ d^T \nabla c_i(x) = 0, i \in \mathcal{E};\ d^T \nabla c_i(x) \geq 0, i \in \mathcal{A}(x) \cap \mathcal{I}\}$$

$$\mathcal{C}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \{d \in \mathcal{F}(\boldsymbol{x}^*)|\ d^T \nabla c_i(\boldsymbol{x}^*) = 0, i \in \mathcal{A}(\boldsymbol{x}^*) \cap \mathcal{I},\ \lambda_i^* > 0\}$$

$$d \in \mathcal{C}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \begin{cases} d^T \nabla c_i(\boldsymbol{x}^*) = 0, & i \in \mathcal{E} \\ d^T \nabla c_i(\boldsymbol{x}^*) = 0, & i \in \mathcal{A}(\boldsymbol{x}^*) \cap \mathcal{I},\ \lambda_i^* > 0 \\ d^T \nabla c_i(\boldsymbol{x}^*) \geq 0, & i \in \mathcal{A}(\boldsymbol{x}^*) \cap \mathcal{I},\ \lambda_i^* = 0 \end{cases}$$

- Primero se probara que si $\boldsymbol{x}^*$ es un minimizador local del problema original, entonces es un punto critico de $\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^*; \mu)$.
- Como $\boldsymbol{x}^*$ es un minimizador local del problema original, entonces satisface las KKT's. Luego

$$
\begin{aligned}
\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) &= 0 \\
c_i(\boldsymbol{x}^*) &= 0, i \in \mathcal{E}
\end{aligned}
$$

Luego, evaluando el Lagrangiano aumentado en $\boldsymbol{x}^*$ y usando las KKT's anteriores

$$
\begin{aligned}
\nabla_{\boldsymbol{x}} \mathcal{L}_A(\boldsymbol{x}^*, \boldsymbol{\lambda}^*; \mu) &= \nabla f(\boldsymbol{x}^*) - \sum_{i \in \mathcal{E}} \left(\lambda_i^* - \mu_k \cancelto{0}{c_i(\boldsymbol{x}^*)} \right) \nabla c_i(\boldsymbol{x}^*) \\
&= \nabla f(\boldsymbol{x}^*) - \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(\boldsymbol{x}^*) \\
&= \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \\
&= 0
\end{aligned}
$$

por lo que $\boldsymbol{x}^*$ es un punto critico de $\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^*; \mu)$, independientemente del valor de μ

- Ahora se probara que es un minimo local, es decir, hay que probar que $\mathcal{L}_A(\boldsymbol{x}^*, \boldsymbol{\lambda}^*; \mu)$ es positiva definida.. Para ello se usara reduccion al absurdo!.
- Supongamos $\nabla^2_{\boldsymbol{xx}}\mathcal{L}_A(\boldsymbol{x}^*, \boldsymbol{\lambda}^*; \mu)$ no es positiva definida, entonces existe $\omega \neq 0$ tal que

$$\omega^T \nabla^2_{\boldsymbol{xx}}\mathcal{L}_A(\boldsymbol{x}^*, \boldsymbol{\lambda}^*; \mu)\omega \quad \leq \quad 0$$

$$
\begin{aligned}
\nabla^2_{\boldsymbol{xx}} \mathcal{L}_A(\boldsymbol{x}^*, \boldsymbol{\lambda}^*; \mu) &= \nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) + \mu \sum_{i \in \mathcal{E}} \cancelto{0}{c_i(\boldsymbol{x}^*)} \nabla^2_{\boldsymbol{xx}} c_i(\boldsymbol{x}^*) \\
&\quad + \nabla c_i(\boldsymbol{x}^*) \nabla c_i(\boldsymbol{x}^*)^T \\
&= \nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) + \mu A^T A
\end{aligned}
$$

con $A^T = [\nabla c_1(\boldsymbol{x}^*), \cdots, \nabla c_{|\mathcal{E}|}(\boldsymbol{x}^*)]$.

Entonces

$$\begin{aligned} \omega^T \left(\nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) + \mu A A^T \right) \omega &\leq 0 \\ \|A\omega\|^2 &\leq -\frac{1}{\mu} \omega^T \nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \omega \end{aligned}$$

haciendo tender $\mu \to \infty$ se tiene

$$\|A\omega\|^2 \leq 0$$

Luego $A\omega = 0$, ie, $\nabla c_i(\boldsymbol{x}^*)^T \omega = 0, \ i \in \mathcal{E}$ y por tanto $\omega \in \mathcal{C}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$

Como $\omega \in \mathcal{C}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ y por tanto $A\omega = 0$ entonces

$$
\begin{aligned}
0 &\geq \omega^T \nabla^2_{\boldsymbol{xx}} \mathcal{L}_A(\boldsymbol{x}^*, \boldsymbol{\lambda}^*; \mu)\omega \\
&= \omega^T \left(\nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) + \mu A A^T\right) \omega \\
&= \omega^T \nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\omega + \mu \cancelto{0}{\|A\omega\|^2}
\end{aligned}
$$

Luego

$$\omega^T \nabla^2_{\boldsymbol{x}\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\omega \quad \leq \quad 0$$

Lo que es una contradiccion con las condiciones de segundo orden, ie, pues $\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ es positiva definida en $\mathcal{C}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$, con $\boldsymbol{x}^*$ un minimo local del problema original.

Problema de optimizacion

$$\begin{aligned} & \min_{\boldsymbol{x}} f(\boldsymbol{x}) \\ s.a. \qquad & c_i(\boldsymbol{x}) \geq 0, \; i \in \mathcal{I} \end{aligned}$$

Para usar el metodo del Lagrangiano Aumentado, se introducen variables de holguras, de este modo se obtiene el siguiente problema:

$$
\begin{aligned}
& \min_{\boldsymbol{x},\boldsymbol{s}} f(\boldsymbol{x}) \\
s.a. \qquad & c_i(\boldsymbol{x}) - s_i = 0, \;\; i \in \mathcal{I} \\
& s_i \geq 0, \;\; i \in \mathcal{I}
\end{aligned}
$$

Ahora podemos calcular el Lagrangiano aumentado, solo para las restricciones de igualdad, y tenemos el siguiente problema:

$$\min_{\boldsymbol{x},\boldsymbol{s}} \mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k) = f(\boldsymbol{x}) - \sum_{i\in\mathcal{I}} \lambda_i^k (c_i(\boldsymbol{x}) - s_i) + \frac{\mu_k}{2} \sum_{i\in\mathcal{I}} (c_i(\boldsymbol{x}) - s_i)^2$$

$$s_i \geq 0, \; i \in \mathcal{I}$$

Ahora podemos resolver el problema de optimizacion, respecto a $\boldsymbol{x}, \boldsymbol{s}$ y luego es posible eliminar la variable de holgura $\boldsymbol{s}$ usando un metodo de proyeccion.

$$\begin{aligned} \partial_{s_i}\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k) &= \lambda_i^k - \mu_k(c_i(\boldsymbol{x}) - s_i) = 0 && (4) \\ s_i &= c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k}. && (5) \end{aligned}$$

es decir

$$s_i = \begin{cases} c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} & \text{si } c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} \geq 0, (\text{ ó } \lambda_i^k \leq \mu_k c_i(\boldsymbol{x})) \\ 0 & \text{si } c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} < 0, (\text{ ó } \lambda_i^k > \mu_k c_i(\boldsymbol{x})) \end{cases}$$

Luego (el signo igual de la comparación puede estar en cualquiera de la opciones, pero se ubica con el signo '$<$' convenientemente)

- Si $c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} > 0$, entonces $c_i(\boldsymbol{x}) - s_i = \frac{\lambda_i^k}{\mu_k}$

$$
\begin{aligned}
& -\sum_{i\in\mathcal{I}} \lambda_i^k (c_i(\boldsymbol{x}) - s_i) + \frac{\mu_k}{2} \sum_{i\in\mathcal{I}} (c_i(\boldsymbol{x}) - s_i)^2 \\
= \quad & -\sum_{i\in\mathcal{I}} \lambda_i^k \frac{\lambda_i^k}{\mu_k} + \frac{\mu_k}{2} \sum_{i\in\mathcal{I}} \left(\frac{\lambda_i^k}{\mu_k}\right)^2 \\
= \quad & -\frac{1}{2} \sum_{i\in\mathcal{I}|c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} > 0} \frac{(\lambda_i^k)^2}{\mu_k}
\end{aligned}
$$

- Si $c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} \leq 0$, entonces $s_i = 0$

$$
\begin{aligned}
& -\sum_{i\in\mathcal{I}} \lambda_i^k (c_i(\boldsymbol{x}) - s_i) + \frac{\mu_k}{2} \sum_{i\in\mathcal{I}} (c_i(\boldsymbol{x}) - s_i)^2 \\
= & \sum_{i\in\mathcal{I} | c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} \leq 0} -\lambda_i^k c_i(\boldsymbol{x}) + \frac{\mu_k}{2} (c_i(\boldsymbol{x}))^2
\end{aligned}
$$

Basado en los casos anteriores, el Lagrangiano Aumentado se puede escribir como sigue

$$
\begin{aligned}
\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k) &= \nabla f(\boldsymbol{x}) + \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} \leq 0} -\lambda_i^k c_i(\boldsymbol{x}) + \frac{\mu_k}{2} \left(c_i(\boldsymbol{x})\right)^2 \\
&\quad - \frac{1}{2} \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}) - \frac{\lambda_i^k}{\mu_k} > 0} \frac{(\lambda_i^k)^2}{\mu_k}
\end{aligned}
$$

- En la practica usamos un metodo iterativo para estimar el optimo $\boldsymbol{x}_k$ del subproblema.
- En cada iteracion se deben garantizar las restricciones $s_i \geq 0$, para lo cual se usa un metodo de proyeccion (iterativo).
- Conocidos $(\boldsymbol{x}_k, \lambda_i^k, \mu_k)$ en la iteracion actual, ie, $\boldsymbol{x}_k$ corresponde al valor actual durante el proceso iterativo y λ_i^k, μ_k son fijos, se identifican el las $s_i < 0$ (que dependen de $(\boldsymbol{x}_k, \lambda_i^k, \mu_k)$ segun (5)) y se proyectan a 0,

Finalmente, el subproblema para el Lagrangiano Aumentado se puede escribir como sigue:

$$\begin{aligned}\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k) \quad &:= \quad \nabla f(\boldsymbol{x}) + \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} \leq 0} -\lambda_i^k c_i(\boldsymbol{x}) + \frac{\mu_k}{2} (c_i(\boldsymbol{x}))^2 \\ & \quad - \frac{1}{2} \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} > 0} \frac{(\lambda_i^k)^2}{\mu_k}\end{aligned}$$

$$\boldsymbol{x}_{k+1} \quad = \quad \arg\min_{\boldsymbol{x}} \mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k)$$

Que se puede escribir de forma compacta, si se define la funcion

$$\Psi(a,b,c) = \begin{cases} -ba + \frac{c}{2}a^2 & \text{si } a - \frac{b}{c} \leq 0, \\ -\frac{b^2}{2c} & \text{si } a - \frac{b}{c} > 0 \end{cases}$$

y se obtiene simplemente

$$\begin{aligned}
\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k) &= f(\boldsymbol{x}) + \sum_{i \in \mathcal{I}} \Psi(c_i(\boldsymbol{x}), \lambda_i^k, \mu_k) \\
\nabla \mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k) &= \nabla f(\boldsymbol{x}) + \sum_{i \in \mathcal{I}} \nabla \Psi(c_i(\boldsymbol{x}), \lambda_i^k, \mu_k) \\
&= \nabla f(\boldsymbol{x}) + \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} \leq 0} (-\lambda_i^k + \mu_k c_i(\boldsymbol{x})) \nabla c_i(\boldsymbol{x})
\end{aligned}$$

- Para actualizar los λ_i^k en la próxima iteración, suponemos que se ha calculado el optimo $\boldsymbol{x}_k$ en el problema actual.
- Ahora podemos comparar el gradiente del Lagrangiano con el gradiente del Lagrangiano aumentado evaluados en los optimos $\boldsymbol{x}^*$, $\boldsymbol{x}_k$ del problema original y del subproblema respectivamente,

es decir

$$\begin{aligned}
\nabla \mathcal{L}_A(\boldsymbol{x}_k, \boldsymbol{\lambda}^k; \mu_k) &= \nabla f(\boldsymbol{x}_k) - \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} > 0} 0 \\
&+ \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} \leq 0} (-\lambda_i^k + \mu_k c_i(\boldsymbol{x}_k)) \nabla c_i(\boldsymbol{x}_k) \\
\nabla \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) &= \nabla f(\boldsymbol{x}^*) - \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}^*) \leq 0} \lambda_i^* \nabla c_i(\boldsymbol{x}^*) \\
&- \sum_{i \in \mathcal{I} | c_i(\boldsymbol{x}^*) > 0} \cancelto{0}{\lambda_i^*} \nabla c_i(\boldsymbol{x}^*) \qquad 0, \text{ pues } \lambda_i^* c_i(\boldsymbol{x}^*) = 0
\end{aligned}$$

Haciendo la comparacion

$$\lambda_i^* \approx \begin{cases} \lambda_i^k - \mu_k c_i(\boldsymbol{x}_k) & \text{si } c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} \leq 0 \\ 0 & \text{si } c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} > 0 \end{cases} \quad (6)$$

$$\lambda_i^{k+1} = \begin{cases} \lambda_i^k - \mu_k c_i(\boldsymbol{x}_k) & \text{si } c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} \leq 0 \\ 0 & \text{si } c_i(\boldsymbol{x}_k) - \frac{\lambda_i^k}{\mu_k} > 0 \end{cases} \tag{7}$$

$$\lambda_i^{k+1} = \max(0, \lambda_i^k - \mu_k c_i(\boldsymbol{x}_k)) \tag{8}$$

Algorithm 2 Algoritmo de Lagrangiano Aumentado (restricciones de desigualdad)

Dado un punto inicial x_0^s, $\boldsymbol{\lambda}^0$, $\tau_0, \mu_0 > 0$
for $k = 0, 1, 2, \cdots$ **do**
  Encontrar un minimizador aproximado $\boldsymbol{x}_k$ de $\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}^k; \mu_k)$ iniciando en x_k^s y se termina cuando $\|\nabla_{\boldsymbol{x}} \mathcal{L}_A(\boldsymbol{x}_k, \boldsymbol{\lambda}^k; \mu_k)\| \leq \tau_k$
  if converge **then**
    Parar el algoritmo con solucion $\boldsymbol{x}_k$
  end if
  Actulizar $\lambda_i^{k+1} = \max(0, \lambda_i^k - \mu_k c_i(\boldsymbol{x}_k))$
  Seleccionar $\mu_{k+1} > \mu_k$
  Seleccionar τ_{k+1}
  Seleccionar un nuevo punto inicial $\boldsymbol{x}_{k+1}^s$
end for