Programacion Cuadratica Secuencial

Oscar S. Dalmau Cedeño
dalmau@cimat.mx

15 de noviembre de 2018

1. SQP con restricciones de igualdad
    - Metodo SQP local
    - SQP Framework
2. SQP con restricciones de igualdad y desigualdad
3. Cuasi Newton para SQP: Damped BFGS

Dado el problema con restricciones de igualdad:

$$\min f(\boldsymbol{x}) \tag{1}$$

$$s.a.: \; c_i(\boldsymbol{x}) = 0, \; i \in \mathcal{E} \tag{2}$$

Calculemos su Lagrangiano y las condiciones KKTs correspondientes.

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^T \boldsymbol{c}(\boldsymbol{x}) \tag{3}$$

donde $\boldsymbol{c}(\boldsymbol{x}) \overset{def}{=} [c_i(\boldsymbol{x})]_{i\in\mathcal{E}} \in \mathbb{R}^{|\mathcal{E}|}$. Luego

$$\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \nabla f(\boldsymbol{x}) - \nabla \boldsymbol{c}(\boldsymbol{x}) \boldsymbol{\lambda} = \mathbf{0} \tag{4}$$

$$\boldsymbol{c}(\boldsymbol{x}) = \mathbf{0} \tag{5}$$

donde $\nabla \boldsymbol{c}(\boldsymbol{x}) = [\nabla c_i(\boldsymbol{x})]_{i\in\mathcal{E}} \in \mathbb{R}^{n\times|\mathcal{E}|} \overset{def}{=} \mathbf{A}(\boldsymbol{x})^T$.

Ahora podemos usar el metodo de Newton para resolver el problema

$$F(\boldsymbol{x}, \boldsymbol{\lambda}) = \mathbf{0} \tag{6}$$

donde

$$\begin{aligned} F(\boldsymbol{x}, \boldsymbol{\lambda}) &= \begin{bmatrix} \nabla_x \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) \\ \boldsymbol{c}(\boldsymbol{x}) \end{bmatrix} \\ &= \begin{bmatrix} \nabla f(\boldsymbol{x}) - \nabla \boldsymbol{c}(\boldsymbol{x}) \boldsymbol{\lambda} \\ \boldsymbol{c}(\boldsymbol{x}) \end{bmatrix} \\ &= \begin{bmatrix} \nabla f(\boldsymbol{x}) - \mathbf{A}(\boldsymbol{x})^T \boldsymbol{\lambda} \\ \boldsymbol{c}(\boldsymbol{x}) \end{bmatrix} \end{aligned}$$

Se resuelve el siguiente sistema con respecto a $[\boldsymbol{p}_k^x; \boldsymbol{p}_k^\lambda]$

$$F'(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{p}_k^\lambda \end{bmatrix} = -F(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) \qquad (7)$$

Luego, se tiene el siguiente esquema iterativo (Newton)

$$\begin{bmatrix} \boldsymbol{x}_{k+1} \\ \boldsymbol{\lambda}_{k+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{\lambda}_k \end{bmatrix} + \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{p}_k^\lambda \end{bmatrix} \tag{8}$$

$$F'(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{p}_k^\lambda \end{bmatrix} = -F(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) \tag{9}$$

$$\begin{bmatrix} \nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) & -\mathbf{A}(\boldsymbol{x}_k)^T \\ \mathbf{A}(\boldsymbol{x}_k) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{p}_k^\lambda \end{bmatrix} = - \begin{bmatrix} \nabla f(\boldsymbol{x}_k) - \mathbf{A}(\boldsymbol{x}_k)^T \boldsymbol{\lambda}_k \\ \boldsymbol{c}(\boldsymbol{x}_k) \end{bmatrix}$$

Definiendo

$$
\begin{aligned}
f_k &\overset{def}{=} f(\boldsymbol{x}_k) && (10)\\
\nabla f_k &\overset{def}{=} \nabla f(\boldsymbol{x}_k) && (11)\\
\mathcal{L}_k &\overset{def}{=} \mathcal{L}(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) && (12)\\
\nabla_{\boldsymbol{x}} \mathcal{L}_k &\overset{def}{=} \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) && (13)\\
\nabla^2_{\boldsymbol{xx}} \mathcal{L}_k &\overset{def}{=} \nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}_k, \boldsymbol{\lambda}_k) && (14)\\
\boldsymbol{c}_k &\overset{def}{=} \boldsymbol{c}(\boldsymbol{x}_k) && (15)\\
\mathbf{A}_k^T &\overset{def}{=} \mathbf{A}(\boldsymbol{x}_k)^T \overset{def}{=} [\nabla c_i(\boldsymbol{x}_k)]_{i \in \mathcal{E}} && (16)
\end{aligned}
$$

$$\begin{bmatrix} \nabla^2_{\boldsymbol{x}\boldsymbol{x}}\mathcal{L}_k & -\mathbf{A}_k^T \\ \mathbf{A}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{p}_k^\lambda \end{bmatrix} = -\begin{bmatrix} \nabla f_k - \mathbf{A}_k^T \boldsymbol{\lambda}_k \\ \boldsymbol{c}_k \end{bmatrix} \quad (17)$$

Supuestos

1. El jacobiano de las restricciones $\mathbf{A}(\boldsymbol{x})$ tiene rango completo
2. La matriz $\mathcal{L}(\boldsymbol{x}, \lambda)$ es positiva definida en el espacio tangente de las restricciones, ie, $d^T \nabla^2_{xx} \mathcal{L}(\boldsymbol{x}, \lambda) d > 0$ para toda $d \neq 0$ tal que $\mathbf{A}(\boldsymbol{x}) d = 0$

Recordatorio: Karush-Kuhn-Tucker matrix

Lemma

Let $\mathbf{A}$ have full row rank, and assume that the reduced Hessian matrix $\mathbf{Z}^T\mathbf{G}\mathbf{Z}$ is positive definite. Then the KKT matrix

$$\mathbf{K} = \begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \tag{18}$$

is nonsingular, and hence there is a unique vector pair $(\boldsymbol{x}^*, \lambda^*)$ satisfying the KKT conditions.

Here $\mathbf{Z} \in \mathbb{R}^{n\times(n-m)}$ is a matrix whose columns are a basis for the null space of $\mathbf{A}$. That is, $\mathbf{Z}$ has full rank and satisfies $\mathbf{A}\mathbf{Z} = 0$.

SQP Framework

$$\begin{bmatrix} \nabla^2_{\boldsymbol{xx}}\mathcal{L}_k & -\mathbf{A}_k^T \\ \mathbf{A}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{p}_k^\lambda \end{bmatrix} = -\begin{bmatrix} \nabla f_k - \mathbf{A}_k^T \boldsymbol{\lambda}_k \\ \boldsymbol{c}_k \end{bmatrix} \quad (19)$$

$$\begin{bmatrix} \nabla^2_{\boldsymbol{xx}}\mathcal{L}_k & -\mathbf{A}_k^T \\ \mathbf{A}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{\lambda}_k + \boldsymbol{p}_k^\lambda \end{bmatrix} = -\begin{bmatrix} \nabla f_k \\ \boldsymbol{c}_k \end{bmatrix} \quad (20)$$

$$\begin{bmatrix} \nabla^2_{\boldsymbol{xx}}\mathcal{L}_k & -\mathbf{A}_k^T \\ \mathbf{A}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k^x \\ \boldsymbol{\lambda}_{k+1} \end{bmatrix} = -\begin{bmatrix} \nabla f_k \\ \boldsymbol{c}_k \end{bmatrix} \quad (21)$$

El sistema anterior se puede escribir como sigue

$$\begin{bmatrix} \nabla^2_{\boldsymbol{xx}}\mathcal{L}_k & -\mathbf{A}_k^T \\ \mathbf{A}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_k^x \\ l_k \end{bmatrix} = -\begin{bmatrix} \nabla f_k \\ \boldsymbol{c}_k \end{bmatrix} \tag{22}$$

El sistema anterior se puede ver con las KKTs de un problema cuadratico, donde $\hat{\mathcal{L}}_k(\boldsymbol{p}, l)$ es el Lagrangiano correspondiente

$$\begin{aligned} \nabla_{\boldsymbol{p}} \hat{\mathcal{L}}_k(\boldsymbol{p}, l) = \nabla^2_{\boldsymbol{xx}} \mathcal{L}_k \boldsymbol{p} + \nabla f_k - \mathbf{A}_k^T l &= \mathbf{0} \qquad (23) \\ \mathbf{A}_k \boldsymbol{p} + \boldsymbol{c}_k &= \mathbf{0} \qquad (24) \end{aligned}$$

luego

$$
\begin{aligned}
\hat{\mathcal{L}}_k(\boldsymbol{p}, l) &= \frac{1}{2}\boldsymbol{p}^T \nabla^2_{\boldsymbol{xx}} \mathcal{L}_k \boldsymbol{p} + \nabla f_k^T \boldsymbol{p} - l^T(\mathbf{A}_k \boldsymbol{p} + \boldsymbol{c}_k) && (25)\\
\mathbf{A}_k \boldsymbol{p} + \boldsymbol{c}_k &= \mathbf{0} && (26)
\end{aligned}
$$

que conduce al problema cuadratico equivalente

$$\begin{aligned} \boldsymbol{p}_k &= \arg\min_{\boldsymbol{p}} \frac{1}{2}\boldsymbol{p}^T \nabla^2_{\boldsymbol{xx}} \mathcal{L}_k \boldsymbol{p} + \nabla f_k^T \boldsymbol{p} && (27) \\ & s.a: \ \mathbf{A}_k \boldsymbol{p} + \boldsymbol{c}_k = \mathbf{0} && (28) \end{aligned}$$

o lo que es lo mismo

$$\begin{aligned} \boldsymbol{p}_k &= \arg\min_{\boldsymbol{p}} \frac{1}{2}\boldsymbol{p}^T \nabla^2_{\boldsymbol{xx}} \mathcal{L}_k \boldsymbol{p} + \nabla f_k^T \boldsymbol{p} && (29) \\ & s.a: \ \nabla c_i(\boldsymbol{x}_k)^T \boldsymbol{p} + c_i(\boldsymbol{x}_k) = 0, \ i \in \mathcal{E} && (30) \end{aligned}$$

y si l_k son los multiplicadores de Lagrange del problema anterior, entonces $\boldsymbol{\lambda}_{k+1} = l_k$.

Por otro lado,

$$\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \nabla f(\boldsymbol{x}) - \mathbf{A}(\boldsymbol{x})^T \boldsymbol{\lambda} \tag{31}$$

en el punto $\boldsymbol{x}_k, \boldsymbol{\lambda}_k$ cumple

$$\begin{aligned} \nabla_{\boldsymbol{x}} \mathcal{L}_k \boldsymbol{p} &= (\nabla f_k - \mathbf{A}_k^T \boldsymbol{\lambda}_k)^T \boldsymbol{p} && (32) \\ &= \nabla f_k^T \boldsymbol{p} - \boldsymbol{\lambda}_k^T \mathbf{A}_k^T \boldsymbol{p} && (33) \end{aligned}$$

como $A_k^T \boldsymbol{p} = -\boldsymbol{c}_k$ entonces

$$\nabla_{\boldsymbol{x}} \mathcal{L}_k^T \boldsymbol{p} = \nabla f_k^T \boldsymbol{p} + \boldsymbol{\lambda}_k^T \boldsymbol{c}_k \tag{34}$$

y como $\boldsymbol{\lambda}_k^T \boldsymbol{c}_k$ es una constante, entonces el problema (27)-(28) es equivalente a

$$\begin{aligned} \boldsymbol{p}_k &= \arg\min_{\boldsymbol{p}} q(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T \nabla^2_{\boldsymbol{xx}} \mathcal{L}_k \boldsymbol{p} + \nabla_{\boldsymbol{x}} \mathcal{L}_k^T \boldsymbol{p} && (35) \\ & s.a: \ \mathbf{A}_k \boldsymbol{p} + \boldsymbol{c}_k = \mathbf{0} && (36) \end{aligned}$$

donde $\boldsymbol{\lambda}_{k+1} = l_k$ y l_k son los multiplicadores del Lagrange del problema anterior.

Algorithm 1 Algoritmo SQP con restricciones de igualdad

Dado $\boldsymbol{x}_0, \boldsymbol{\lambda}_0$
for $k = 0, 1, 2, \cdots$ (Hasta Convergencia) **do**
  Hallar $\nabla^2_{\boldsymbol{xx}}\mathcal{L}_k,\ \nabla\mathcal{L}_k,\ \mathbf{A}_k,\ \boldsymbol{c}_k$, segun (10)-(16)
  Obtener $\boldsymbol{p}_k, l_k$, resolviendo (35)-(36) o el equivalente (27)-(28)
  Actualizar $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{p}_k$
  Actualizar $\boldsymbol{\lambda}_{k+1} = l_k$
end for

El algoritmo se puede ver de la siguiente forma:

1. Se obtiene la aproximacion cuadratica del Lagrangiano alrededor de $\boldsymbol{x}_k$ con $\boldsymbol{\lambda}_k$ fijo ie,

$$\mathcal{L}(\boldsymbol{x}_k + \boldsymbol{p}, \boldsymbol{\lambda}_k) \approx \frac{1}{2}\boldsymbol{p}^T \nabla^2_{\boldsymbol{x}\boldsymbol{x}} \mathcal{L}_k \boldsymbol{p} + \nabla_{\boldsymbol{x}} \mathcal{L}_k^T \boldsymbol{p} + \mathcal{L}_k \quad (37)$$

2. Se obtiene la aproximacion lineal de las restricciones,

$$\boldsymbol{c}(\boldsymbol{x}_k + \boldsymbol{p}) \approx \mathbf{A}_k^T \boldsymbol{p} + \boldsymbol{c}_k \quad (38)$$

$$
\begin{aligned}
\min f(\boldsymbol{x}) & & \quad (39)\\
s.a.:\ c_i(\boldsymbol{x}) &= 0,\ i \in \mathcal{E} & \quad (40)\\
c_i(\boldsymbol{x}) &\geq 0,\ i \in \mathcal{I} & \quad (41)
\end{aligned}
$$

Para este problema se sigue el mismo esquema anterior, es decir, se obtiene el paso $\boldsymbol{p}_k$ minimizando la aproximacion cuadratica del Lagrangiano en $\boldsymbol{x}_k$, sujeto a la aproximación lineal de las restricciones, esto es:

$$\boldsymbol{p}_k = \arg\min_{\boldsymbol{p}} \frac{1}{2}\boldsymbol{p}^T \nabla^2_{\boldsymbol{x}\boldsymbol{x}}\mathcal{L}_k \boldsymbol{p} + \nabla f_k^T \boldsymbol{p} \tag{42}$$

$$s.a: \ \nabla c_i(\boldsymbol{x}_k)^T \boldsymbol{p} + c_i(\boldsymbol{x}_k) = 0, \ i \in \mathcal{E} \tag{43}$$

$$\nabla c_i(\boldsymbol{x}_k)^T \boldsymbol{p} + c_i(\boldsymbol{x}_k) \geq 0, \ i \in \mathcal{I} \tag{44}$$

y si l_k son los multiplicadores de Lagrange del problema anterior, entonces $\boldsymbol{\lambda}_{k+1} = l_k$.
Para resolver el subproblema (42)-(44) se puede usar el **metodo se conjuntos activos.**

Teorema

Suponga que $\boldsymbol{x}^*$ es una solucion del problema de optimizacion con restricciones de igualdad y desigualdad y se cumplen las KKTs para algun λ^*. Suponga que se cumplen las LICQ, las resctricciones de complementareidad estrictas y las condiciones suficientes de segundo orden en $(\boldsymbol{x}^*, \lambda^*)$. Entonces, si $(\boldsymbol{x}_k, \lambda_k)$ es suficientemente cercano a $(\boldsymbol{x}^*, \lambda^*)$ entonces existe una solucion local del subproblema cuadratico cuyo conjunto activo A_k es el mismo que el conjunto activo $A(\boldsymbol{x}^*)$ del problema no lineal en el punto $\boldsymbol{x}^*$.

Damped BFGS

En muchas aplicaciones no es facil acceder o calcular $\nabla^2_{\boldsymbol{xx}}\mathcal{L}_k$. En este caso se sugiere usar el metodo **cuasi Newton BFGS** para aproximar el Hessiano,

$$\mathbf{B}_k \approx \nabla^2_{\boldsymbol{xx}}\mathcal{L}_k$$

en funcion de los gradiente, es decir,

$$\begin{aligned}
\boldsymbol{s}_k &= \boldsymbol{x}_{k+1} - \boldsymbol{x}_k \\
\boldsymbol{y}_k &= \nabla_{\boldsymbol{x}}\mathcal{L}_{k+1} - \nabla_{\boldsymbol{x}}\mathcal{L}_k \\
\mathbf{B}_{k+1} &= \mathbf{B}_k - \frac{\mathbf{B}_k \boldsymbol{s}_k \boldsymbol{s}_k^T \mathbf{B}_k}{\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k} + \frac{\boldsymbol{y}_k^T \boldsymbol{y}_k}{\boldsymbol{s}_k^T \boldsymbol{y}_k}
\end{aligned}$$

satisfaciendo la condicion de curvatura $\boldsymbol{s}_k^T \boldsymbol{y}_k > 0$.

Dado que la condicion de curvatura no se puede garantizar que se cumpla, se propone actualizar $\mathbf{B}_{k+1}$ usando BFGS si la condicion de curvatura es suficientemente positiva, es decir, si

$$\boldsymbol{s}_k^T \boldsymbol{y}_k \quad \geq \quad \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k$$

donde $\theta \in [0, 1]$ es un parametro, por ejemplo $\theta = 0{,}2$.

En otro caso, se propone usar una combinación lineal

$$\begin{aligned}
\boldsymbol{s}_k &= \boldsymbol{x}_{k+1} - \boldsymbol{x}_k \\
\boldsymbol{y}_k &= \nabla_{\boldsymbol{x}} \mathcal{L}_{k+1} - \nabla_{\boldsymbol{x}} \mathcal{L}_k \\
\boldsymbol{r}_k &= \theta_k \boldsymbol{y}_k + (1-\theta_k) \mathbf{B}_k \boldsymbol{s}_k \\
\mathbf{B}_{k+1} &= \mathbf{B}_k - \frac{\mathbf{B}_k \boldsymbol{s}_k \boldsymbol{s}_k^T \mathbf{B}_k}{\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k} + \frac{\boldsymbol{r}_k^T \boldsymbol{r}_k}{\boldsymbol{s}_k^T \boldsymbol{r}_k}
\end{aligned}$$

donde

$$\theta_k = \begin{cases} 1 & \text{si } \boldsymbol{s}_k^T \boldsymbol{y}_k \geq \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k \\ \frac{(1-\theta)\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k}{\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k - \boldsymbol{s}_k^T \boldsymbol{y}_k} & \text{si } \boldsymbol{s}_k^T \boldsymbol{y}_k < \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k \end{cases}$$

- Si $\theta_k = 1$ entonces se obtiene la actualización BFGS.
- Si $\theta_k = 0$ entonces $\mathbf{B}_{k+1} = \mathbf{B}_k$.
- Si $0 < \theta_k < 1$ entonces se obtiene una interpolación entre $\mathbf{B}_k$ y BFGS.

- Si $\boldsymbol{s}_k^T\boldsymbol{y}_k < \theta \boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k$ entonces se garantiza la condicion de curvatura y por tanto $\mathbf{B}_{k+1}$ es positiva definida, ie

$$
\begin{aligned}
\boldsymbol{s}_k^T\boldsymbol{r}_k &= \boldsymbol{s}_k^T\left(\frac{(1-\theta)\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k}{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}\boldsymbol{y}_k+\frac{\theta\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}\mathbf{B}_k\boldsymbol{s}_k\right)\\
&= \frac{(1-\theta)\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k}{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}\boldsymbol{s}_k^T\boldsymbol{y}_k+\frac{\theta\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k\\
&= \frac{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k}{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}\left((1-\theta)\boldsymbol{s}_k^T\boldsymbol{y}_k+\theta\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k\right)\\
&= \theta\frac{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}{\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k-\boldsymbol{s}_k^T\boldsymbol{y}_k}\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k\\
&= \theta\boldsymbol{s}_k^T\mathbf{B}_k\boldsymbol{s}_k > 0
\end{aligned}
$$

- Si $\boldsymbol{s}_k^T \boldsymbol{y}_k < \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k$ entonces $0 < \theta_k < 1$.

$$\theta_k = \frac{(1-\theta)\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k}{\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k - \boldsymbol{s}_k^T \boldsymbol{y}_k}$$

Como

$$s_k^T \boldsymbol{y}_k < \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k < \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k$$

entonces $\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k - s_k^T \boldsymbol{y}_k > 0$ y por tanto $\theta_k > 0$.

- Por otro lado

$$
\begin{array}{rcl}
s_k^T \boldsymbol{y}_k & < & \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k \\
0 & < & \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k - s_k^T \boldsymbol{y}_k \\
(1-\theta) \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k & < & \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k - s_k^T \boldsymbol{y}_k \\
\theta_k = \dfrac{(1-\theta) \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k}{\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k - \boldsymbol{s}_k^T \boldsymbol{y}_k} & < & 1
\end{array}
$$

Algorithm 2 Actualizacion Damped BFGS

Dada la matriz $\mathbf{B}_k$ simetrica y definida positiva y $\theta = 0{,}2$
Calcular s_k, y_k y $r_k = \theta_k y_k + (1-\theta_k) B_k s_k$ donde

$$\theta_k = \begin{cases} 1 & \text{si } \boldsymbol{s}_k^T \boldsymbol{y}_k \geq \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k \\ \frac{(1-\theta)\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k}{\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k - \boldsymbol{s}_k^T \boldsymbol{y}_k} & \text{si } \boldsymbol{s}_k^T \boldsymbol{y}_k < \theta \boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k \end{cases}$$

Actualizar

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \frac{\mathbf{B}_k \boldsymbol{s}_k \boldsymbol{s}_k^T \mathbf{B}_k}{\boldsymbol{s}_k^T \mathbf{B}_k \boldsymbol{s}_k} + \frac{\boldsymbol{r}_k^T \boldsymbol{r}_k}{\boldsymbol{s}_k^T \boldsymbol{r}_k}$$