Ejercicio 1

a) $\begin{cases} \min \|x\|^2 \\ \text{st } \sum_i x_i = 1 \end{cases}$

Escribimos el Lagrangeano del problema

$$L(x,\lambda) = \|x\|^2 + \lambda\left(\sum_i x_i - 1\right)$$

Calculemos los gradientes de L

$$\nabla_x L(x,\lambda) = 2x + \lambda \underbrace{\mathbb{1}}_{\text{vector de 1's}}$$

$$\nabla_\lambda L(x,\lambda) = \sum_i x_i - 1$$

Igualamos a 0

$$\Rightarrow \begin{cases} 2x + \lambda \mathbb{1} = 0 & (1) \\ \sum_i x_i = 1 & (2) \end{cases}$$

De (1) tenemos que $x_i = -\frac{\lambda}{2}$

Introducimos lo anterior en (2)

$$\Rightarrow \sum_{i=1}^{n} \frac{-\lambda}{2} = 1 \Rightarrow -n\frac{\lambda}{2} = 1 \Rightarrow \lambda = -\frac{2}{n}$$

$$\Rightarrow x_i = \frac{-\left(\frac{-2}{n}\right)}{2} = \frac{1}{n} \Rightarrow \boxed{\bar{x}_i = 1/n}$$

Verifiquemos la cond. de 2da orden:

$\nabla_{xx} L = 2\,Id > 0 \Rightarrow \bar{x}_i = 1/n$ es la solución del problema

b) $\min \sum x_i$
st $\|x\|^2 = 1$

El Lagrangeano $L(x,\lambda)$

$L(x,\lambda) = \sum x_i + \lambda(\|x\|^2 - 1)$

$\nabla_x L(x,\lambda) = \mathbb{1} + 2\lambda x$

$\nabla_\lambda L(x,\lambda) = \|x\|^2 - 1$

Igualamos los ∇ a 0

$$\begin{cases} \mathbb{1} + 2\lambda x = 0 & (1) \\ \|x\|^2 = 1 & (2) \end{cases}$$

De (1): $x_i = \frac{-1}{2\lambda}$

Introduciendo en (2)

$$\sum x_i^2 = \sum \left(\frac{-1}{2\lambda}\right)^2 = \frac{n}{4\lambda^2} = 1 \Rightarrow \lambda = \pm \frac{\sqrt{n}}{2}$$

$$\Rightarrow x_i = \frac{-1}{2\left(\pm\frac{\sqrt{n}}{2}\right)} = \mp \frac{1}{\sqrt{n}}$$

Escaneado con CamScanner

Verifiquemos la condicion de 2^{do} orden:

$$\nabla_{xx} L(\bar{x}, \bar{\lambda}) = 2\lambda Id = \mp 2 \frac{\sqrt{n}}{2} Id = \mp \sqrt{n}\, Id$$

$\nabla_{xx} L(\bar{x}, \bar{\lambda}) > 0 \Leftrightarrow \mp\sqrt{n} > 0 \Leftrightarrow$ los $\bar{x}_i = \dfrac{-1}{\sqrt{n}}$ son solución

La solución optima del problema es

$$\boxed{\bar{x}_i = -1/\sqrt{n}}$$

c) $\begin{cases} \min & \|x\|^2 \\ \text{st} & x^T Q x = 1 \end{cases}$ con $Q \succeq 0$ y $Q^T = Q$

El Lagrangeano del problema es:

$$L(x,\lambda) = \|x\|^2 + \lambda(x^T Q x - 1)$$

Sus gradientes son:

$$\begin{cases} \nabla_x L(x,\lambda) = 2x + 2\lambda Q x \\ \nabla_\lambda L(x,\lambda) = x^T Q x - 1 \end{cases}$$

Como $\nabla L(\bar{x},\bar{\lambda}) = 0$ es una condición necesaria para que $\bar{x}$ sea óptimo, busquemos el $\bar{x}$ que la cumpla

$$\Rightarrow \nabla L(\bar{x},\bar{\lambda}) = 0 \Leftrightarrow \begin{cases} x + \lambda Q x = 0 \\ x^T Q x - 1 = 0 \end{cases} \Leftrightarrow \begin{cases} Qx = -\frac{1}{\lambda} x & (1) \\ x^T Q x = 1 & (2) \end{cases}$$

Llamemos α_i al i-esimo valor propio de Q. De la ec. (1) sabemos que $\bar{x}$ deberá ser un v.a.p. de Q asociado a $\alpha_i = -\frac{1}{\lambda}$

Si introducimos (1) en (2) se tiene:

$$x^T Q x = x^T\left(-\frac{1}{\lambda} x\right) = -\frac{1}{\lambda}\|x\|^2 = 1$$

$$\Rightarrow \|x\|^2 = -\lambda = \frac{1}{\alpha_i}$$

Por lo tanto, sabemos que $\bar{x}$ además de ser vap de Q deberá tener norma $\frac{1}{\sqrt{\alpha_i}}$.

Sabemos que el problema tiene mínimo porque la función es convexa. Es posible buscar cuál de todos los candidatos optimiza el problema

Lo anterior lleva a que el $\bar{x}$ que optimiza el problema sea el vector propio asociado al valor propio más grande(α) de Q y que además su norma sea $\frac{1}{\sqrt{\alpha}}$

Ejercicio 2

$$(P)\begin{cases} \min \sum_i \alpha_i x_i \\ \text{s.t. } \prod_i x_i^{\alpha_i} = 1 \\ \quad x_i > 0 \quad \forall_i = 1, \dots, n \end{cases}$$

con $\alpha_i \geq 0$
y $\sum_i \alpha_i = 1$

Tomemos que $y_i = \ln x_i \Rightarrow x_i = e^{y_i}$

Apliquemos Ln a ambos lados de la restriccion

$$\Rightarrow \ln\left(\prod_i x_i^{\alpha_i}\right) = \ln(1)$$

$$\Rightarrow \sum_i \ln x_i^{\alpha_i} = 0$$

$$\Rightarrow \boxed{\sum_i \alpha_i y_i = 0}$$

Aplicar Ln a ambos lados y que la restriccia sea la mismo es valido ya que Ln es monotono creciente.

Apliquemos el cambio de variable $x_i = e^{y_i}$ sobre la funcion a minimizar:

$$\sum_i \alpha_i x_i = \sum_i \alpha_i e^{y_i}$$

Notar que si $x_i > 0$ siempre es valido tomar $y_i = \ln x_i$

Un problema equivalente en terminos de y_i es:

$$\begin{cases} \min \sum_i \alpha_i e^{y_i} = f(y) \\ \text{st} \ \sum \alpha_i y_i = 0 = h(y) \end{cases} \qquad (Q)$$

Si hallamos los y_i optimos, luego tenemos los x_i optimos como $x_i = e^{y_i}$

El Lagrangeano de (Q) es:

$$L(y,\lambda) = \sum_i \alpha_i e^{y_i} + \lambda \sum_i \alpha_i y_i$$

Su gradiente respecto a y es:

$$\nabla_y L(y,\lambda) = \begin{pmatrix} \alpha_1(e^{y_1}+\lambda) \\ \alpha_2(e^{y_2}+\lambda) \\ \vdots \\ \alpha_n(e^{y_n}+\lambda) \end{pmatrix}, \text{ Es decir } (\nabla_y L)_j = \alpha_j(e^{y_j}+\lambda)$$

Notar que $\nabla_y L = 0 \Leftrightarrow e^{y_i} = -\lambda \ \forall i$

$$\Rightarrow \boxed{y_i = \ln(-\lambda)}$$

El gradiente respecto a λ es:

$$\nabla_\lambda L(y,\lambda) = h(y) = \sum_i \alpha_i y_i = \sum_i \alpha_i \ln(-\lambda)$$

$$= \ln(-\lambda) \underbrace{\sum_i \alpha_i}_{1} = \ln(-\lambda)$$

Notar que $\nabla_\lambda L = 0 \Leftrightarrow \boxed{\lambda = -1}$

si $\lambda = -1 \Rightarrow \boxed{\bar{y}_i = 0}$

Verifiquemos la condicion de segundo orden:

$$\nabla_{yy} L(y,\lambda) = \begin{pmatrix} \alpha_1 e^{y_1} & 0 & \cdots & & \\ 0 & \alpha_2 e^{y_2} & 0 & \cdots & \\ \vdots & 0 & \alpha_3 e^{y_3} & 0 & \cdots \\ & \vdots & 0 & \ddots & \\ & & \vdots & & \alpha_n e^{y_n} \end{pmatrix}$$

$$\nabla_{yy} L(\bar{y},\lambda) = \begin{pmatrix} \alpha_1 & & & 0 \\ & \alpha_2 & & \\ & & \alpha_3 & \\ 0 & & & \ddots \\ & & & & \alpha_n \end{pmatrix}$$

$\nabla_{yy} L(\bar{y},\lambda) > 0 \Leftrightarrow \alpha_i > 0 \ \forall i$

Pero tenemos que $\alpha_i > 0 \ \forall i \Rightarrow \nabla_{yy}$ es definido positivo

$\Rightarrow \bar{y}_i = 0$ es minimo.

Si $\bar{y}_i = 0$ es minimo de (Q) $\Rightarrow$ $\boxed{\bar{x}_i = 1 \text{ es minimo de (P)}}$

Escaneado con CamScanner

Para probar que $\prod_i x_i^{\alpha_i} \leq \sum_i \alpha_i x_i$ es necesario resolver el siguiente problema de optimización:

$$\begin{cases} \min \sum \alpha_i x_i = f(x) \\ \text{st: } \prod_i x_i^{\alpha_i} = K \\ \quad x_i > 0 \quad \forall i = 1, \dots n, \end{cases} \qquad (B)$$

Este problema se resuelve de forma analoga a (P) y la solucion optima se obtiene en

$$\boxed{\bar{x}_i = K} \quad y \quad \boxed{f(\bar{x}) = \sum \alpha_i K = K}$$

Supongamos que para un x dado se cumple

$$\prod_i x_i^{\alpha_i} = C$$

Sabemos $\sum \alpha_i x_i$ sera como muy chico C

$$\Rightarrow \prod_i x_i^{\alpha_i} = C \leq \sum \alpha_i x_i$$

$$\Rightarrow \boxed{\prod_i x_i^{\alpha_i} \leq \sum_i \alpha_i x_i}$$

Ejercicio 3

$$\begin{cases} \min \sum_{i=1}^{m} \|x-a_i\|^2 = f(x) & \text{con } \hat{a} = \frac{1}{m}\sum_j a_j \\ \text{st } \|x\|^2 = 1 \end{cases}$$

El lagrangeano del problema es:

$$L(x,\lambda) = \sum_j^m \|x-a_j\|^2 + \lambda(\|x\|^2 - 1)$$

Calculemos el gradiente de L

$$* \nabla_x L(x,\lambda) = 2\sum_j^m (x-a_j) + 2\lambda x$$

$$= 2\sum_j^m x - 2\underbrace{\sum_j^m a_j}_{m\hat{a}} + 2\lambda x$$

$$= 2mx - 2\hat{a} + 2\lambda x = 2[x(m+\lambda) - \hat{a}]$$

$$\nabla_x L(x,\lambda) = 0 \Rightarrow \boxed{x = \frac{\hat{a}}{m+\lambda}}^{(1)} \text{ si } \hat{a} \neq 0$$

$$* \nabla_\lambda L(x,\lambda) = \|x\|^2 - 1$$

$$\nabla_\lambda L(x,\lambda) = 0 \Leftrightarrow \boxed{\|x\|^2 = 1} \Rightarrow \sum_i^m x_i^2 = 1$$

Juntando ambos condiciones se tiene:

$$\sum x_i^2 = \sum\left(\frac{\hat{a}_i}{m+\lambda}\right)^2 = \frac{1}{(m+\lambda)^2}\sum \hat{a}_i^2 = \frac{1}{(m+\lambda)^2}\|\hat{a}\|^2 = 1$$

$$\frac{1}{(m+\lambda)^2}\|\hat{a}\|^2 = 1 \iff m+\lambda = \pm\|\hat{a}\|$$

$$\Rightarrow \bar{\lambda} = -m \pm \|\hat{a}\|$$

Vuelvo a introducir en (1)

$$\Rightarrow x = \frac{\hat{a}}{\cancel{m} - \cancel{m} \pm \|a\|} = \pm \frac{\hat{a}}{\|\hat{a}\|}$$

$$\Rightarrow \boxed{\bar{x} = \pm \frac{\hat{a}}{\|\hat{a}\|}}$$

Verifico la condición de 2^{do} orden.

$$\nabla_{xx} L(x,\lambda) = 2(m+\lambda)\, Id$$

$$\nabla_{xx} L(\bar{x},\bar{\lambda}) > 0 \iff \underline{m+\bar{\lambda} > 0}$$

Veamos cuando se cumple esta condicion:

$$m+\bar{\lambda} = \begin{cases} m+(-m+\|\hat{a}\|) = \|\hat{a}\| > 0 & \text{si } \hat{a} \neq 0 \\ m+(-m-\|\hat{a}\|) = -\|\hat{a}\| < 0 & \text{si } \hat{a} \neq 0 \end{cases}$$

Lo anterior implica que

$$\begin{cases} \bar{x} = \dfrac{\hat{a}}{\|\hat{a}\|} & \text{es mimimo} \\ x = -\dfrac{\hat{a}}{\|\hat{a}\|} & \text{es maximo} \end{cases}$$

La solución del problema es $\boxed{\bar{x} = \frac{\hat{a}}{\|\hat{a}\|}}$

¿Qué pasa si $\hat{a}=0$?

Recordar que se debia cumplir

$$\begin{cases} x(m+\lambda)-\hat{a}=0 \\ \|x\|^2=1 \end{cases} \overset{\hat{a}=0}{\Rightarrow} \begin{cases} x(m+\lambda)=0 \\ \|x\|^2=1 \end{cases}$$

Entonces si se cumple que $\lambda=-m$

$\Rightarrow$ todo x de modulo 1 cumple las condiciones. Entoces todo punto es candidato a minima

Veamos que si $\hat{a}=0$ entonces la función es constante sujeta a los $\|x\|=1$

$$f(x)=\sum_{j=1}^{m}\|x-a_j\| = m\|x\|^2+\sum_{j=1}^{m}\|a_j\|^2-2x^T\underbrace{\sum_{j=1}^{M}a_j}_{=\hat{a}=0}$$

$$=m\|x\|^2+\sum_{j=1}^{m}\|a_j\|^2$$

De esta forma se tiene que

$$f(x)=m+\sum_{j=1}^{m}\|a_j\| \quad \text{si} \quad \|x\|=1$$

Por lo tanto, todo punto que cumple la restricción es minimo

Ejercicio 4

a) $$\begin{cases} \min\limits_{x,y,z} \frac{1}{2}\|Ax-\alpha\|^2 + \frac{1}{2}\|Bx-\beta\|^2 + \frac{1}{2}\|Cz-\gamma\|^2 \\ \text{st. } x=y \\ \quad\;\; x=z \end{cases}$$

El lagrangeano del problema es:

$$L(x,y,z,\lambda) = \frac{1}{2}\|Ax-\alpha\|^2 + \frac{1}{2}\|By-\beta\|^2 + \frac{1}{2}\|Cz-\gamma\|^2 + \lambda_1(x-y) + \lambda_2(x-z)$$

Las condiciones que se derivan de $\nabla L = 0$ son:

$$\begin{cases} A^T(Ax-\alpha) + \lambda_1 + \lambda_2 = 0 \\ B^T(By-\beta) - \lambda_1 = 0 \\ C^T(Cz-\gamma) - \lambda_2 = 0 \\ x = y \\ x = z \end{cases}$$

$\Rightarrow$ Se debera cumplir:

$$A^T(Ax-\alpha) + B^T(Bx-\beta) + C^T(Cx-\gamma) = 0$$

$$\Rightarrow \boxed{\bar{x} = \bar{y} = \bar{z} = (A^TA + B^TB + C^TC)^{-1}(A^T\alpha + B^T\beta + C^T\gamma)}$$

b) Sea $w = [x, y, z]$ y

$$\begin{cases} \min\limits_{w \in \mathbb{R}^{3n}} \frac{1}{2} \|Dw - \delta\|^2 \\ \text{st } Hw = 0 \end{cases} \quad \text{con } \delta = [\alpha, \beta, \gamma]$$

* Primeros veamos que H es:

$$\boxed{H = \left(\begin{array}{c|c|c} Id & -Id & 0 \\ \hline Id & 0 & -Id \end{array}\right) \quad \text{con } H \in M^{2n \times 3n}}$$

Se tiene que $Hw = H\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y \\ x - z \end{pmatrix}$ por b que la restricción es equivalente

* Ahora veamos que D es:

$$\boxed{D = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix} \quad \text{con } D \in M^{3n \times 3n}}$$

Ver que $\frac{1}{2} \|Dw - \delta\| = \frac{1}{2} \left\| \begin{pmatrix} Ax \\ By \\ Cz \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \right\|^2$

$$= \frac{1}{2}\left(\sum_i ((Ax)_i - \alpha_i)^2 + \sum_i ((By)_i - \beta_i)^2 + \sum_i ((Cz)_i - \gamma_i)^2 \right)$$

$$= \frac{1}{2}\left(\|Ax - \alpha\|^2 + \|By - \beta\|^2 + \|Cz - \gamma\|^2 \right)$$

c) El Lagrangeano es:

$$L(w,\lambda) = \frac{1}{2}\|Dw-\delta\|^2 + \lambda^T H w$$

Su gradiente:

$$\nabla L_w(w,\lambda) = D^T(Dw-\delta) + H^T\lambda$$

$$\nabla L_\lambda(w,\lambda) = Hw$$

Entonces,
$$\nabla_L(\bar{w},\bar{\lambda}) = 0 \Leftrightarrow \begin{cases} D^T(D\bar{w}-\delta) + H^T\bar{\lambda} = 0 & (1) \\ H\bar{w} = 0 & (2) \end{cases}$$

Despejamos $\bar{w}$ de (1),

$$\Rightarrow D\bar{w} - \delta = -(D^T)^{-1} H^T \bar{\lambda} \Rightarrow \bar{w} = D^{-1}\delta - \overbrace{D^{-1}(D^T)^{-1}}^{(D^TD)^{-1}} H^T \lambda$$

$$\Rightarrow \bar{w} = D^{-1}\delta - (D^TD)^{-1} H^T \bar{\lambda}$$

Introducimos $\bar{w}$ en (2)

$$\Rightarrow H\left(D^{-1}\delta - (D^TD)^{-1} H^T \bar{\lambda}\right) = 0$$

$$\Rightarrow \boxed{\bar{\lambda} = \left(H(D^TD)^{-1}H^T\right)^{-1} H D^{-1}\delta}$$

Tenemos que $\bar{w}$ será

$$\bar{w} = D^{-1}\delta - (D^TD)^{-1} H^T \left(H(D^TD)^{-1}H^T\right)^{-1} H D^{-1}\delta$$

d) El Lagrangeano aumentado (L_c) es:

$$L_c(w,\lambda) = \frac{1}{2}\|Dw-\delta\|^2 + \lambda^T Hw + \frac{\tau}{2}\|Hw\|^2$$

Para calcular el minimo, calculamos ∇L_c y lo igualamos a 0

$$\Rightarrow \nabla L_c = D^T(Dw-\delta) + H^T\lambda + \tau H^T H w$$

$$\nabla L_c = 0 \Leftrightarrow (D^T D + \tau H^T H) w = D^T\delta - H^T\lambda$$

$$\Rightarrow \boxed{\bar{w} = (D^T D + \tau H^T H)^{-1}(D^T\delta - H^T\lambda)}$$

Implementación

Al resolver el problema de forma exacta se obtiene la siguiente solución:

$$\mathbf{x}^* = \begin{bmatrix} -0.45209767 \\ -0.11645923 \\ -0.00416928 \\ -0.06615063 \\ 0.15325308 \\ 0.12561526 \\ -0.0518167 \\ -0.00062787 \\ -0.04588603 \\ 0.00636417 \end{bmatrix}$$

La siguiente gráfica corresponde al error de cada uno de los métodos respecto al óptimo teórico(x^*), es decir se gráfica $\frac{||x^k - x^*||_2}{||x^*||_2}$ en escala vertical logarítmica.

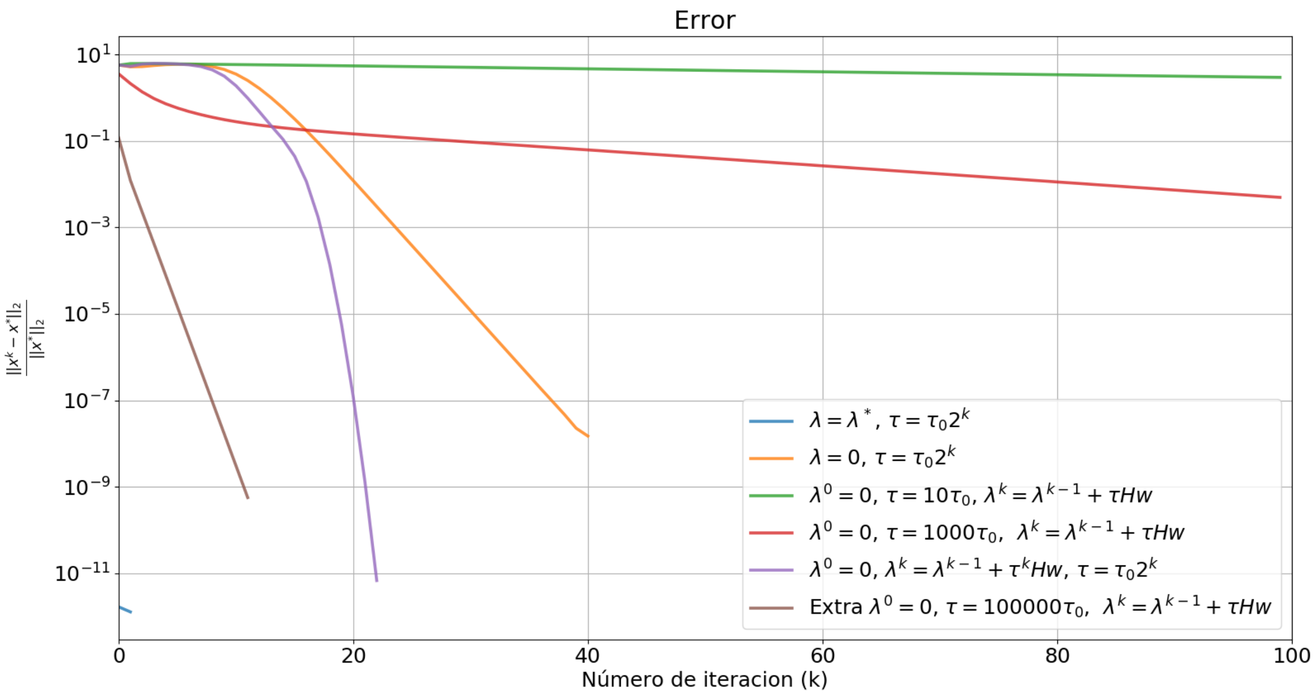

Figure 1: Evolución del error respecto al valor óptimo (x^*) en cada uno de los 5 métodos propuestos

A primera vista esta gráfica puede parecer incorrecta ya que las curvas no terminan, pero esto se debe a que algoritmo finaliza si $\frac{\|w^k - w^{k-1}\|}{\|w^k\|} < 1e{-}8$ o si se alcanzan las 100 iteraciones. En la tabla 1 se muestra en una tabla la cantidad de iteraciones requeridas para cada uno de los métodos implementados.

Método	**Numero de Iteraciones**
Penalización cuadrática $\lambda = \lambda^*$, $\tau = \tau_0 2^k$	1
Penalización cuadrática $\lambda = 0$, $\tau = \tau_0 2^k$	40
Multiplicadores $\lambda^0 = 0$, $\tau = 10\tau_0$, $\lambda^k = \lambda^{k-1} + \tau Hw$	100
Multiplicadores $\lambda^0 = 0$, $\tau = 1000\tau_0$, $\lambda^k = \lambda^{k-1} + \tau Hw$	100
Combinado $\lambda^0 = 0$, $\lambda^k = \lambda^{k-1} + \tau^k Hw$, $\tau = \tau_0 2^k$	22

Table 1: Iteraciones requeridas por cada uno de los métodos.

A partir de la gráfica y la tabla se aprecia que la elección el λ en el método de penalización cuadrática es clave para mejorar la tasa de convergencia, cuando se usa el lambda óptimo el método converge inmediatamente, mientras que si esta elección no es exacta ($\lambda = 0$) se tarda 40 iteraciones.

En cuanto los métodos de multiplicadores, ocurrió que el algoritmo no converge en menos de 100 iteraciones para los τ propuestos en la letra.
Dado que mientras se implementaba se observo que a mayor valor de τ se tenia mayor velocidad, surgió la curiosidad de si el aumentar aun mas el τ, por ejemplo a $100.000\tau_0$ se obtendrá una convergencia mayor. Al probar lo anterior se observo (ver Figura 1) que para este nuevo valor de τ el algoritmo converge rápidamente (11 iteraciones) al punto óptimo.
Se concluye que valor de τ es relevante para el método anterior.

En ultimo lugar se tiene el método combinado, el cual utiliza las virtudes de ambos métodos. Por un lado se aprovecha de la buena actualización de λ que realiza el método de los multiplicadores, y por otro lado utiliza un valor de τ con aumento exponencial que es la clave del método de penalización cuadrática. De esta forma se obtiene una velocidad de convergencia superior a tener uno de los dos parámetros fijos.

A continuación, se presentan los puntos obtenidos por cada método (el subíndice se corresponde con el orden en el que han sido presentado previamente).

$$x_1 = \begin{bmatrix} -0.45209767 \\ -0.11645923 \\ -0.00416928 \\ -0.06615063 \\ 0.15325308 \\ 0.12561526 \\ -0.0518167 \\ -0.00062787 \\ -0.04588603 \\ 0.00636417 \end{bmatrix}, x_2 = \begin{bmatrix} -0.45209767 \\ -0.11645923 \\ -0.00416928 \\ -0.06615063 \\ 0.15325307 \\ 0.12561526 \\ -0.0518167 \\ -0.00062787 \\ -0.04588603 \\ 0.00636417 \end{bmatrix}, x_3 = \begin{bmatrix} 0.49428732 \\ -0.359564 \\ -0.86401806 \\ -0.27124434 \\ -0.03539417 \\ 0.17024413 \\ -0.51730138 \\ 0.54697099 \\ 0.06391845 \\ -0.08557211 \end{bmatrix}$$

$$x_4 = \begin{bmatrix} -0.45180941 \\ -0.11718768 \\ -0.00445706 \\ -0.06611046 \\ 0.15273706 \\ 0.1260817 \\ -0.05162811 \\ -0.0011988 \\ -0.04811751 \\ 0.00630212 \end{bmatrix}, x_5 = \begin{bmatrix} -0.45209767 \\ -0.11645923 \\ -0.00416928 \\ -0.06615063 \\ 0.15325308 \\ 0.12561526 \\ -0.0518167 \\ -0.00062787 \\ -0.04588603 \\ 0.00636417 \end{bmatrix}$$

Lo importante a destacar de los puntos obtenidos es que en todos los casos en el que el algoritmo llega a su condición de parada (en otras palabras, converge) debido a que $\epsilon < 1x10^{-8}$ se obtiene el punto óptimo x^*.
Esto era de esperar ya que estamos en las hipótesis de la **Proposición 1** dada en el teórico, ya que que en todos los casos se toman $\{\lambda^k\}$ acotados y $\{c^k\}$ crecientes tendiendo a infinito. Por lo tanto, al cumplirse las hipótesis, teóricamente se tiene que todo punto limite de la sucesión x^k es un mínimo del problema original. Se observa que lo anterior se verifica de forma practica.

Código

Los fragmentos de código relevantes para el método combinado se encuentra a continuación. En caso de requerir más detalle, junto con este *PDF* se adjunta el archivo *code.py* que contiene la totalidad de código fuente.

```python
import numpy as np

def calc_x_from_w(w):
    xyz = np.reshape(w, (3,-1))
    x = np.mean(xyz, axis=0)
    return x

def calc_error(wk, wk_1):
    error = np.linalg.norm(wk - wk_1) / np.linalg.norm(wk)
    return error

def print_info(k, xk):
    print('El numero de iteraciones es: ', k,'\n')
    print('El punto al que converge es: \n', xk[k])
    print('-'*50, '\n\n')

def w_lagrangeando_aumentado(D, H, S, tauk, lambda_ref):
    return np.linalg.inv(D.T@D+tauk*H.T@H) @ (D.T@S - H.T@lambda_ref)

def metodo_combinado(D, H, S, tau, epsilon=1e-8, MAX_ITER=100,
                                                        verbose=True):
    xk = list()
    ek = list()
    lambda_k = np.zeros(H.shape[0])
    for k in range(MAX_ITER):
        tauk = tau(k)
        lambda_ref = lambda_k
        w_ = w_lagrangeando_aumentado(D, H, S, tauk, lambda_ref)
        xk.append(calc_x_from_w(w_))
        if k>0:
            ek.append(error:=calc_error(xk[k], xk[k-1]))
        if k>0 and error<epsilon:
            break
        lambda_k += tauk*H@w_

    if verbose:
        print_info(k, xk)

    xk = np.array(xk)
    return xk, ek
```

```
xk_h, ek_h = metodo_combinado(D, H, S, lambda k:2**k*tau0)
```