Obligatorio 2: Lagrange

Lucía Bouza - 4.289.797-0

Ejercicio 1

a) $\min \|X\|^2$ sujeto a $\sum x_i = 1$

$$\|x\|^2 = \left(\cancel{\sqrt{\textstyle\sum x_i^2}}\right)^{\cancel{2}} = \sum x_i^2$$

$$\nabla f(x) = \nabla \sum x_i^2 = 2 \cdot X$$

$$\nabla h(x) = \nabla \left(\sum x_i - 1\right) = 1$$

condición necesaria de Lagrange

$$2 \cdot x^* + \lambda^* \cdot 1 = 0$$

Restricción

$$\sum x_i - 1 = 0$$

$$\begin{cases} 2 \cdot x_i + \lambda^* = 0 \longrightarrow x_i = -\dfrac{\lambda^*}{2} \\ \sum x_i - 1 = 0 \end{cases}$$

$$\sum -\frac{\lambda^*}{2} - 1 = 0$$

$$n \cdot -\frac{\lambda^*}{2} - 1 = 0$$

$$\boxed{\lambda^* = -\frac{2}{n}}$$

$$x_i = -\frac{-2/n}{2}$$

$$\boxed{x_i = \frac{1}{n}}$$

condición de 2do orden

$\nabla^2 f(x) = 2 \cdot I$ es semidefinida positiva en las direcciones factibles

$f(x)$ y $h(x)$ son convexas $\longrightarrow$ el mínimo es global.

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

$$\nabla f(x) = \nabla \sum x_i = 1$$
$$\nabla h(x) = \nabla \|x\|^2 - 1 = 2 \cdot x$$

condición necesaria de Lagrange

$$1 + \lambda^* \cdot 2 \cdot x_i^* = 0$$

Restricción

$$\|x\|^2 - 1 = 0 \approx \sum x_i^2 - 1 = 0$$

$$\begin{cases} 1 + \lambda^* \cdot 2 \cdot x_i^* = 0 \longrightarrow x_i^* = -\frac{1}{2\lambda^*} \longrightarrow \boxed{x_i^* = -\frac{1}{\sqrt{n}}} \\ \sum x_i^2 - 1 = 0 \end{cases}$$

$$\sum \left(-\frac{1}{2\lambda^*}\right)^2 - 1 = 0$$

$$n \cdot \frac{1}{4\lambda^{*2}} = 1$$

$$\frac{n}{4} = \lambda^{*2} \longrightarrow \boxed{\frac{\sqrt{n}}{2} = \lambda^*}$$

condición de 2do orden

$\nabla^2 f(x) = 0$ es semidefinido positivo en las direcciones factibles

$f(x)$ y $h(x)$ son convexas $\longrightarrow$ es un mínimo global

c) $\min \|x\|^2$ sujeto a $x^T.Q.x = 1$ / Q simétrica y definida positiva.

$\nabla f(x) = 2x$

$\nabla h(x) = Qx$

condición necesaria de Lagrange

$$\begin{cases} 2x + \lambda.Qx = 0 \\ x^T.Q.x - 1 = 0 \end{cases} \longrightarrow (2.I + \lambda Q).x = 0$$

$$\cancel{\lambda}.\left(\frac{2}{\lambda}I + Q\right)x = 0$$

x vector propio

$-\frac{2}{\lambda}$ valor propio

como tengo n λ_i valores propios → tengo n posible x soluciones del problema original, que son los vectores propios de Q.

a mayor valor propio, menor la norma del vector propio asociado → la solución a $\min \|x\|^2$ es el vector propio de mayor valor propio.

Ejercicio 2

Ejercicio 2

$\alpha_1, \dots \alpha_n$ escalares positivos $\sum \alpha_i = 1$

$\min \sum \alpha_i . x_i$ sujeto a $\prod x_i^{\alpha_i} = 1$ $x_i > 0$ $i \in 0 \dots n$

haremos cambio de variable $y_i = \ln x_i$

$f(x) = \sum \alpha_i . x_i$.

con el cambio de variable: $f(y) = \sum \alpha_i . e^{y_i}$

$\nabla f(y) = \sum \alpha_i . e^{y_i}$

Para $h(x)$ hacemos el siguiente desarrollo:

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

$\sum \ln(x_i^{\alpha_i}) = 0$

$\sum \alpha_i . \ln x_i = 0$

$\sum \alpha_i . y_i = 0 \longrightarrow h(y) = \sum \alpha_i . y_i$

$\nabla h(y) = \alpha$

condición necesaria de lagrange

$\sum \alpha_i . e^{y_i^*} + \lambda^* . \alpha_i = 0$

Restricción

$\sum \alpha_i y_i^* = 0$

$$\begin{cases} \sum \alpha_i . e^{y_i^*} + \lambda^* . \alpha_i = 0 \longrightarrow \alpha_i \left(e^{y_i^*} + \lambda^* \right) = 0 \\ \sum \alpha_i . y_i^* = 0 \end{cases}$$

$\alpha_i > 0$ $\lambda^* = - e^{y_i^*}$

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

sustituyo en la segunda ecuación

$$\sum \alpha_i \cdot y_i^* = 0 \longrightarrow \left.\begin{array}{l}\sum \alpha_i \cdot (\ln(-\lambda)) = 0 \\ \alpha_i > 0 \ \forall i\end{array}\right\} \begin{array}{l}\ln(-\lambda) = 0 \\ -\lambda = 1 \\ \boxed{\lambda = -1}\end{array}$$

$$y_i = \ln(1) = \ln(x_i)$$

$$\boxed{x_i = 1}$$

condición de 2do orden

$\nabla^2 f(x) = 0$ es semidefinida positiva

$f(x)$ y $h(x)$ son convexas $\longrightarrow$ el mínimo encontrado es global.

Demostrar que para cualquier conjunto de x_i positivos y $\alpha_i > 0 \ \forall i$ y $\sum \alpha_i = 1$ se cumple

$$\prod x_i^{\alpha_i} \leq \sum \alpha_i \cdot x_i$$

Demostración

utilizando los resultados anteriores, podemos ver que la restricción ahora será $\prod x_i^{\alpha_i} = \beta$ siendo β positivo.

Podemos reutilizar todo lo realizado, pero la restricción ahora se verá así:

$$\sum \alpha_i \cdot y_i = \ln(\beta)$$

al realizar la sustitución de $y_i^* = \ln(-\lambda^*)$ en la ecuación, nos queda:

$$\sum \alpha_i \cdot \ln(-\lambda^*) = \ln(\beta)$$

$$\ln(-\lambda^*) \cdot \underbrace{\sum \alpha_i}_{1} = \ln(\beta)$$

$$\boxed{\beta = -\lambda^*}$$

$$\left.\begin{array}{l} y_i = \ln(-\lambda^*) = \ln(\beta) \\ y_i = \ln(x_i) \end{array}\right\} \boxed{x_i = \beta}$$

entonces: $\min \sum \alpha_i \cdot x_i = \sum \alpha_i \cdot \beta = \beta \cdot \underbrace{\sum \alpha_i}_{1} = \beta$

y habíamos dicho que $\prod x_i^{\alpha_i} = \beta$

entonces $\prod x_i^{\alpha_i} \leq \sum \alpha_i \cdot x_i$

L.Q.Q.D

Ejercicio 3

m vectores $a_1 \cdots a_m \in \mathbb{R}^n$

$\hat{a}$ centroide ; $\hat{a} = \frac{1}{m} \cdot \sum a_i$

$\min \sum_1^m \|x - a_J\|^2$ sujeto a $\|x\|^2 = 1$

$f(x) = \|x - a_J\|^2 = \sum_1^m (x - a_J)^2 = \sum_1^m (x^2 + a_J^2 - 2.x.a_J)$

$\nabla f(x) = \sum_J 2x - 2a_J$

$\nabla h(x) = 2x$

condición necesaria de lagrange

$\sum_J (2x - 2a_J) + \lambda \cdot 2x = 0$

Restricción

$\|x\|^2 - 1 = 0$

$$\begin{cases} \sum (2x^* - 2a_J) + \lambda^* 2x^* = 0 \\ \|x^*\|^2 - 1 = 0 \end{cases}$$

$\cancel{2} \cdot \sum^m x^* - \cancel{2} \sum^m a_J + \cancel{2} \lambda^* x^* = 0$

$m \cdot x^* - \hat{a} \cdot m + \lambda^* \cdot x^* = 0$

$x^*(m + \lambda^*) = \hat{a} \cdot m$

$x^* = \frac{\hat{a} \cdot m}{m + \lambda^*}$

sustituyo en 2da ecuación

$\left\| \frac{m \cdot \hat{a}}{m + \lambda^*} \right\|^2 - 1 = 0 \qquad \left(\frac{m}{m + \lambda^*}\right)^2 \cdot \|\hat{a}\|^2 = 1$

$$m^2 \cdot \|\hat{a}\|^2 = m^2 + \lambda^{*2} + 2 \cdot m \cdot \lambda^*$$

$$\lambda^{*2} + 2m \cdot \lambda^* + m^2 (1 - \|\hat{a}\|^2) = 0$$

$$\lambda^* = \frac{-2m \pm \sqrt{4m^2 - 4 \cdot (1 - \|\hat{a}\|^2) \cdot m^2}}{2}$$

$$\lambda^* = \frac{-2m \pm \sqrt{\cancel{4m^2} - \cancel{4m^2} + 4 \cdot m^2 \cdot \|\hat{a}\|^2}}{2}$$

$$\lambda^* = \frac{-\cancel{2}m \pm \cancel{2}m \cdot \|\hat{a}\|}{\cancel{2}}$$

$$\boxed{\lambda^* = -m \pm m \cdot \|\hat{a}\|}$$

$$\longrightarrow x = \frac{\hat{a} \cdot \cancel{m}}{\cancel{m} + (-\cancel{m} + \cancel{m} \cdot \|\hat{a}\|}$$

$$\boxed{x = + \frac{\hat{a}}{\|\hat{a}\|}}$$

$$\longrightarrow x = \frac{\hat{a} \cdot \cancel{m}}{-\cancel{m} + (\cancel{m} - m \cdot \|\hat{a}\|)}$$

$$\boxed{x = - \frac{\hat{a}}{\|\hat{a}\|}}$$

condición 2do orden

$$\nabla_x^2 L(x, \lambda) = m - \lambda$$

pruebo con ambos λ^* encontrados

$$\nabla_x^2 L(x, \lambda_1^*) = \cancel{m} - (+\cancel{m} + m \cdot \|\hat{a}\|) = -m \cdot \|\hat{a}\|$$

es negativo $\longrightarrow x = + \frac{\hat{a}}{\|\hat{a}\|}$ es un máximo global.

$$\nabla_x^2 L(x, \lambda_2^*) = \cancel{m} - (\cancel{m} - m \cdot \|\hat{a}\|) = m \cdot \|\hat{a}\|$$

es positivo $\longrightarrow x = - \frac{\hat{a}}{\|\hat{a}\|}$ es un mínimo global si $\hat{a} = 0 \longrightarrow \frac{1}{m} \cdot \sum a_J = 0 \longrightarrow \sum a_J = 0$

desarrollando $\nabla L(x, \lambda)$:

$$\sum 2x^* - 2a_J + \lambda \cdot 2x^* = 0$$

$$\cancel{2}m \cdot x^* \underbrace{- \hat{a} \cdot m}_{0} + \cancel{2}\lambda \cdot x^* = 0$$

$$m \cdot x^* + \lambda \cdot x^* = 0$$

$$x^* (m + \lambda) = 0 \longrightarrow \boxed{\lambda = -m}$$

Restricción $\Rightarrow \|x\|^2 = 1$

condición de segundo orden

$\nabla_x^2 L(x, \lambda) = (m + \lambda) \cdot I + 2 \cdot I$ } $\nabla_x^2 L(x, \lambda^*) = 2 \cdot I$

con λ óptimo $\lambda = -m$

es positiva para todo punto que cumple la restricción $\|x\|^2 = 1 \longrightarrow$ todos los puntos de la restricción son mínimos.

Ejercicio 4

a) $$\nabla f(x,y,z) = \begin{pmatrix} \nabla_x f \\ \nabla_y f \\ \nabla_z f \end{pmatrix} = \begin{pmatrix} A^T.(Ax-\alpha) \\ B^T.(By-\beta) \\ C^T(Cz-\gamma) \end{pmatrix}$$

a partir de la restricción $x=y=z$ genero 3 restricciones:

$$h_1(x,y,z) = x-y=0$$
$$h_2(x,y,z) = x-z=0$$
$$h_3(x,y,z) = y-z=0$$

$$\nabla h_1(x,y,z) = \begin{pmatrix} \nabla_x h_1 \\ \nabla_y h_1 \\ \nabla_z h_1 \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 0 & 0 & \cdots & 0 \end{pmatrix}}^{n}$$

$$\nabla h_2(x,y,z) = \underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ -1 & -1 & \cdots & -1 \end{pmatrix}}_{n}$$

$$\nabla h_3(x,y,z) = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \end{pmatrix}}_{n}$$

sistema ecuaciones

$$\begin{cases} A^T(Ax-\alpha)+\lambda_1+\lambda_2=0 \\ B^T(By-\beta)-\lambda_1+\lambda_3=0 \\ C^T(Cz-\gamma)-\lambda_2-\lambda_3=0 \\ x=y=z \end{cases}$$

$$\lambda_1 = -A^T(Ax-\alpha) - \lambda_2$$

$$\lambda_3 = -B^T(Bx-\beta) - A^T(Ax-\alpha) - \lambda_2$$

$$\cancel{\lambda_2} = C^T(Cx-\gamma) + B^T(Bx-\beta) + A^T(Ax-\alpha) + \cancel{\lambda_2}$$

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

$$C^TCx + B^TBx + A^TAx = C^T\cdot\gamma + B^T\cdot\beta + A^T\cdot\alpha$$

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

b) Para poder reescribir el problema en función de D, δ y H siendo $\delta = [\alpha, \beta, \gamma]$, las matrices D y H deberían ser:

$$D = \begin{array}{cc} \begin{array}{|c|c|c|} \hline A & 0 & 0 \\ \hline 0 & B & 0 \\ \hline 0 & 0 & C \\ \hline \end{array} & 3n \\ 3n & \end{array} \qquad H = \begin{array}{cc} \begin{array}{|c|c|c|} \hline I & -I & 0 \\ \hline I & 0 & -I \\ \hline 0 & I & -I \\ \hline \end{array} & 3n \\ 3n & \end{array}$$

hagamos cuentas:

$$D\cdot w - \delta = \begin{pmatrix} A\cdot x + \cancel{0}\cdot y + \cancel{0}\cdot z \\ \cancel{0}\cdot x + B\cdot y + \cancel{0}\cdot z \\ 0x + 0\cdot y + C\cdot z \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} Ax - \alpha \\ By - \beta \\ Cz - \gamma \end{pmatrix}$$

$$H\cdot w = \begin{pmatrix} I\cdot x & -Iy & +\cancel{0} \\ Ix & +\cancel{0} & -I\cdot z \\ \cancel{0} & Iy & -Iz \end{pmatrix} = \begin{pmatrix} x - y \\ x - z \\ y - z \end{pmatrix}$$

el problema es equivalente al original.

c) $\nabla_w f(w) = \Delta^T . (\Delta w - \delta)$

$\nabla_w h(w) = H^T$

sistema de ecuaciones

$$\begin{cases} \Delta^T(\Delta w - \delta) + H^T . \lambda^* = 0 \\ H^T = 0 \end{cases}$$

de la primer ecuación despejamos λ^*

- multiplicamos por H todo el término ya que H no es invertible.
- Además, en el cálculo de la solución numérica $H . H^T$ es una matriz singular por lo que tampoco es invertible. Para el cálculo de la solución numérica no consideremos una de las condiciones de H, por lo que H puede de tamaño $3n \times 2n$

$$H . \Delta^T(\Delta w - \delta) + H . H^T . \lambda^* = 0$$

$$\boxed{\lambda^* = - H . \Delta^T(\Delta w - \delta) . (H . H^T)^{-1}}$$

d) $L_T(w, \lambda) = \frac{1}{2} \cdot \| \Delta w - \delta \|^2 + \lambda . H . w + \frac{T}{2} \cdot \| H w \|^2$

calculo $\nabla_w L_T(w, \lambda) = 0$

$$\Delta^T(\Delta w - \delta) + \lambda . H + \frac{T}{\cancel{2}} . \cancel{2} . H^T . H . w = 0$$

$$\Delta^T . \Delta . w - \Delta^T \delta + \lambda H + T . H^T H w = 0$$

$$\boxed{w = (\Delta^T \delta - \lambda H) . (\Delta^T \Delta + T . H^T . H)^{-1}}$$

Parte E

Se resuelve numéricamente lambda* y la solución exacta al problema a través del siguiente código.

```
###############################################
#####SOLUCION lambda* analítico #######
###############################################

def lambdaAsterisco (w, D, H, delta):
    aux1 = np.linalg.inv(np.dot(H, H.T))
    aux2 = np.dot(D,w) - delta
    aux3 = np.dot(D.T,aux2)
    aux4 = np.dot(-H,aux3)
    return(np.dot(aux4,aux1))

###############################################
#####SOLUCION exacta analítica #######
###############################################
def solucionExacta(A, B, C, alpha, betha, gamma):
    aux1 = np.dot(A.T,alpha) + np.dot(B.T, betha) + np.dot(C.T,gamma)
    aux2 = np.dot(A.T,A) + np.dot(B.T,B) + np.dot(C.T,C)
    x = np.dot(aux1, np.linalg.inv(aux2))
    return x
```

El resultado para x es: [-0.45209767 -0.11645923 -0.00416928 -0.06615063 0.15325308 0.12561526 -0.0518167 -0.00062787 -0.04588603 0.00636417]

Para el caso de lambda*, como vimos en la resolución analítica, queda en función de w. Para este ejercicio, tomamos como vector w el vector de ceros de tamaño 30.

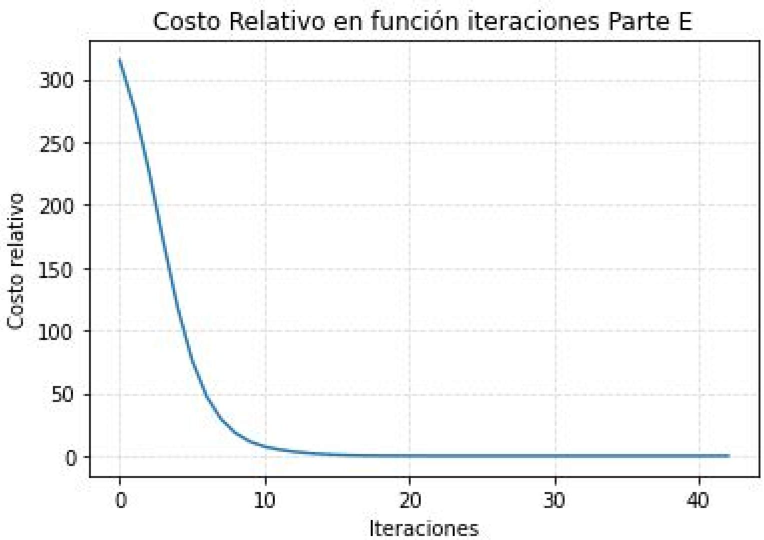
Para este ejercicio, para hallar el mínimo w en cada paso, utilizamos la resolución analítica calculada en el ejercicio 4.d. El código de dicha resolución se expone a continuación:

```
def SolAnaliticaLagrangeano(H, D, delta, lambdaa, t):
    aux1 = np.dot(D.T,delta) - np.dot(lambdaa,H)
    aux2 = np.linalg.inv(np.dot(D.T, D) + t*np.dot(H.T, H))
    return np.dot(aux1, aux2)
```

La resolución de este problema, con penalización cuadrática, utilizando lambda*, da como resultado:[-0.45209767 -0.11645923 -0.00416928 -0.06615063 0.15325308 0.12561526 -0.0518167 -0.00062787 -0.04588602 0.00636417]

Lo cual es una aproximación excelente a la solución exacta (se marca en naranja el valor en el que difieren)

Podemos ver a continuación la gráfica del error relativo en función de las iteraciones. En la gráfica se puede observar que el error rápidamente converge a cero, siempre con pendiente negativa, sin oscilar.

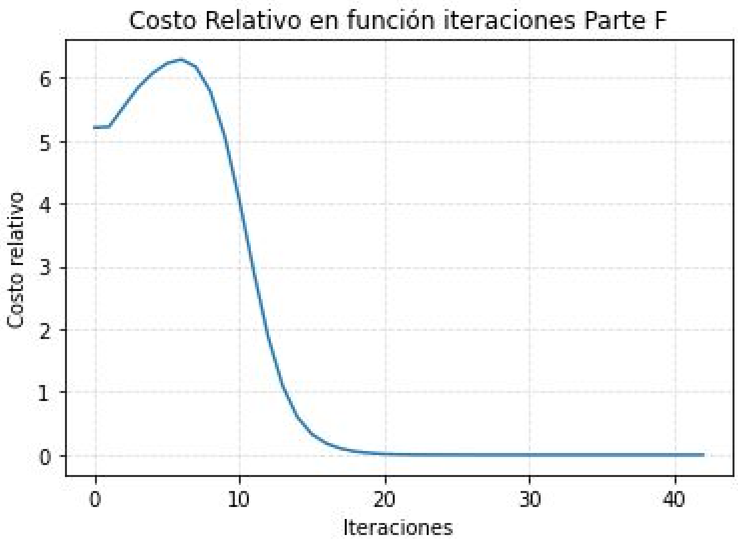


Parte F

En esta sección utilizamos el método de penalización cuadrática, con lambda = 0, y el mismo factor de aumento de tau que en el ejercicio E. La solución obtenida es: [-0.45209767 -0.11645923 -0.00416928 -0.06615063 0.15325307 0.12561526 -0.0518167 -0.00062787 -0.04588603 0.00636417]

La aproximación a la solución exacta también es excelente (marcado en naranja en el valor que difieren).

Podemos ver a continuación la gráfica del error relativo en función de las iteraciones. En la gráfica se puede observar que primero el error aumenta pero luego rápidamente converge a cero, aunque con una pendiente un poco menos inclinada que en el ejercicio anterior. Comparando con el ejercicio E, vemos que para llegar al error relativo aceptable demoran la misma cantidad de iteraciones, pero al tener la solución exacta de lambda*, se converge más rápidamente.

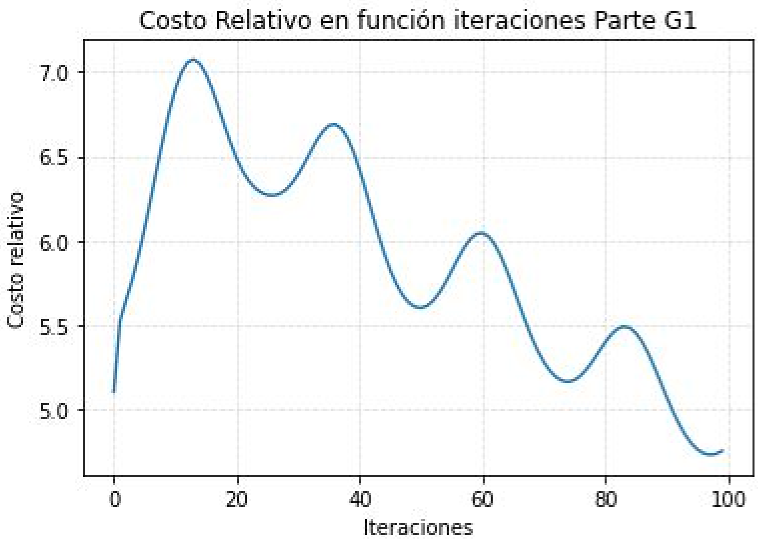


Parte G

En esta sección utilizamos el método de multiplicadores de Lagrange, donde se actualiza lambda y tau se mantiene constante. Se realizan dos pruebas diferentes, con dos valores diferentes de tau, uno 100 veces mayor que el otro. Mostraremos los resultados para ambos valores.

Tau = (1/||D||)*10

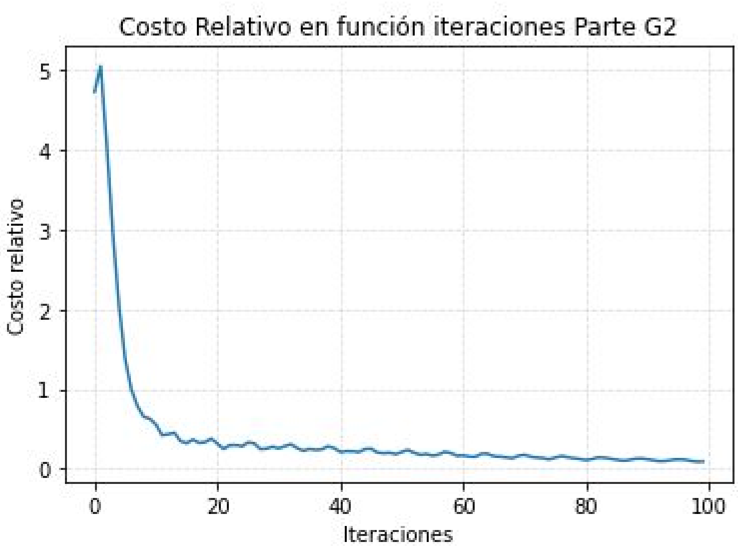
Solución: [1.04021777 -0.67576836 -1.70060387 -0.08835587 -0.03017521 0.72156588
 -0.23350877 0.36764798 -0.17237106 0.19497707]



Podemos observar que el costo relativo promedialmente disminuye, pero lo va haciendo oscilando, es decir aumentando y decreciendo.

Tau = (1/||D||)*1000

Solución: [-0.45235219 -0.12966503 -0.02206411 -0.05739355 0.13338149 0.13834899
 -0.03787135 -0.00180887 -0.07388246 -0.00109606]



Podemos ver que el costo relativo disminuye y converge a cero de una forma mucho más veloz que para Tau más pequeño. Además el error relativo que se alcanza es considerablemente menor al alcanzado con el Tau 100 veces más chico.

Nota: tanto para este ejercicio como para la parte H, al utilizar la resolución analítica para w en cada paso, los resultados no fueron los esperados: la solución no se aproximaba a la solución exacta y el error no disminuye. Se utilizó entonces Descenso por gradiente con Armijo para hallar w en cada paso.

Parte H

En esta sección combinamos el método de los multiplicadores con el método de penalización cuadrática. El código relacionado al método es el siguiente:

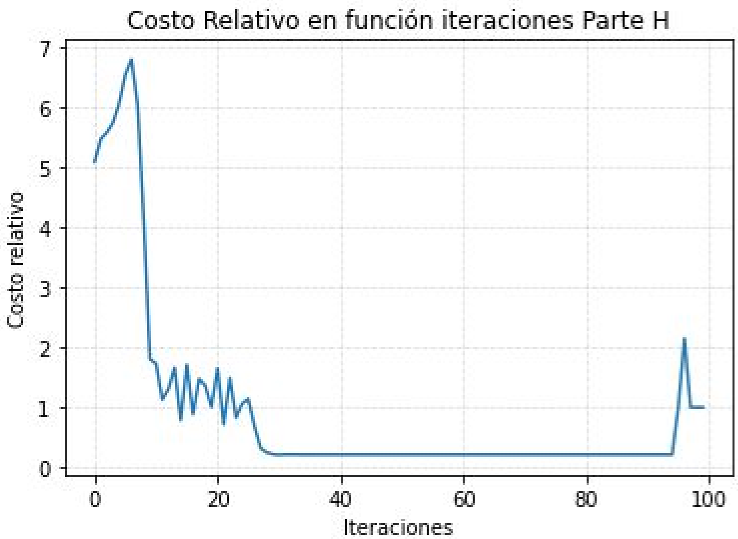
```
def combinado(H, D, delta, lambdaa0, t0, solution):
    #iteraciones
    iteraciones_min_f = 80
    iteraciones_min_lagrangeano = 1000
    #epsilons
    epsilon_f = 1e-8
    epsilon_lagrangeano = 1e-4
    #inicialización variables
    w = w_k = np.zeros(30)
    t = t0
    lambdaa = lambdaa0
    vector_errorRelativo = list()

    for k in range(iteraciones_min_f):
        t = 2 * t
        lambdaa = lambdaa + t * np.dot(H, w)
        w = w_k
        w_k = DescensoPorGradiente(H, w, D, delta, lambdaa, t, iteraciones_min_lagrangeano, epsilon_lagrangeano)
        #w_k = SolAnaliticaLagrangeano(H, D, delta, lambdaa, t)

        'guardo error relativo'
        x_k = (w_k[0:10] + w_k[10:20] + w_k[20:30]) /3
        errorRelativo = np.linalg.norm(solution - x_k)/np.linalg.norm(solution)
        vector_errorRelativo.append(errorRelativo)

        if np.linalg.norm(w_k - w) / np.linalg.norm(w_k) < epsilon_f:
            break
    return w_k, vector_errorRelativo
```

La solución obtenida es la siguiente: [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]

Observando más detenidamente, el vector w es:
[2.70472700e+42 -4.01527451e+41 -4.53900198e+41 -3.40313283e+41
-2.04080464e+41 9.50020750e+41 -4.38826187e+41 1.64342822e+42
7.53091356e+41 1.36039588e+42 -1.35236350e+42 2.00763725e+41
2.26950099e+41 1.70156641e+41 1.02040232e+41 -4.75010375e+41
2.19413094e+41 -8.21714110e+41 -3.76545678e+41 -6.80197939e+41
-1.35236350e+42 2.00763725e+41 2.26950099e+41 1.70156641e+41
1.02040232e+41 -4.75010375e+41 2.19413094e+41 -8.21714110e+41
-3.76545678e+41 -6.80197939e+41]

X, Y y Z no están quedando con los mismos valores. De hecho X = -Y- Z

El costo relativo para este método se puede apreciar en la siguiente gráfica:



En un principio disminuye y se acerca a 0, pero luego vuelve a aumentar.

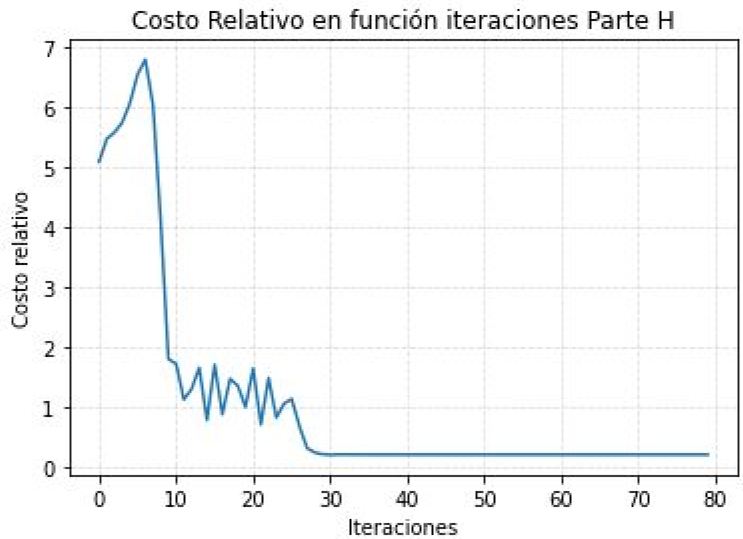
Disminuyendo la cantidad de iteraciones a 80, los resultados son los siguientes:

X = [-0.45404279 -0.19155532 0.02444686 -0.09487543 0.16555149 0.09616699
 -0.01757323 0.03240491 -0.03265633 -0.02929381]

W =
[0.13788799 -0.13645924 0.10908148 -0.29062869 0.42899483 -0.03213832
 0.05994941 -0.14276977 0.10076249 0.24360648 -0.83652971 0.09042076
 0.15091311 -0.57900409 0.53240887 0.06341698 -0.26080321 0.07206135
 0.11033322 -0.45568069 -0.66348666 -0.52862749 -0.18665401 0.58500648
-0.46474924 0.25722231 0.14813411 0.16792316 -0.30906471 0.1241928]

Comparando con la solución exacta: [-0.45209767 -0.11645923 -0.00416928 -0.06615063
0.15325308 0.12561526 -0.0518167 -0.00062787 -0.04588603 0.00636417]

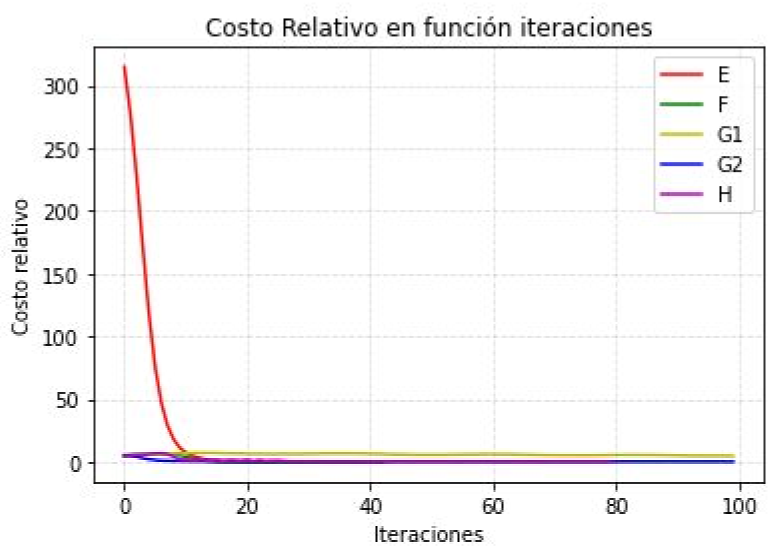
Podemos observar que el valor de X hallado, si bien es menos preciso que para los métodos anteriores, nos da una mejor aproximación al valor obtenido con 100 iteraciones, para el método combinado.



Se supone que lo que sucede es que el error no llega a converger a cero y luego con más iteraciones (entre 80 y 100) tau se dispara y con él la estimación de w.

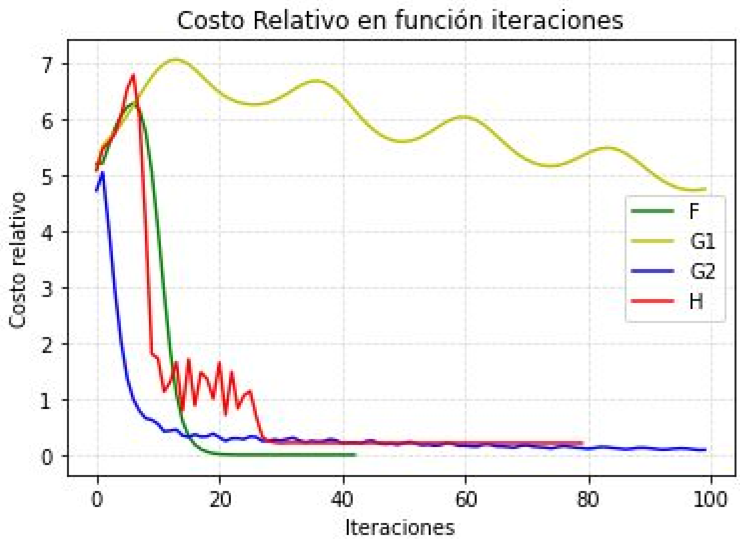
Conclusiones

Graficamos todos los métodos en un mismo plot, para así poder comparar más fácilmente:



El método de penalización cuadrática converge rápidamente, alcanzando los valores de costo relativo aceptables en menos de 50 iteraciones.

Hagamos zoom mirando solamente los casos F, G1, G2 y H para poder sacar conclusiones de ellos:



El método de penalización cuadrática, sigue siendo mejor que el resto de los métodos.
El método de multiplicadores de Lagrange con un Tau más pequeño, demora más en converger que con Tau mayor.
Para el caso del método combinado, podemos ver que alcanza valores de costo relativo similares al método de multiplicadores de Lagrange, sin necesidad de contar con un Tau tan grande (con un Tau 1000 veces menor se alcanzan valores similares en el método combinado).
Recordemos también que para este método fue necesario utilizar menos iteraciones ya que con más de 90 iteraciones comienza a aumentar el costo relativo.