Obligatorio 1

TAO - Teoría y Algoritmia de Optimización - 2022

Instituto de Matemática y Estadística Rafael Laguardia
Facultad de Ingeniería - Universidad de la República

Conformidad

He leído y estoy de acuerdo con las Instrucciones especificadas en la carátula obligatorio. He resuelto por mi propia cuenta los ejercicios, sin recurrir a informes de otros compañeros, o soluciones existentes.
Soy el único autor de este trabajo. El informe y todo programa implementado como parte de la resolución del obligatorio son de mi autoría y no incluyen partes ni fragmentos tomados de otros informes u otras fuentes, salvo las excepciones mencionadas.

Christian Diaz - cdiaz@fing.edu.uy
C.I.: 4.639.064-7
Estudiante del Doctorado en Ingeniería Estructural

Ejercicio 1: Convexidad

Para demostrar la convexidad de funciones se utilizará la siguiente definición:
Sean $\mathbf{u},\ \mathbf{v}\ \in\ \mathbb{R}^n$.

$$\Rightarrow\ g(\mathbf{x})\text{es convexa} \Leftrightarrow g(t\mathbf{u}+(1-t)\mathbf{v}) \leq tg(\mathbf{u})+(1-t)g(\mathbf{v}) \qquad \forall t\ \in\ [0,1] \tag{1}$$

Mientras que para demostrar la convexidad de conjuntos se utilizará la siguiente definición:
Sean $C \subset \mathbb{R}^n$ y $p, q\ \in\ C$

$$\Rightarrow\ C \text{ es convexo} \Leftrightarrow \lambda.p+(1-\lambda).q \quad \in C \quad \forall \lambda \in [0,1] \tag{2}$$

Parte a)

Sea $g(\mathbf{x})=\sum_i w_i f_i(\mathbf{x})$ con $w_i \geq 0$ y $f_i(\mathbf{x})$ funciones convexas.

La desigualdad de la Ec. 1 es equivalente a la Ec. 3

$$tg(\mathbf{u})+(1-t)g(\mathbf{v})-g(t\mathbf{u}+(1-t)\mathbf{v}) \geq 0 \tag{3}$$

Demostración

Usando la definición de $g(\mathbf{x})$ para $\mathbf{u},\ \mathbf{v}\ \in\ \mathbb{R}^n$ y $t\ \in\ [0,1]$ en la Ec. 3
$tg(\mathbf{u})+(1-t)g(\mathbf{v})-g(t\mathbf{u}+(1-t)\mathbf{v}) = t\sum_i w_i f_i(\mathbf{u})+(1-t)\sum_i w_i f_i(\mathbf{v})-\sum_i w_i f_i(t\mathbf{u}+(1-t)\mathbf{v})$

Sacando $\sum_i w_i$ de factor común:
$= \sum_i w_i\left(tf_i(\mathbf{u})+(1-t)f_i(\mathbf{v})-f_i(t\mathbf{u}+(1-t)\mathbf{v})\right)$

Se definen los términos A_i como: $A_i = tf_i(\mathbf{u})+(1-t)f_i(\mathbf{v})-f_i(t\mathbf{u}+(1-t)\mathbf{v})$ queda:
$= \sum_i w_i A_i$

Como $f_i(\mathbf{x})$ son funciones convexas, cada sumando $A_i \geq 0$. $\Rightarrow$ como cada $w_i \geq 0$ el producto de dos factores mayores o iguales a cero (entiéndase $w_i A_i$) será siempre mayor o igual a cero, y por consiguiente, la suma de sumandos todos mayores o iguales a cero, será mayor o igual a cero.
$\Rightarrow \sum_i w_i A_i \geq 0\ \blacksquare$

Parte b)

Sea $h(\mathbf{x}) = max\left\{f_i(\mathbf{x}) : i=1,...,k\right\}$ $\Rightarrow$ será convexa sii se cumple la desigualdad de la Ec. 1.

Demostración

Usando la definición de $h(\mathbf{x})$ para $\mathbf{u},\ \mathbf{v}\ \in\ \mathbb{R}^n$ y $t\ \in\ [0,1]$ en el término de la derecha de la Ec. 1.
$th(\mathbf{u})+(1-t)h(\mathbf{v})=t.max\,\{f_i(\mathbf{u}):i=1,...,k\}+(1-t).max\,\{f_i(\mathbf{v}):i=1,...,k\}$

De aquí en más se omitirá la aclaración de que los subíndices i toman valores $i=1,...,k$ para facilitar el seguimiento de la demostración.
Dado que t y $(1-t)$ son escalares, pueden introducirse dentro de la función $max\{\}$ sin alterar el resultado, entonces la ecuación sigue:
$=max\,\{t.f_i(\mathbf{u}):i=1,...,k\}+max\,\{(1-t).f_i(\mathbf{v}):i=1,...,k\}$

La función máximo cumple con la siguiente propiedad: $max\{\sum_j f_j\}\leq\sum_j(max\{f_j\})$, entonces el resultado anterior es:
$\geq max\{t.f_i(\mathbf{u})+(1-t).f_i(\mathbf{v})\}$

Como $f_i(\mathbf{x})$ son funciones convexas, entonces el resultado anterior es:
$\geq max\{f_i(t\mathbf{u}+(1-t).\mathbf{v})\}=h(t\mathbf{u}+(1-t).\mathbf{v})$

Obteniéndose el siguiente resultado y demostrando la desigualdad de la Ec. 1:
$th(\mathbf{u})+(1-t)h(\mathbf{v})\geq h(t\mathbf{u}+(1-t).\mathbf{v})$ ■

Parte c)

Sea $L(\mathbf{x})=f(\mathbf{Ax}+\mathbf{b})\ \Rightarrow$ será convexa sii se cumple la desigualdad de la Ec. 1.

Demostración

Se define la función lineal $g(\mathbf{x})=\mathbf{Ax}+\mathbf{b}\ \Rightarrow\ L(\mathbf{x})=f(g(\mathbf{x}))$.
Usando la definición de $L(\mathbf{x})$ para $\mathbf{u},\ \mathbf{v}\ \in\ \mathbb{R}^n$ y $t\ \in\ [0,1]$ en el término de la izquierda de la Ec. 1.
$L(t.\mathbf{u}+(1-t).\mathbf{v})=f(g(t.\mathbf{u}+(1-t).\mathbf{v}))$

Por la linealidad de $g(\mathbf{x})$
$=f(t.g(\mathbf{u})+(1-t).g(\mathbf{v}))$

Como $f(x)$ es convexa:
$\leq t.f(g(\mathbf{u}))+(1-t).f(g(\mathbf{v}))$

Utilizando la definición de $g(\mathbf{x})$
$=t.f(\mathbf{Au}+\mathbf{b})+(1-t).f(\mathbf{Av}+\mathbf{b})$

Utilizando la definición de $L(\mathbf{x})$
$=t.L(\mathbf{u})+(1-t).L(\mathbf{v})$

Obteniéndose el siguiente resultado y demostrando la desigualdad de la Ec. 1:
$L(t.\mathbf{u}+(1-t).\mathbf{v})\leq t.L(\mathbf{u})+(1-t).L(\mathbf{v})$ ■

Parte d)

Sea $Y = \bigcap_i X_i$ con $i = 1, .., k$ con $k \in \mathbb{N}$ y X_i conjuntos convexos.
$\Rightarrow$ Y será convexa sii se cumple la condición de la Ec. 2.

Demostración

Sean $p,\ q\ \in\ Y\ \Rightarrow\ p,\ q \in X_i\ \ \forall\ i = 1, ...k$ por la definición del conjunto Y.

$\Rightarrow$ como X_i son convexos:
$s = \lambda.p + (1 - \lambda).q \in X_i\ \ \forall i = 1, ..., k$

Nuevamente por la definición del conjunto Y, si $s \in X_i\ \ \forall i = 1, ..., k$:
$\Rightarrow s = \lambda.p + (1 - \lambda).q \in Y\ \Rightarrow\ Y$ es convexo ■

Parte e)

Sea $B(c, r) = \{x : ||x - c|| \leq r\}$.
$\Rightarrow$ $B(c, r)$ será convexa sii se cumple la condición de la Ec. 2.

Demostración

Sean $p,\ q\ \in\ B(c, r)$
$\Rightarrow$ por definición: $||p - c|| \leq r$ y $||q - c|| \leq r$

Se define además, $s = \lambda.p + (1 - \lambda).q$.
$||s - c|| = ||\lambda.p + (1 - \lambda).q - c||$

Se suma y se resta $c.\lambda$
$= ||\lambda.p - c.\lambda + (1 - \lambda).q - c + c.\lambda||$

Se saca y $(1 - \lambda)$ de factor común:
$= ||\lambda.(p - c) + (1 - \lambda).(q - c)||$

por la desigualdad triangular se tiene que:
$\leq ||\lambda.(p - c)|| + ||(1 - \lambda).(q - c)||$

Dado que λ y $(1 - \lambda)$ son escalares, se pueden sacar para afuera de la norma.
$= \lambda||(p - c)|| + (1 - \lambda)||(q - c)||$

Puesto que p y q pertenecen a la bola de radio r y centro c:
$\leq \lambda.r + (1 - \lambda).r$

Aplicando distributiva y operando:
$= r$

Obteniéndose el siguiente resultado y demostrando la desigualdad de la Ec. 2:
$||s - c|| \leq r$ ■

Parte f)

Sea $C = \{\sum_i \theta_i x_i : \sum_i \theta_i = 1, \theta_i \geq 0, \forall i\}$ la envolvente de los puntos $x_1, x_2, ..., x_k \ \in \mathbb{R}^n$
$\Rightarrow \ C$ será convexo sii se cumple la condición de la Ec. 2.

Demostración

Sean $a, \ b \ \in \ C$
$\Rightarrow$ por definición:
$a = \sum_i \alpha_i x_i : \ \sum_i \alpha_i = 1, \ \alpha_i \geq 0, \ \forall i$
$b = \sum_i \beta_i x_i : \ \sum_i \beta_i = 1, \ \beta_i \geq 0, \ \forall i$

Se define: $s = \lambda.a + (1-\lambda).b$
$\Rightarrow \ s = \lambda \sum_i \alpha_i x_i + (1-\lambda) \sum_i \beta_i x_i$

Aplicando distributiva para llevar los escalares λ y $(1-\lambda)$ dentro de la sumatoria:
$= \sum_i \lambda \alpha_i x_i + \sum_i (1-\lambda)\beta_i x_i$

Se define: $\delta_i = \lambda\alpha_i + (1-\lambda)\beta_i \ \ \forall i$:
$= \sum_i \delta_i x_i$ **resultado 1.**

Analizando δ_i:
$\delta_i = \lambda\alpha_i + (1-\lambda)\beta_i$:

$$\left.\begin{array}{c} 0 \leq \lambda \leq 1 \\ \alpha_i \geq 0 \\ \beta_i \geq 0 \end{array}\right\} \left.\begin{array}{c} \lambda\alpha_i \geq 0 \\ (1-\lambda)\beta_i \end{array}\right\} \rightarrow \delta_i \geq 0 \ \forall i \ \textbf{resultado 2}$$

Analizando $\sum_i \delta_i$:
$\sum_i \delta_i = \sum_i (\lambda\alpha_i + (1-\lambda)\beta_i)$

Aplicando las sumatorias por separado, y sacando de factor común λ y $(1-\lambda)$
$= \lambda \sum_i \alpha_i + (1-\lambda) \sum_i \beta_i$

Usando que $\sum_i \alpha_i = 1$ y $\sum_i \beta_i = 1$:
$= \lambda + (1-\lambda) = \lambda$

Se obtiene el **resultado 3:**
$\sum_i \delta_i = 1$

Entonces, a partir de los **resultados 1, 2 y 3** se obtiene que el punto s:
$s = \delta_i x_i, \ \sum_i \delta_i = 1, \ \delta_i \geq 0 \quad \Rightarrow \quad s \in C$ ■

Ejercicio 2: Interpretación geométrica

Parte a)

Sea $f(x,y) = -log(y^2 - x^2)$ se demostrará su convexidad:

$$\begin{aligned} f(x,y) &= -log(y^2 - x^2) \\ &= -log\left((y+x)(y-x)\right) \\ &= -log(y+x) - log(y-x) \end{aligned}$$

Definiendo las siguientes funciones:

$$\left.\begin{aligned} g_1(x,y) = -log(y+x) \\ g_2(x,y) = -log(y-x) \end{aligned}\right\} \Rightarrow \ f(x,y) = g_1(x,y) + g_2(x,y)$$

Entonces si $g_1(x,y)$ y $g_2(x,y)$ son convexas, $f(x,y)$ será convexa por ser la suma de dos funciones convexas con coeficientes positivos (la unidad), como se demostró en el apartado a) del Ej. 1.

Se definen los mapeadores lineales $T_1(\mathbf{u})$ y $T_2(\mathbf{u})$:
$T_1(\mathbf{u}) = \mathbf{A}_1\mathbf{u} + \mathbf{b}$
$T_2(\mathbf{u}) = \mathbf{A}_2\mathbf{u} + \mathbf{b}$
Donde:
$\mathbf{A}_1 = (1,1)$
$\mathbf{A}_2 = (-1,1)$
$\mathbf{b} = (0,0)^T$
$\mathbf{u} = (x,y)^T$

Esto transforma las funciones $g_i(x,y)$ a lo siguiente:
$g_1(x,y) = -log(y+x) = -log(T_1(\mathbf{u}))$
$g_2(x,y) = -log(y-x) = -log(T_2(\mathbf{u}))$

Como se demostró en el apartado c) del Ej.1: Si $-log(x)$ es convexa $\Rightarrow \ -log(\mathbf{A}x + \mathbf{b})$ es convexa.
De esta forma, solo resta demostrar que $l(u) = -log(u) : \mathbb{R}^2 \to \mathbb{R}$ es convexa.
Utilizando la siguiente proposición:
Si $\nabla^2 l(\mathbf{u}) \succeq 0 \ \forall \ u \ \in \ Dom(l) \ \Rightarrow \ l(\mathbf{u})$ es convexa.

Aplicando la proposición a la $l(u)$:

$$\nabla^2(-log(u)) = \frac{\partial^2(-log(u))}{\partial u^2} = \frac{\partial - 1/u}{\partial u} = \frac{1}{u^2}$$

$\Rightarrow \ \nabla^2(-log(u)) \succeq 0 \ \forall u \neq 0$ dado que $u^2 \geq 0$.

Se analiza entonces el dominio de $f(x,y) = -log(y^2 - x^2)$:

$$\begin{cases} x^2 + y^2 \leq 1 \\ 2x - y \leq 1 \\ y \geq 1/2 \\ x \geq 0 \end{cases}$$

Considerando que u esta restrigido a que $u = x + y$ y $u = y - x$ hay dos opciones para analizar:

Opción 1: $u = x + y \neq 0 \Leftrightarrow x \neq -y$, entonces dado que tanto x como y son mayores o iguales a cero, $x = -y$ solo en el valor 0, pero $y \geq 1/2$ ■

Opción 2: $u = y - x \neq 0 \Leftrightarrow y \neq x$, pero dado que el valor de y esta restringido a $y \geq 2x$ solo podrían ser iguales cuando toman el valor 0, pero nuevamente $y \geq 1/2$ ■
.
Entonces la función de costo $f(x, y)$ es convexa. ■

Parte b)

Sea el conjunto factible $X := \{x^2 + y^2 \leq 1/2,\ 2x - y \leq 0,\ y \geq 1/2,\ x \geq 0\}$ se demostrará su convexidad:

El conjunto factible X puede considerarse como $X = \bigcap_i X_i$ con $i = 1, 2, 3, 4$:

$$\text{Considerando: } \begin{cases} X_1 = (x, y) : x^2 + y^2 \leq 1 \\ X_2 = (x, y) : 2x - y \leq 1 \\ X_3 = (x, y) : y \geq 1/2 \\ X_4 = (x, y) : x \geq 0 \end{cases}$$

Entonces, por la demostración del apartado d) del Ej. 1, si cada X_i es convexo, entonces X será convexo.

Analizando X_1:
Considerando la demostración del apartado e): *Toda bola de radio r y centro c es convexa*, entonces X_1 es convexa por ser una bola en $\mathbb{R}^2$ de radio 1 y centro en el origen. ■

Analizando X_2:
Se utiliza la Ec. 2 considerando $\mathbf{u} = (x_1, y_1)$ y $\mathbf{v} = (x_2, y_2)\ \in\ X_2$
Esto significa: $[2x_1 - y_1 \leq 0, 2x_2 - y_2 \leq 0]$

Sea $\mathbf{w} = \lambda.\mathbf{u} + (1 - \lambda)v$ con $\lambda \in [0, 1]$

Sustituyendo $\mathbf{u}$ y $\mathbf{v}$ y operando:
$\mathbf{w} = (\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2)$

$\Rightarrow \mathbf{w} \in X_4 \Leftrightarrow 2\lambda x_1 + 2(1 - \lambda)x_2 - \lambda y_1 - (1 - \lambda)y_2 \leq 0$:

Sacando λ y $(1 - \lambda)$ de factor común, se obtiene:
$\lambda(2x_1 - y_1) + (1 - \lambda)(2x_2 - y_2)$

Analizando ambos sumando por separado:
$\lambda(2x_1 - y_1) \geq 0$ Dado que $2x_1 - y_1 \geq 0$ pues $\mathbf{u} \in X_2$ y además $\lambda \in [0, 1]$
$(1 - \lambda)(2x_2 - y_2) \geq 0$ Dado que $2x_2 - y_2 \geq 0$ pues $\mathbf{v} \in X_2$ y además $(1 - \lambda) \in [0, 1]$

Como la suma de términos positivos, es positivo $\Rightarrow \mathbf{w} \in\ X_2\ \Rightarrow\ X_2$ es convexo. ■

Analizando X_3:
Se utiliza la Ec. 2 considerando $\mathbf{u} = (x_1, y_1)$ y $\mathbf{v} = (x_2, y_2) \in X_3$
Esto significa: $[y_1, y_2 \geq 1/2]$

Sea $\mathbf{w} = \lambda.\mathbf{u} + (1 - \lambda)v$ con $\lambda \in [0, 1]$

Sustituyendo **u** y **v** y operando:
$\mathbf{w} = (\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2)$

$\Rightarrow \mathbf{w} \in X_4 \Leftrightarrow \lambda y_1 + (1 - \lambda)y_2) \geq 1/2.$

Dado que $y_1, y_2 \geq 1/2$:
$\lambda y_1 + (1 - \lambda)y_2) \geq \lambda 1/2 + (1 - \lambda)1/2$

Sacando el término $1/2$ de factor común y operando:
$= 1/2(\lambda + (1 - \lambda)) = 1/2$

Obteniéndose de esta forma que $\lambda y_1 + (1 - \lambda)y_2) \geq 1/2$
Entonces $\mathbf{w} \in X_3 \;\Rightarrow\; X_3$ es convexo. ∎

Analizando X_4:
Se utiliza la Ec. 2 considerando $\mathbf{u} = (x_1, y_1)$ y $\mathbf{v} = (x_2, y_2) \in X_4 \Rightarrow [x_1, x_2 \geq 0]$

Sea $\mathbf{w} = \lambda.\mathbf{u} + (1 - \lambda)v$ con $\lambda \in [0, 1]$

Sustituyendo **u** y **v** y operando:
$\mathbf{w} = (\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2)$

$\Rightarrow \mathbf{w} \in X_4 \Leftrightarrow \lambda x_1 + (1 - \lambda)x_2 \geq 0 \Rightarrow$ analicemos ambos sumando por separado:
$\lambda x_1 \geq 0$ dado que $x_1 \geq 0$ y $\lambda \in [0, 1]$
$(1 - \lambda)x_2 \geq 0$ dado que $x_2 \geq 0$ y $(1 - \lambda) \in [0, 1]$

Como la suma de términos positivos, es positivo $\Rightarrow \mathbf{w} \in \; X_4 \;\Rightarrow\; X_4$ es convexo. ∎

Parte c y d)

Dado el siguiente problema:

$$min\ f(x, y) = -log(y^2 - x^2)$$

$$s.t. \begin{cases} x^2 + y^2 \leq 1 \\ 2x - y \leq 1 \\ y \geq 1/2 \\ x \geq 0 \end{cases}$$

En la Fig. 1 se bocetan en diferentes colores las cuatro restricciones que se aplican sobre el espacio $\mathbb{R}^2$. Allí se puede apreciar que la zona delimitada con el polígono negro es la región factible.

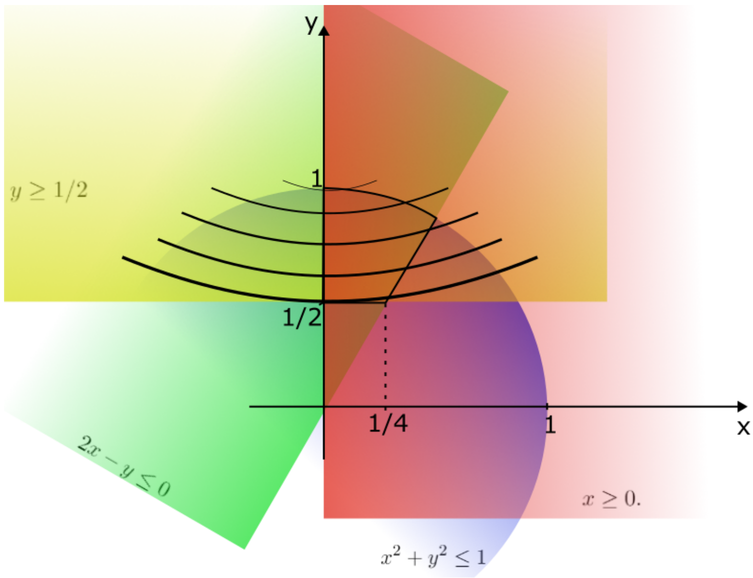

Figura 1: Región factible y curvas de nivel para el ejercicio 2.

Además, se aprecian las curvas de nivel que son aquellas curvas negras de concavidad positiva, centradas en el eje y. El nivel de estas curvas esta asociado directamente al espesor de las mismas, por lo que se puede apreciar que disminuyen a medida que se alejan del eje x. También se observa que para un valor y=cte, el valor funcional crece a medida que se aleja del eje y. De esta forma se puede deducir gráficamente que el punto de la región factible donde se da el mínimo valor funcional tiene que ser el *(0,1)*, al ser el punto mas alejado del eje x sobre el eje y que es la dirección de decrecimiento del funcional.

Para corroborar se implemento un algoritmo que grafique la región factible y las curvas de nivel para diferentes valores, obteniéndose la Figura 2.

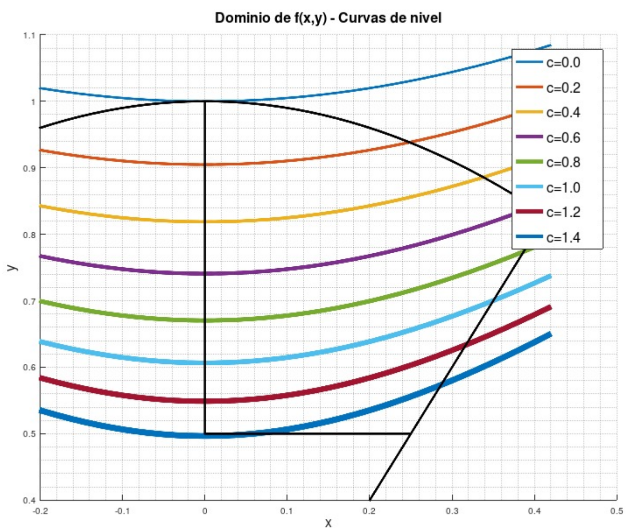

Figura 2: Región factible y curvas de nivel para el ejercicio 2 calculadas.

Ejercicio 3: Puntos críticos y óptimos globales

Parte a)

Dado el siguiente problema:

$$min 4x^4 - x^3 - 4x^2 + 1$$

$$s.t. -1 \leq x \leq 1$$

La Figura 3 muestra el valor funcional para la región factible. En ella se observan claramente dos mínimos locales, el primero en el entorno del valor $x_1 = -0{,}38$ y el segundo exactamente cuando $x_2 = 0{,}81$. Este último, se corresponde con el mínimo global ya que $f(x_2) \leq f(x) \forall x \in [-1, 1]$.

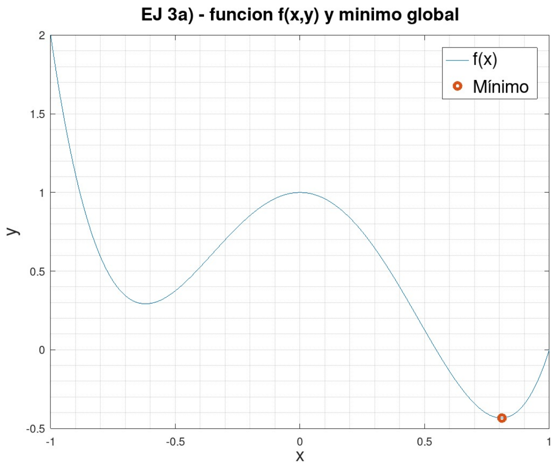

Figura 3: Valor funcional de la función costo para la región factible y localización del mínimo.

Por otro lado, el gradiente de la función de costo se anula en 3 puntos (al ser un polinomio de tercer orden) y coincide con que todos ellos caen en la región factible. Dos de estos corresponden a los mínimos ya mencionados, mientras que el tercero se da en $x_3 = 0$ y coincide con un máximo local, ya que $\varepsilon : f(x_3) \geq f(x) \forall x \in [x - \varepsilon, x + \varepsilon]$.

Parte b)

Dado el siguiente problema:

$$min x^3$$

$$s.t. -1 \leq x \leq 1$$

La Figura 4 muestra el valor funcional para la región factible. En ella se observan claramente un mínimo global sobre una de las restricciones ($x_1 = -1$). Se debe destacar que en x_1 se encuentra un mínimo global puesto que no existe otro valor funcional menor *dentro de la región factible*. No obstante, si la región factible creciera hacia $-x$, el punto del dominio donde se da el mínimo global también se movería hacia $-x$.

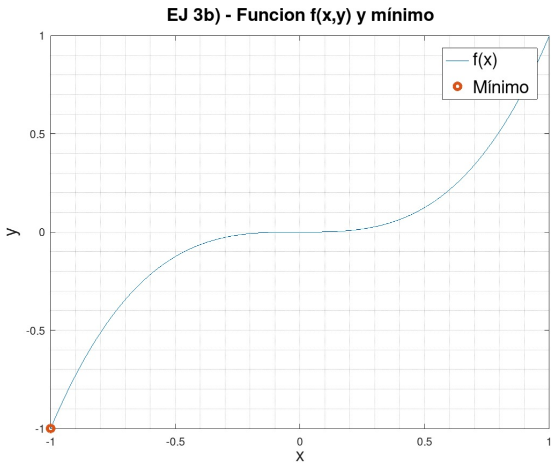

Figura 4: Valor funcional de la función costo para la región factible y localización del mínimo.

Cabe destacar que para esta función de costo, el punto donde se da el mínimo no coincide con el punto donde se anula el gradiente. La anulación del gradiente indica que la tangente a la función en dicho punto tiene pendiente cero, pero no da información sobre concavidad. El punto donde se anula el gradiente en este caso tiene concavidad positiva hacia $+x$ y negativa hacia $-x$, por lo que no es un extremo relativo, sino un punto silla.

Parte c)

Dado el siguiente problema:

$$min(x-a)^2+1$$

$$s.t. -1 \leq x \leq 1$$

En este caso se tiene que la función costo es una parábola parametrizada en el valor de a. La anulación del gradiente en este caso si nos da el punto en el cual se dará el mínimo de la función puesto que la función es convexa. Además si el punto donde se anula el gradiente de la función, se da dentro de la región factible coincidirá con punto donde se encuentre el mínimo global.

Al derivar la función de costo e igualar a cero se encuentra que el mínimo se dará en a, por lo que el problema debe dividirse en dos situaciones: a) cuando $a \in$ region factible, b) cuando $a \notin$ región factible.

Caso a) Cuando $a \in$ región factible, será en este punto que se encuentre el mínimo y además será el mínimo global, como se ve en la Figura 5.

Caso b) Cuando $a \notin$ región factible, el mínimo se dará en la restricción que se encuentre activa. Es decir, si $a \leq -1$ el mínimo de la función se encontrará en $x_1 = -1$ como se puede ver en la Figura 6. Mientras que si $a \geq 1$ el mínimo se encontrara en $x_2 = 1$ como se ve en la Figura 7.

Cabe destacar que en todos los casos el mínimo que se encuentra es global, aunque no siempre coincide con el punto donde se anula el gradiente como ya se expresó anteriormente.

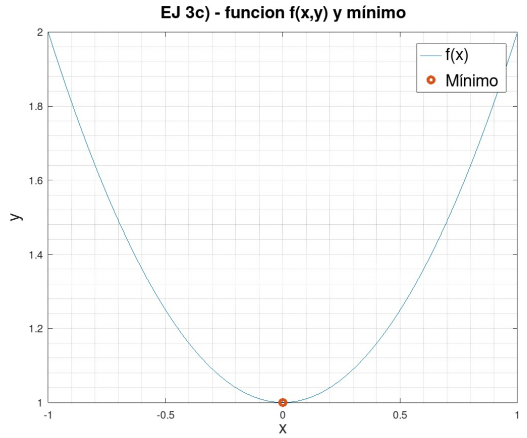

Figura 5: Valor funcional de la función costo para la región factible y localización del mínimo para a=0.

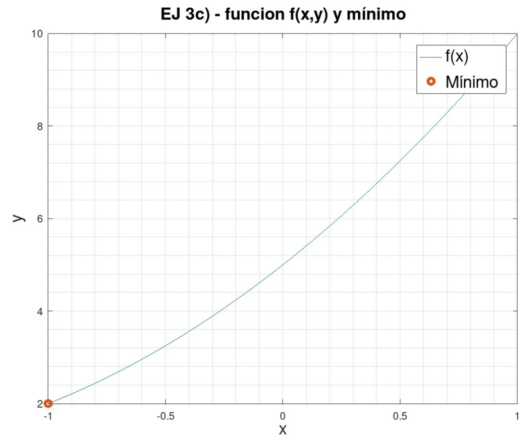

Figura 6: Valor funcional de la función costo para la región factible y localización del mínimo para a=-1.

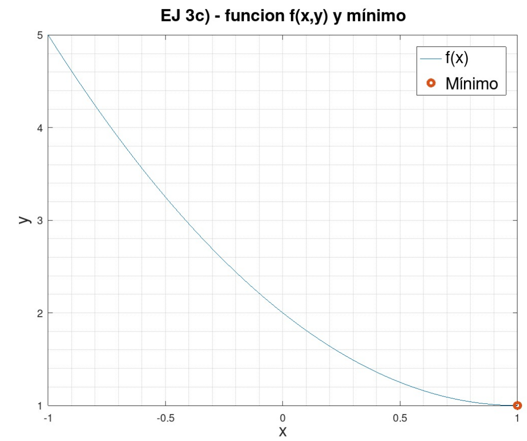

Figura 7: Valor funcional de la función costo para la región factible y localización del mínimo para a=1.

Parte d)

Dado el siguiente problema:

$$min||x - \bar{x}||^2 + 1$$

$$s.t. x \in [0,1]^2$$

Al graficar las curvas de nivel para la función costo, se obtiene la Figura 8 donde el espesor de cada curva tiene una correspondencia directa con el nivel que representa.De esta forma se puede apreciar que en la región factible el mínimo se encuentra en el punto $\mathbf{x} = (0{,}5, 0)$

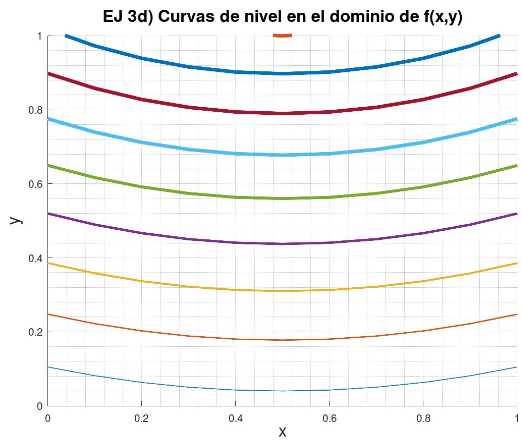

Figura 8: Curvas de nivel en la región factible para el problema d).

Cada restricción implica la zona factible de cada componente de la variable. Es decir, el conjunto S_1 restringe el rango de valores que puede tomar la primer componente de la variable $\mathbf{x}$ en el rango de $[0, 1]$ y esto se puede ver en la Figura 9 como el conjunto rojo. Por otro lado, el conjunto S_2 restringe el valor posible de la segunda componente de la variable $\mathbf{x}$ en el rango $[0, 1]$ que en la misma figura se ve como la zona azul. La zona donde ambos conjuntos se intersectan es conocida como la región factible ($[0, 1]^2$) de la función de costo.

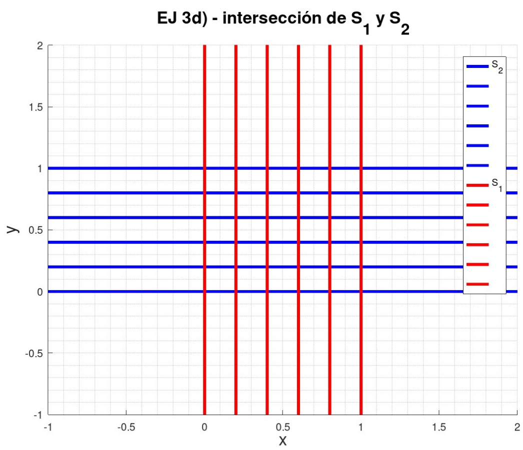

Figura 9: Intersección de los conjuntos que determinan la región factible.

En la Figura 10 se muestra el valor funcional de función de costo en la región factible. Además, se marcó también el punto donde se da el mínimo dentro de la región factible que coincide con el mencionado anteriormente.

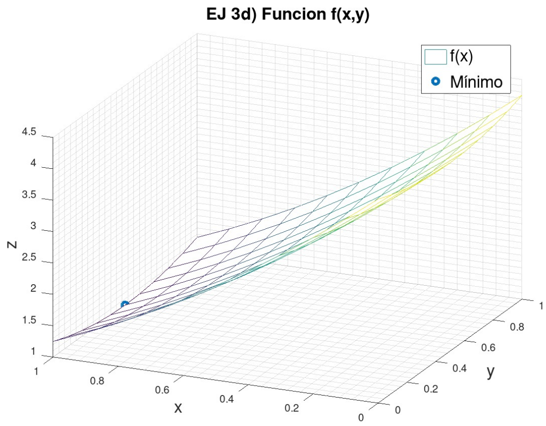

Figura 10: Valor funcional de la función costo para la región factible y localización del mínimo..

Al realizar el gradiente de la función de costo se obtiene lo siguiente:

$$\nabla||(x,y)-(2,1/2)||^2 =< (x-2, y-1/2), (x-2, y-1/2) >$$
$$(2(x-2), 2(y-1/2)) = 0 \Leftrightarrow x = 2 \text{ y } y = 1/2$$

Por lo que se obtiene que el mínimo valor de la función costo se da en el punto $(2, 1/2)$ que queda fuera de la región factible. Por lo que la restricción S_1 está activa, y el mínimo en la región factible se dará en una de sus fronteras. En este caso, la frontera activa es $x = 1$, como se aprecia en la Figura 10.

Para analizar si la función costo es convexa se analiza si la matriz conformada por las segundas derivadas es semidefinida positiva.

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Del resultado anterior se puede determinar que la matriz Hessiana es definida positiva, dado que el $|\nabla^2 f| > 0$, lo que indica que la función de costo es convexa.

Para el análisis de la región factible se utilizará el resultado de la parte f del Ej. 1.

Sean:

$p_1 = (0, 0)$
$p_2 = (0, 1)$
$p_3 = (1, 1)$
$p_4 = (1, 0)$

La región factible $[0, 1]^2$ se puede describir equivalentemente como la envolvente de los puntos p_i con $i = 1, 2, 3$ y 4. Entonces como la envolvente de k cantidad de puntos es convexa $\Rightarrow$ la región factible también es convexa.

Ejercicio 4: Descenso por gradiente

Parte a)

Condición de optimalidad para un problema de minimización sin restricciones:

$\Rightarrow$ si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable y convexa.
$\Rightarrow$ $\nabla f(x^*) = 0$ es condición necesaria y suficiente para que x^* sea mínimo global.

Sea $f = ||\mathbf{Ax\text{-}b}||^2$ con $\mathbf{A} \in \mathbb{R}^{n\times n}$ y $\mathbf{b} \in \mathbb{R}^n$.
Solución analítica del problema se halla como sigue:
$f = ||\mathbf{Ax\text{-}b}||^2$
$= [(\mathbf{Ax\text{-}b})^T(\mathbf{Ax\text{-}b})]$

Por propiedad de matrices:
$= [(\mathbf{x}^T\mathbf{A}^T - \mathbf{b}^T)(\mathbf{Ax\text{-}b})]$

Realizando el producto:
$= [\mathbf{x}^T\mathbf{A}^T\mathbf{A}x - \mathbf{x}^T\mathbf{A}^T\mathbf{b} - \mathbf{b}^T\mathbf{Ax} + \mathbf{b}^T\mathbf{b}]$

Por propiedad de matrices:
$= [\mathbf{x}^T\mathbf{A}^T\mathbf{A}x - 2\mathbf{b}^T\mathbf{Ax} + \mathbf{b}^T\mathbf{b}]$

Llamando $\mathbf{M} = \mathbf{A}^T\mathbf{A}$ y derivando aplicando reglas de derivadas de matrices:
$\nabla||\mathbf{Ax\text{-}b}||^2 = (\mathbf{M} + \mathbf{M}^T)x - 2\mathbf{A}^T\mathbf{b}$

Para determinar la solución se debe igual el gradiente hallado a $\vec{0}$:
$\nabla||\mathbf{Ax\text{-}b}||^2 = (\mathbf{M} + \mathbf{M}^T)x - 2\mathbf{A}^T\mathbf{b} = \vec{0}$
$\Leftrightarrow$ $\mathbf{x}^* = (\mathbf{M} + \mathbf{M}^T)^{-1}2\mathbf{A}^T\mathbf{b}$
$= [\mathbf{A}^T\mathbf{A} + (\mathbf{A}^T\mathbf{A})^T]^{-1}2\mathbf{A}^T\mathbf{b}$
$= [2\mathbf{A}^T\mathbf{A}]^{-1}2\mathbf{A}^T\mathbf{b}$
$\Rightarrow$ $\mathbf{x}^* = [\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}^T\mathbf{b}$

Parte b)

Resolviendo computacionalmente la expresión alcanzada en el apartado anterior para las matrices dadas como datos extra se obtiene que $\mathbf{x}^* = (-0{,}13285, 0{,}12793)^T$

Parte c, d y e)

Para todos los casos (paso fijo, paso decreciente, paso exacto y regla de armijo) se utilizó la misma configuración inicial y los mismos criterios de parada para que los resultados fueron comparables.

Como configuración inicial se seleccionó un punto “semilla” de forma aleatoria dentro de la región factible.

Posteriormente se calculó el funcional para ese valor inicial y el gradiente, además de inicializar todos los vectores donde se almacenarán los datos que después se graficarán y compararán.

La forma de cálculo del punto siguiente (x^{k+1}) es la habitual: $x^{k+1} = x^k + s^k d^k$, siendo: s^k y d^k el paso y la dirección correspondiente para la iteración k respectivamente.

Las condiciones de paradas utilizadas fueron 2:
a) La cantidad máxima de iteraciones no debe superar un valor preestablecido. Esto evita loops infinitos por errores o por divergencia.
b) La variación del funcional entre dos iteraciones consecutivas, normalizadas según el funcional del valor semilla, sea menor a una tolerancia preestablecida. Esto funciona correctamente dependiendo de la función de costo, sin embargo en este caso, como la función de costo es convexa, funciona correctamente. Matemáticamente esta condición se ve como: $||f(x^{k+1}) - f(x^k)||/||f(x^0)|| \leq tolerancia$

En la Figura 12 se puede apreciar el error relativo en función de la iteración para tres de los métodos solicitados (Paso fijo, Paso exacto y Regla de armijo), mientras que para ver el error relativo del método de paso decreciente hay que referirse a la Figura **??**. Esto se debe a que por momentos el error es tan grande que no permite comparar con los otros métodos en la misma gráfica puesto que el resto de los modelos quedan de esta escala.

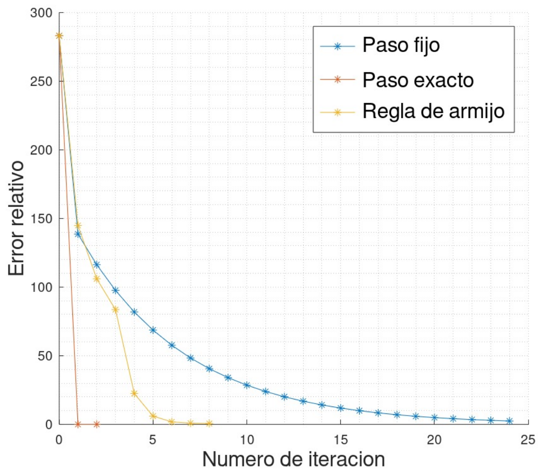

Figura 11: Error relativo en función de la iteración.

Como el punto de inicio afecta la performance de los métodos, se realizaron 3 corridas con puntos de inicio aleatorios, se calcularon los tiempos totales y se promediaron. Los tiempos totales consumidos en cada modelo se pueden ver en la siguiente tabla:

Método	Tiempo (ms)	Tiempo (ms)	Tiempo (ms)	Tiempo promedio (ms)
Paso fijo	0.078	0.069	0.105	0.084
Paso Decreciente	0.071	0.078	0.068	0.072
Paso exacto	0.086	0.094	0.083	0.087
Regla de Armijo	0.127	0.142	0.137	0.135

En las Figuras 13 14 y 15 se aprecian los tiempos consumidos en cada iteración por

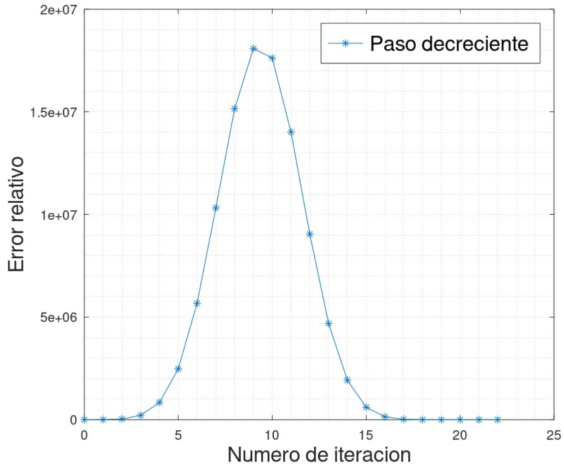

Figura 12: Error relativo en función de la iteración.

cada método, para cada uno de los 3 intentos ejecutados.

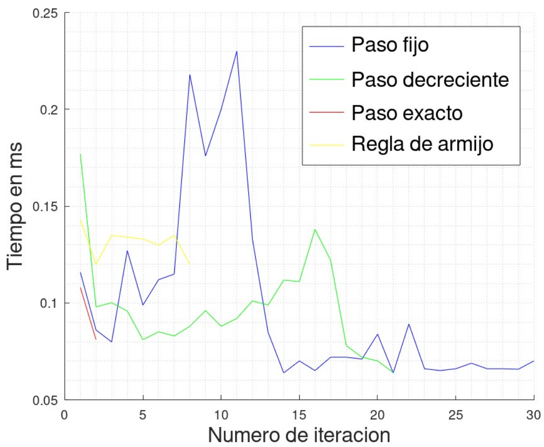

Figura 13: Tiempo consumido por cada método en cada iteración.

Comparando los tiempos consumidos entre los cuatro métodos se pueden hacer algunos comentarios.

El método con mejor performance en promedio es el que se obtiene por el paso decreciente, no obstante, en algunos casos, el método con paso fijo resuelve el problema más rápido. Lo que indica la fuerte dependencia con el punto de partida. El método por paso exacto, los sigue de cerca en performance aunque siempre le toma algo mas de tiempo en determinar la solución. Finalmente, el método por la regla de Armijo siempre toma algo más de tiempo en calcular la solución, haciendolo un método globalmente más lento en comparación.

Estos resultados también pueden observarse en las Figuras 13 14 y 15 donde en general, el método con la regla de Armijo tiene un consumo de tiempo algo mayor por iteración que el resto, mientras que el paso decreciente, tambien en términos generales es el que tiene menores tiempos de cómputo por iteración.

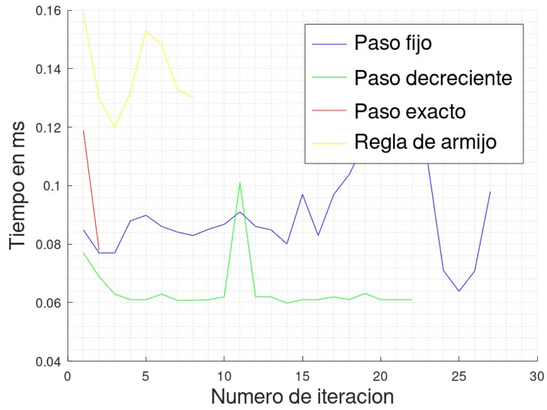

Figura 14: Tiempo consumido por cada método en cada iteración.

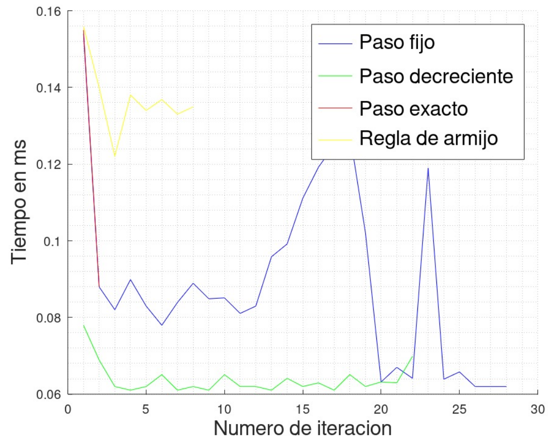

Figura 15: Tiempo consumido por cada método en cada iteración.

No obstante, cabe resaltar que al ser tiempos de ejecución tan cortos, los tiempos de computos asociados al almacenamiento de las variables es considerable, por lo que se deberían testear los métodos en circunstancias en las que esto no tenga tanto peso relativo para tener una comparación justa entre metodos.

Los códigos asociados a los ejercicios se adjuntan en formato .m a la entrega.

Ejercicio 5: Problemas equivalentes

Parte a)

Sea el problema:

$$(P0)\ min_{\mathbf{x}\in\mathbb{R}^n}\ \frac{1}{2}||\mathbf{Ax}\text{-}\mathbf{b}||^2 + \sum_i |x_i|$$

$P0$ será convexo sii, la función de costo es convexa y su conjunto de restricciones es convexo.

La función de costo definida por $f(\mathbf{x}) = \frac{1}{2}||\mathbf{Ax}\text{-}\mathbf{b}||^2 + \sum_i |x_i|$ será convexa si cumple con la desigualdad de la Ecuación 1.

Demostración

Usando la definición de $f(\mathbf{x})$ para $\mathbf{u},\ \mathbf{v}\ \in\ \mathbb{R}^n$ y $t\ \in\ [0,1]$ en el término de la izquierda de la Ec. 1
$f(t\mathbf{u} + (1-t)\mathbf{v}) = \frac{1}{2}||\mathbf{A}(t\mathbf{u} + (1-t)\mathbf{v})||^2 + \sum_i |tu_i + (1-t)v_i|$

Operando dentro de la norma, sacando escalares de factor común y separando la sumatoria:
$= \frac{1}{2}||t\mathbf{Au} + (1-t)\mathbf{Av} - (t\mathbf{b} + (1-t)\mathbf{b})||^2 + t\sum_i |u_i| + (1-t)\sum_i |v_i|$

Aplicando la desigualdad triangular:
$\leq \frac{1}{2}||t\mathbf{Au} - t\mathbf{b}||^2 + \frac{1}{2}||(1-t)\mathbf{Av} - (1-t)\mathbf{b})||^2 + t\sum_i |u_i| + (1-t)\sum_i |v_i|$

Sacando t y $(1-t)$ de factor comun:
$= t\left(\frac{1}{2}||\mathbf{Au}\text{-}\mathbf{b}||^2 + \sum_i |u_i|\right) + (1-t)\left(\frac{1}{2}||\mathbf{Av}\text{-}\mathbf{b}||^2 + \sum_i |v_i|\right)$

Sustituyendo se cumple la desigualdad buscada y se demuestra la convexidad de $f(\mathbf{x})$:
$= tf(\mathbf{u}) + (1-t)f(\mathbf{v})$ ■

Parte b)

Sea el problema:

$$(QP)\ min_{\mathbf{x},\ \mathbf{t}\in\mathbb{R}^n}\ \frac{1}{2}||\mathbf{Ax}\text{-}\mathbf{b}||^2 + \sum_i t_i$$

$$s.t. : x_i \leq t_i$$

$$-x_i \leq t_i$$

Se dice que (QP) y $(P0)$ son equivalentes sii: x^* es solución de $(P0) \Leftrightarrow x^*$ es solución de (QP)

Demostración $(\Rightarrow)$

Sea $\mathbf{x}^*$ minimiza $(P0)$

Se busca $t \in \mathbb{R}^n$: minimice: $\sum_i t_i : \{s.t. : x_i \leq t_i \;\; y \;\; - x_i \leq t_i\} = \begin{cases} \sum_i x_i^* & si \;\; x_i^* \leq 0 \\ \sum_i -x_i^* & si \;\; x_i^* \leq 0 \end{cases}$

$\Rightarrow \mathbf{x}^*$ también minimiza (QP) $\blacksquare$

Demostración $(\Leftarrow)$

Sea $(\mathbf{x}^*, \mathbf{t}^*)$ minimiza (QP)

Se asume por absurdo que: $\exists \hat{\mathbf{x}} : \dfrac{1}{2}||\mathbf{A}\hat{\mathbf{x}} - b||^2 + \sum_i |\hat{x}_i| < \dfrac{1}{2}||\mathbf{A}\mathbf{x}^* - b||^2 + \sum_i |x_i^*|$

Considerando que la restricción de t es equivalente a $|x_i| \leq t_i$ y que se busca minimizar, se puede sustituir $t_i = |x_i|$.
Sustituyendo entonces $|\hat{x}|$ por $\hat{t_i}$ y $|x_i^*|$ por $|t_i^*|$, se tiene:
$\dfrac{1}{2}||\mathbf{A}\hat{\mathbf{x}} - b||^2 + \sum_i |\hat{t}_i| < \dfrac{1}{2}||\mathbf{A}\mathbf{x}^* - b||^2 + \sum_i |t_i^*|$

Esto es un absurdo, pues el par que minimiza (QP) es $(\mathbf{x}^*, \mathbf{t}^*)$. $\blacksquare$

Parte c)

Para demostrar que el problema (QP) es convexo, se utilizará la desigualdad de la Ec. 1

Demostración

Usando la definición de $f(\mathbf{x},\mathbf{t}) = \dfrac{1}{2}||\mathbf{Ax}\text{-}\mathbf{b}||^2 + \sum_i |t_i|$ como una suma de dos funciones $g(\mathbf{x}) = ||\mathbf{Ax}\text{-}\mathbf{b}||^2$ y $h(\mathbf{t}) = \sum_i |t_i|$
$\Rightarrow f(\mathbf{x},\mathbf{t})$ es la suma con coeficientes positivos $(1/2y1)$ de las funciones $g(\mathbf{x})$ y $h(\mathbf{t})$ que si estas son convexas $\Rightarrow f(\mathbf{x},\mathbf{t})$ es convexa.

Se analiza $h(\mathbf{t})$
Se define $\mathbb{I}$ como la matriz identidad de orden $n \Rightarrow h(\mathbf{t}) = \sum_i |t_i| = \mathbb{I}\mathbf{t} \Rightarrow h(\mathbf{t})$ se puede considerar como un mapeo lineal $\Rightarrow$ es convexa

Se analiza $g(\mathbf{x})$
Se define $g^*(\mathbf{x}) = \mathbf{Ax}\text{-}\mathbf{b}$ como un mapeador lineal de $\mathbb{R}^n \rightarrow \mathbb{R} \Rightarrow g(\mathbf{x}) = ||g^*(\mathbf{x})||^2$
$\Rightarrow$ utilizando el resultado demostrado en la parte c) del Ejercicio 1, solo queda demostrar que la $||\mathbf{x}||^2$ en convexa, $\Rightarrow g(\mathbf{x})$ será convexa, y $\Rightarrow f(\mathbf{x},\mathbf{t})$ será convexa.

En la Figura 16 se muestran dos problemas cuyo dominio es el mismo, y la solución se da en el mismo punto x^* por lo tanto, son dos problemas equivalentes. No obstante, la función de la izquierda es convexa, mientras la de la derecha **no** lo es. Esto demuestra que pueden existir dos problemas equivalentes de minimización uno convexo, y el otro no.

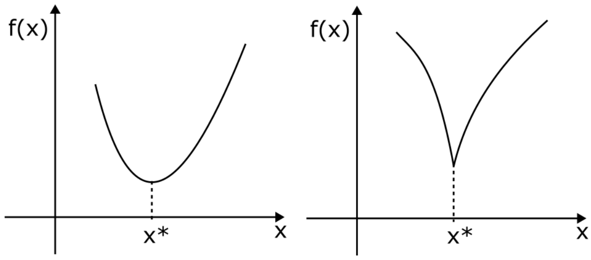

Figura 16: Comparación problemas equivalentes.

Parte d)

Resolviendo el problema (QP) con el solver *cvx* en matlab se obtuvo:
Valor óptimo: $f(x^*) = 6{,}9338$
$x^* = (-0{,}1327, 0{,}1273)^T$
$t^* = (0{,}1327, 0{,}1273)^T$

Estos resultados permiten relacionar a x^* con t^*, mostrando que $\mathbf{t} = |\mathbf{x}|$

Ejercicio 6: Mínimos cuadrados con restricción

Parte a)

Sea el problema:

$$(P0)\ min_{x,y}\{5x^2 + 5y^2 + 5x - 3y - 6xy + 5/4\}$$
$$s.t. : x^2 + y^2 \leq R^2$$

El problema $(P0)$ será convexo $\Rightarrow f(x,y) = 5x^2 + 5y^2 + 5x - 3y - 6xy + 5/4$ es convexa y $C : x^2 + y^2 \leq R^2$ es convexo.
Se analiza primero $f(x,y)$ utilizando que si su matriz Hessiana es definida positiva $\Rightarrow$ $f(x,y)$ es convexa.

$$\nabla^2 f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix}$$

El determinante de la matriz $\nabla^2 f = 64 \Rightarrow f(x,y)$ es convexa.

Se analiza ahora el conjunto C de restricciones. $C = B(c = 0, R) \Rightarrow$ por el resultado de la parte e) del Ej. 1, C es convexo.
$\Rightarrow (P0)$ es un problema convexo.

Parte b)

Dado que es un problema convexo, si se calcula el gradiente e iguala al vector nulo su puede determinar el punto donde se minimiza la función. Este punto se da en $x^* = (-1/2, 0)$. Este punto, se encuentra a una distancia $R = 1/2$ del origen de coordenadas que coincide con el origen del disco factible de soluciones. Esto implica que para valores de radio de región factible $R < 1/2$ el punto donde se minimiza la función de costo queda fuera de esta región, por lo que el mínimo del problema con ese tipo de restricciones quedará sobre la frontera de la región factible, y por lo tanto estará activada la restricción.

Parte c)

Solamente se alcanzó a implementar paso decreciente, no por la complejidad del "line search" sino por escacez de tiempo para trabajar en el obligatorio.
A continuación se muestran los resultados obtenidos para este algoritmo. La Figura 17 muestra la sucesión de puntos determinados por los que evoluciona el algoritmo el algoritmo implementado, se aprecia como en el primer paso se estampa contra la frontera de la región factible y luego evoluciona sobre la misma hasta que se cumple alguno de los criterios de parada.
En la Figura 18 se ve la evolución de la diferencia de los valores funcionales de dos puntos de la sucesión consecutivos. Se aprecia que desciende muy rápidamente en el primer paso

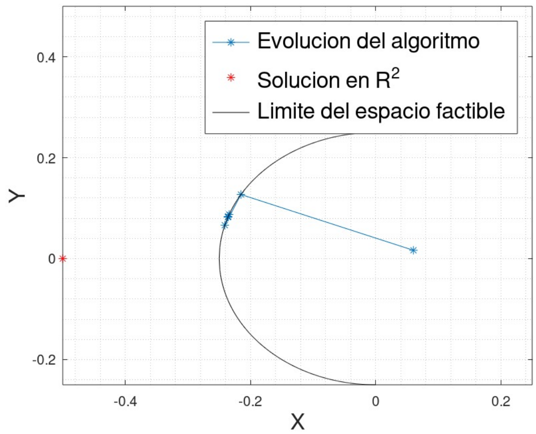

Figura 17: Sucesión de puntos obtenidos en el espacio factible y mínimo.

y posteriormente la evolución es mucho más lenta. Cabe notar que esto es consecuencia de lo poco compleja que es la función de costo y el dominio.

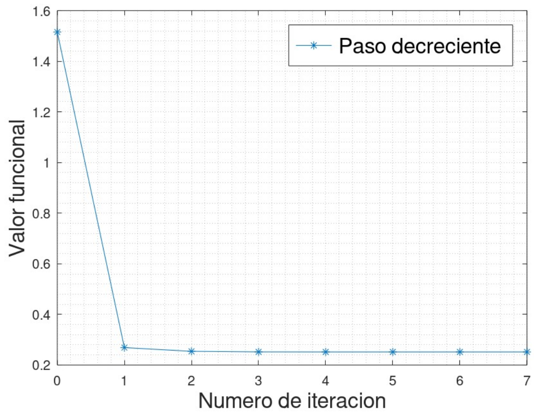

Figura 18: Evolución de la diferencia entre valores funcionales consecutivos.

Por otro lado, se muestra la evolución de la diferencia entre dos puntos consecutivos de la sucesión. Aquí se aprecia un comportamiento similar al de la evolución de los valores funcionales. Nuevamente, esto está fuertemente relacionado con lo amigable que es la función de costo y el conjunto factible.

Parte d)

El nuevo problema propuesto **NO** es un problema convexo, puesto que a pesar de que la función de costo es convexa, **NO** lo es el conjunto factible.

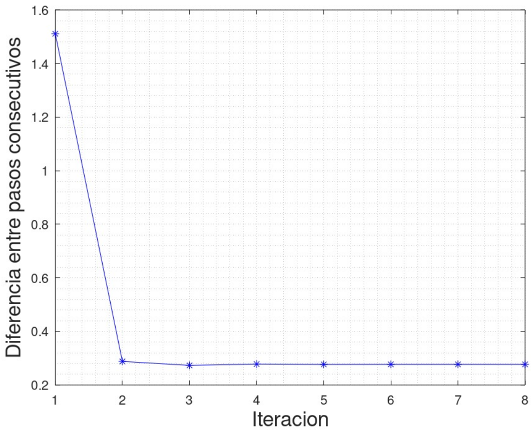

Figura 19: Evolucion de la diferencia de los valores x^k.