Ejercicio 1

a) $g(x) = \sum_i w_i f_i(x)$, $w_i \geq 0$

Sabemos que f_i es convexa $\forall i$

$$\Rightarrow f_i(tx+(1-t)y) \leq t f_i(x) + (1-t) f_i(y)$$

$\forall x, y \in \mathbb{R}^n$
$\forall t \in [0,1]$

Queremos probar que g es convexa.
Es decir, que se cumple

$$\underbrace{g(tx+(1-t)y)}_{} \overset{?}{\leq} t g(x) + (1-t) g(y)$$

trabajamos solo sobre el termino izquierdo

convexidad de f_i

$$\sum_i w_i f_i(tx+(1-t)y) \underset{\circledast}{\leq} \sum_i w_i [t f_i(x) + (1-t) f_i(y)]$$

$$= t \underbrace{\sum_i w_i f_i(x)}_{= g(x)} + (1-t) \underbrace{\sum_i w_i f_i(y)}_{= g(y)}$$

Por lo tanto nos queda

$$g(tx+(1-t)y) \leq t g(x) + (1-t) g(y)$$

$\Rightarrow$ g convexa

$\circledast$ Ademas de convexidad, hay que pedir que $w_i > 0$, lo que esta dado por hipotesis.

b) $h(x) = \max\{f_i(x) : i = 1, \ldots, k\}$

$h(tx + (1-t)y) = \max\{f_i(tx + (1-t)y)\}$

Supongamos que j es el indice de la función que es máxima en $tx + (1-t)y$

$\Rightarrow h(tx + (1-t)y) = f_j(tx + (1-t)y)$

Como f_j es convexa

$f_j(tx + (1-t)y) \leq t f_j(x) + (1-t) f_j(y)$

$\Rightarrow h(tx + (1-t)) \leq t f_j(x) + (1-t) f_j(y)$

Por otro lado sabemos que

$\left.\begin{array}{l} f_j(x) \leq h(x) \\ f_j(y) \leq h(y) \end{array}\right\} \Rightarrow t f_j(x) + (1-t) f_j(y) \leq t h(x) + (1-t) h(y)$

$\Rightarrow h(tx + (1-t)) \leq t h(x) + (1-t) h(y)$

$\Rightarrow h$ convexa

Escaneado con CamScanner

c) $\ell(x) = f_1(Ax+b)$

$\ell(tx+(1-t)y) = f_1(A(tx+(1-t)y))$

Operamos en la expresión anterior:

sumo y resto tb

$\ell(tx+(1-t)y) = f_1(tAx+(1-t)Ay+tb+(1-t)b)$

$= f_1(t(Ax+b)+(1-t)(Ay+b))$

f_1 convexa

$\leq t\underbrace{f_1(Ax+b)}_{\ell(x)}+(1-t)\underbrace{f(Ay+b)}_{\ell(y)}$

$= t\,\ell(x) + (1-t)\,\ell(y)$

$\Rightarrow \ell(tx+(1-t)y) \leq t\,\ell(y)+(1-t)\,\ell(y)$

$\Rightarrow \ell$ convexa

d) $Y = \bigcap_i X_i$

Tomemos dos puntos $z, w \in Y$

$\Rightarrow z, w \in X_i \ \forall i$

Como X_i es convexo $\forall i$

$\Rightarrow q_t = tz + (1-t)w \in X_i \ \forall i$

Por def de interseccòn $q_t \in Y$

$\Rightarrow q_t = tz + (1-t)w \in Y$

Como z, w son cualquier punto de Y y verifican la def de convexidad

$\Rightarrow Y$ es convexo

e) $B(c,r) = \{x : \|c-x\| \leq r\}$

Sea $z, w \in B(c,r)$. Es decir:

$$\begin{cases} \|c-z\| \leq r \\ \|c-w\| \leq r \end{cases}$$

Veamos si $p = tz + (1-t)w \in B(c,r)$

Es decir si

$$\|c-p\| \leq r$$

$$\|c-p\| = \|c - (tz + (1-t)w)\|$$

sumo y resto tc

$$= \|tc + (1-t)c - tz - (1-t)w\|$$

$$= \|t(c-z) + (1-t)(c-w)\|$$

desigualdad triangular

$$\leq \|t(c-z)\| + \|(1-t)(c-w)\|$$

$$= t\underbrace{\|c-z\|}_{\leq r} + (1-t)\underbrace{\|c-w\|}_{\leq r}$$

$$\leq tr + (1-t)r = r$$

$$\Rightarrow \|c - [tz + (1-t)w]\| \leq r$$

$$\Rightarrow tz + (1-t)w \in B(c,r) \quad \forall z, w \in B(c,r) \quad \forall t \in [0,1]$$

$\Rightarrow B(c,r)$ es convexa

f) $C = \{ \sum_i \theta_i x_i : \sum_i \theta_i , \theta_i \geq 0 \ \forall i \}$

Sea $z, w \in C$. Entonces se cumple:

$$\begin{cases} \exists \theta_z / z = \sum_i \theta_{zi} x_i \ \ y \ \ \sum_i \theta_{zi} = 1 \\ \exists \theta_w / w = \sum_i \theta_{wi} x_i \ \ y \ \ \sum_i \theta_{wi} = 1 \end{cases}$$

Veamos si $p = tw + (1-t) z \in C$ $\quad \forall w, z \in C$
$\forall t \in [0,1]$

Sustituimos w, z por su valor

$$p = t \sum_i \theta_{zi} x_i + (1-t) \sum_i \theta_{wi} x_i$$

$$p = \sum_i \underbrace{[t \theta_{zi} + (1-t)\theta_{wi}]}_{\theta_{pi}} x_i = \sum_i \theta_{pi} x_i$$

Si probamos que $\begin{cases} \theta_{pi} \geq 0 \ \forall i \\ \sum \theta_{pi} = 1 \end{cases} \Rightarrow C$ es convexo

Veamos

* $\theta_{pi} \geq 0$: θ_{pi} es la suma de elementos positivos $\Rightarrow \theta_{pi} \geq 0$

* $\sum_i \theta_{pi} = \sum_i t\theta_{zi} + (1-t)\theta_{wi} = t \overbrace{\sum_i \theta_{zi}}^{=1} + (1-t) \overbrace{\sum_i \theta_{wi}}^{=1}$

$= t + (1-t) = 1$

$\Rightarrow C$ es convexo

Ejercicio 2

a) Veamos si la función de costo f es convexa.

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

sugerencia

$$\stackrel{?}{=} -\log(x-y) - \log(x+y)$$

Notar que tanto $x-y$ como $x+y$ son un caso particuar de la expresion vectorial $Ax+b$.

* $x-y$: $A_1 = (1,-1)$ $b_1 = 0$
* $x+y$: $A_2 = (1,1)$ $b_2 = 0$

Sea $g(x) = -\log(x)$ con $g: \mathbb{R} \to \mathbb{R}$

$$\Rightarrow f(x,y) = g(x-y) + g(x+y)$$

$$\Rightarrow f(\vec{x}) = g(A_1\vec{x}) + g(A_2\vec{x})$$

Si probamos que $g(x)$ es convexa $\Rightarrow f$ sera convexa. Esto es aplicando las propiedades a y c de Ejercicio 1.

$$g(x) = -\log(x) \qquad g'(x) = -\frac{1}{x} \qquad g''(x) = \frac{1}{x^2}$$

Ver que $g''(x) > 0 \ \forall x \in \mathbb{R} \Rightarrow g$ es convexa

Como g convexa $\Rightarrow f$ convexa

b) $X = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq 1, x \geq 0, 2x-y \leq 0$
$y \geq 1/2\}$

Intentaremos ver que X es la interseccion de conjuntos convexos.

* $x^2+y^2 \leq 1$ es el conjunto $B(0,1)$ que vimos en el ejercicio 1 que es convexo

* $\left.\begin{array}{l} x \geq 0 \\ 2x-y \leq 0 \\ y \geq 1/2 \end{array}\right\}$ Son todos semiplanos y los semiplanos son convexos

Entonces X es la intersección de 4 conjuntos convexos. Por la parte d de ejercicio 1
$\Rightarrow$ X es convexo

Ejercicio 2

Parte c

A continuación se presentan las curvas de nivel de $f(x, y) = log(y^2 - x^2)$

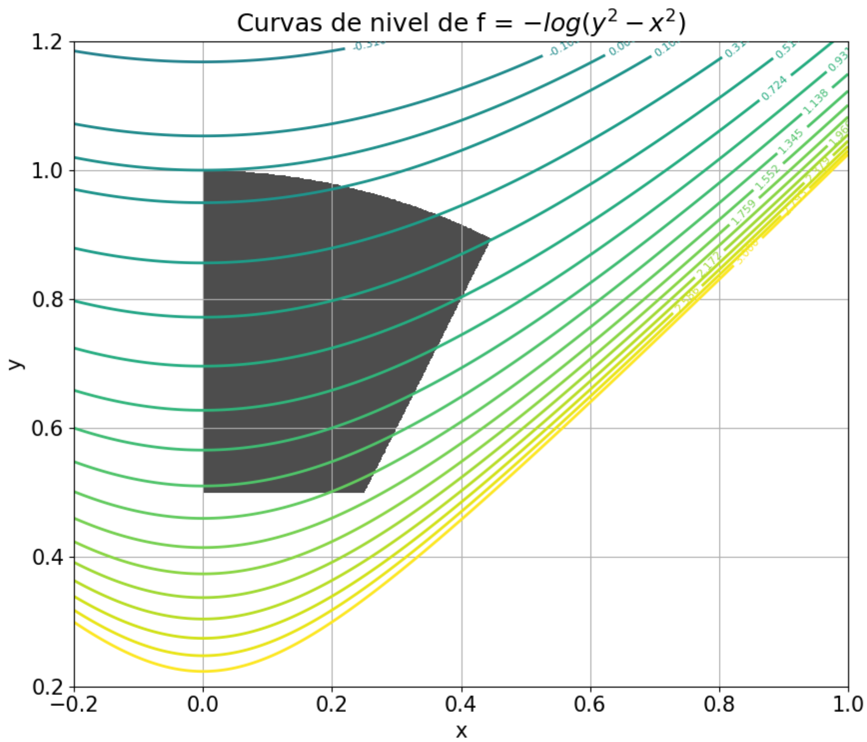


Figure 4: Cuanto mas amarillento el valor de la curva de nivel, mayor el valor asociado. La región gris corresponde a la zona factible.

Parte d

En la Figura 4, se ve que la curva de nivel correspondiente al nivel $f(x, y) = 0$ intersecta al conjunto factible únicamente en el punto $\mathbf{p} = (0, 1)$. Notar que las demás curvas de nivel que intersectan al conjunto poseen un valor mayor, por lo que el mínimo se da en $\mathbf{p}$.

Ejercicio 3

Parte a

$$\min_{x \in R} f(x) = 4x^4 - x^3 - 4x^2 + 1$$

$$s.to : -1 \leq x \leq 1$$

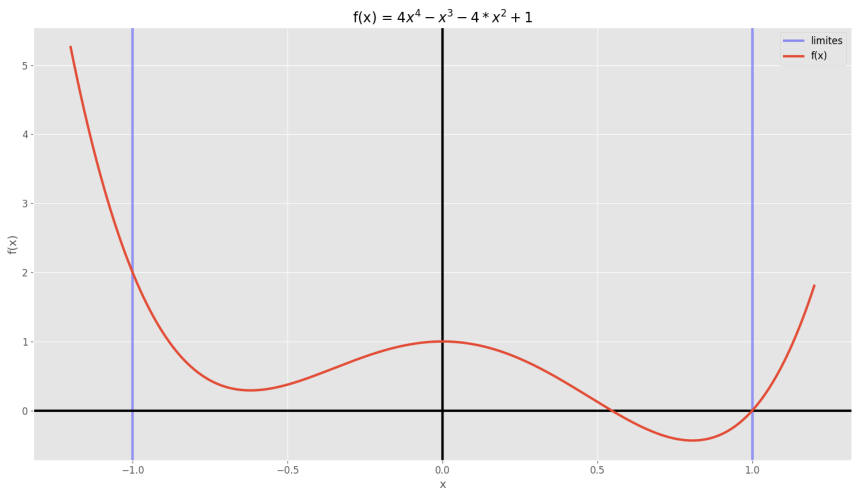


Figure 5

Derivando $f(x)$ se tiene $f'(x) = 16x^3 - 3x^2 - 8x$, la cual se anula en

$$\bar{x} = \begin{cases} -0.62 \\ 0.81 \\ 0 \end{cases}$$

A partir de la gráfica se realiza la siguiente clasificación para estos puntos

$$\bar{x} = \begin{cases} -0.62 & \text{mínimo local} \\ 0.81 & \text{mínimo global} \\ 0 & \text{máximo local} \end{cases}$$

Los tres puntos están dentro del conjunto y el mínimo se da en uno de ellos, en particular en 0.81.

Parte b

$$\min_{x \in R} f(x) = x^3$$

$$s.to : -1 \leq x \leq 1$$

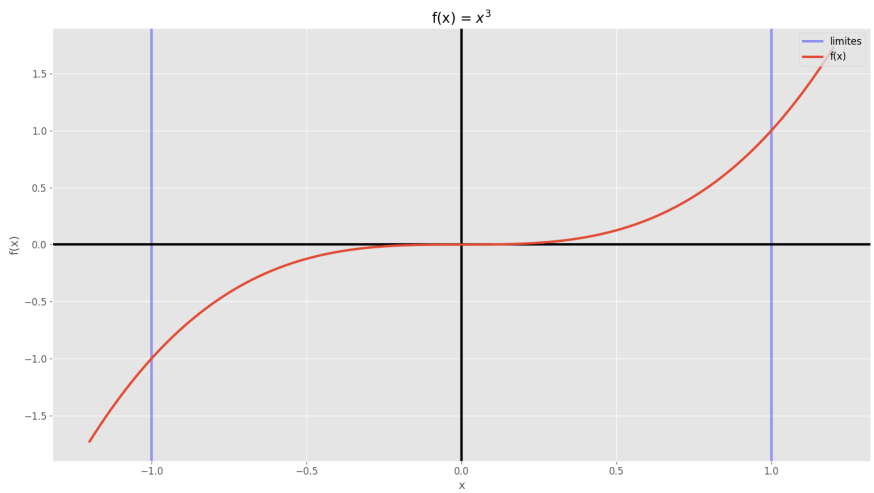


Figure 6

El gradiente de f es $f'(x) = 3x^2$.
Este se anula únicamente en $\bar{x} = 0$, el cual como se ve en la gráfica es un punto silla dentro de la región factible.

El mínimo se da en el extremo -1 del conjunto.

Parte c

$$\min_{x \in R} f(x) = (x - a)^2 + 1$$

$$s.to : -1 \leq x \leq 1$$

Se gráfica la función $f(x)$ para varios valores de a.

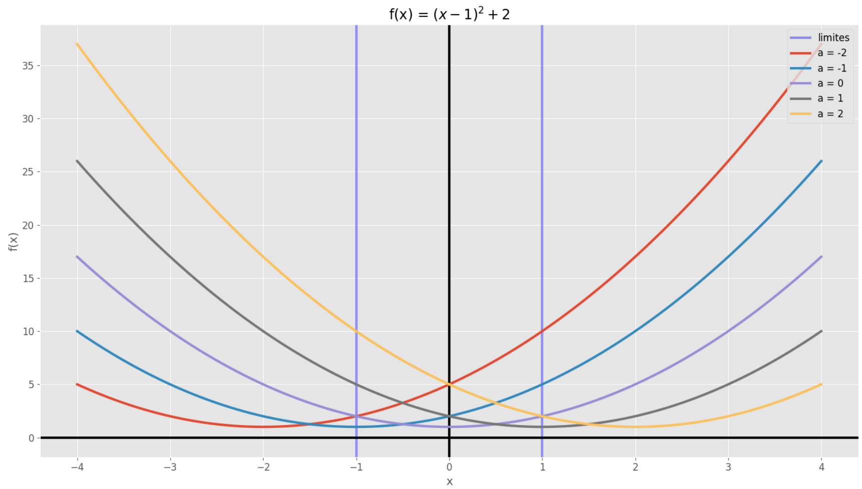


Figure 7

La derivada de la función es $f'(x) = 2(x-a)$, la cual se anula en $\bar{x} = a$. El punto $\bar{x}$ es un mínimo local, observar que el valor de a determina donde se da el mínimo de la parábola f.

Si $|a| > 1$ el mínimo local estará fuera del conjunto factible, por lo tanto, el problema con restricciones tendrá su mínimo en uno de sus extremos.

Existen 3 casos posibles

$$\bar{x} = \begin{cases} -1 & \text{si } a < 1 \\ a & \text{si } |a| \leq 1 \\ 1 & \text{si } a > 1 \end{cases}$$

Parte d

$$\min_{(x)\in R^2} f(x) = \|x - x^*\|^2 + 1$$

$$\text{s. to: } x \in [0,1]^2$$

Dado que $\|.\|^2$ es convexa (se prueba en el Ejercicio 5) y por la propiedad del Ejercicio 1 parte c, se tiene que $\|x - \bar{x}\|^2$ es convexo. Sumarle 1 no altera la convexidad, ya a suma de funciones convexas es convexa, por lo tanto $f(x)$ lo es.

Veamos si $B = [0,1]^2 = S_1 \cap S_2$ es convexo. Basta probar que S_1 y S_2 sean convexos, ya que la intersección de ambos también lo seria. Ambos conjuntos son semiplanos de R^2, por lo que son convexos.

El gradiente de f es $\nabla f(x) = 2(x - x^*)$, anulándose en $\bar{x} = x^*$. Este es un mínimo local, que se encuentra fuera de la región factible.

A continuación se muestran las curvas de nivel junto con la región factible.

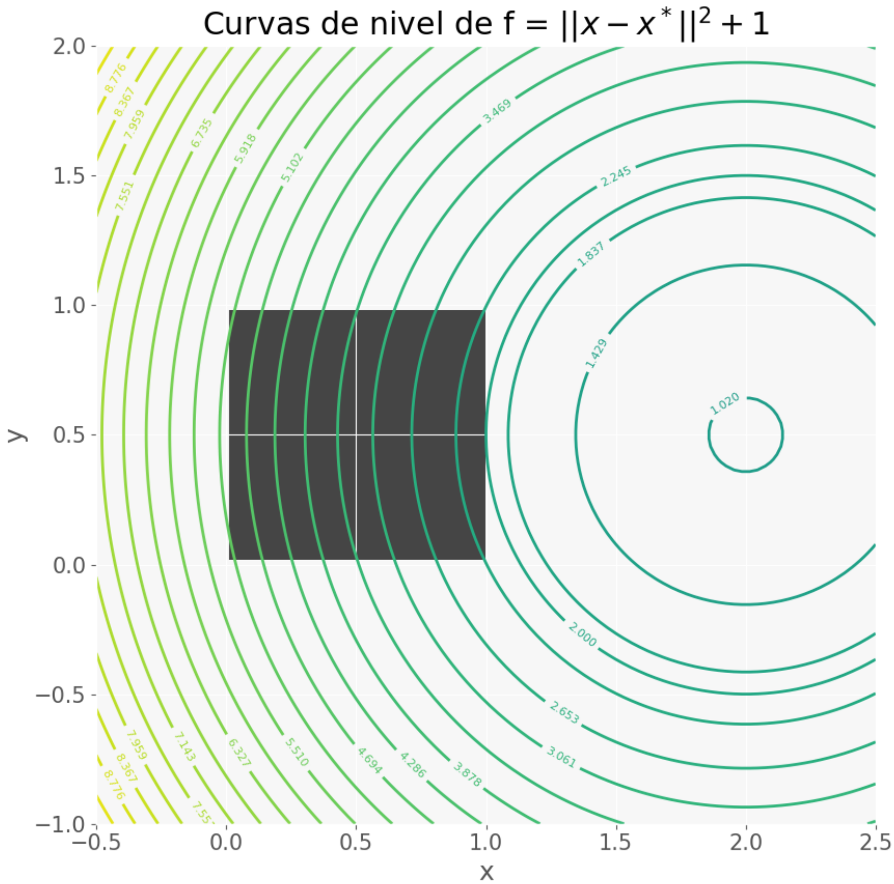


Figure 8: Cuanto mas amarillento el valor de la curva de nivel, mayor el valor asociado. La región gris corresponde a la zona factible.

Gráficamente se puede apreciar que la curva de nivel de menor valor que intersecta a la región factible es aquella asociada al valor 2. Este corte se da en el punto $\bar{x} = (1, \frac{1}{2})$, el cual es el mínimo del problema

Ejercicio 4

Parte a

Dado el problema sin restricciones siguiente:

$$\min_{\mathbf{x}} f(\mathbf{x})$$

Se tiene que la condición necesaria de optimalidad es

$$\nabla f = \mathbf{0}$$

Si además f es convexa, se tiene que es una condición suficiente. Es decir, si se anula el gradiente en un punto $\bar{x}$ ($\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$), entonces este punto es un mínimo global de f.

Veamos que el problema de mínimos cuadrados, $\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2$, es convexo

Tenemos que la función $g(\mathbf{x}) = \|\mathbf{x}\|_2^2$ es convexa (esta prueba se realiza en el Ejercicio 5). Por otro lado, como se probo en el ejercicio 1 parte c, $f(\mathbf{x}) = g(\mathbf{A}\mathbf{x}-\mathbf{b})$ es convexa si g es convexa. Entonces el problema de mínimos cuadrados sin restricciones es convexo.

Ahora calculemos el gradiente de $f(\mathbf{x})$. Sea $g(\mathbf{x}) = \mathbf{A}\mathbf{x}-\mathbf{b}$, entonces $f(\mathbf{x}) = f(g(\mathbf{x})) = \|g\|_2^2$. Por la regla de la cadena

$$\nabla f(\mathbf{x}) = \frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} \frac{\partial f(g)}{\partial g}$$

Recurriendo a "The Matrix Cookbook" se tiene que

$$\frac{\partial f(g)}{\partial g} = \frac{\partial \|g\|_2^2}{\partial g} = 2g$$

$$\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial \mathbf{A}\mathbf{x}+\mathbf{b}}{\partial \mathbf{x}} = \mathbf{A}^T$$

Entonces el gradiente sera

$$\nabla f(\mathbf{x}) = \mathbf{A}^T 2g$$

$$\nabla f(\mathbf{x}) = 2\mathbf{A}^T(\mathbf{A}\mathbf{x}-\mathbf{b}) \tag{1}$$

Busquemos el punto $\bar{\mathbf{x}}$, para esto igualamos la ecuación anterior a 0 y despejamos el valor $\bar{\mathbf{x}}$.

$$\nabla f(\bar{\mathbf{x}}) = \mathbf{0} \Leftrightarrow \mathbf{A}^T(\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}) = \mathbf{0} \Leftrightarrow \mathbf{A}^T\mathbf{A}\bar{\mathbf{x}} - \mathbf{A}^T\mathbf{b} = \mathbf{0} \Leftrightarrow \mathbf{A}^T\mathbf{A}\bar{\mathbf{x}} = \mathbf{A}^T\mathbf{b} \Leftrightarrow$$

$$\bar{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \tag{2}$$

En el caso del problema dado, el punto optimo es

$$\bar{\mathbf{x}} = (-0.13285, 0.12793)^T$$

Implementación

En el archivo ”ej4.py” adjunto en la entrega contiene el código de las 4 implementaciones. Para los 4 tipos de pasos se graficar evolucion del error $\frac{\|\bar{\mathbf{x}}-\mathbf{x_k}\|}{\|\bar{\mathbf{x}}\|}$.

De forma arbitraria se escoge comenzar desde el punto $\mathbf{x_0} = (0, 0)$, ya que esta elección inicial no modifica cualitativamente el resultado (se probó con otros puntos).

En todos los métodos, el criterio de parada elegido es la combinación de:

$$k < MAX\_ITER = 1000$$

$$\frac{\|\mathbf{x_k} - \mathbf{x_{k-1}}\|}{\|\mathbf{x_{k-1}}\|} < \epsilon = 0.0001$$

Se fijo la matriz $\mathbf{D^k}$ de todos estos métodos como la identidad ($\mathbf{D^k} = \mathbf{Id}$), es decir, no se uso ni Newton ni Diagonal Scaling.

En la Figura 1 se presentan las gráficas de evolución del error para cada uno de los tipos de paso, indicándose, en los casos necesarios, los parámetros utilizados.

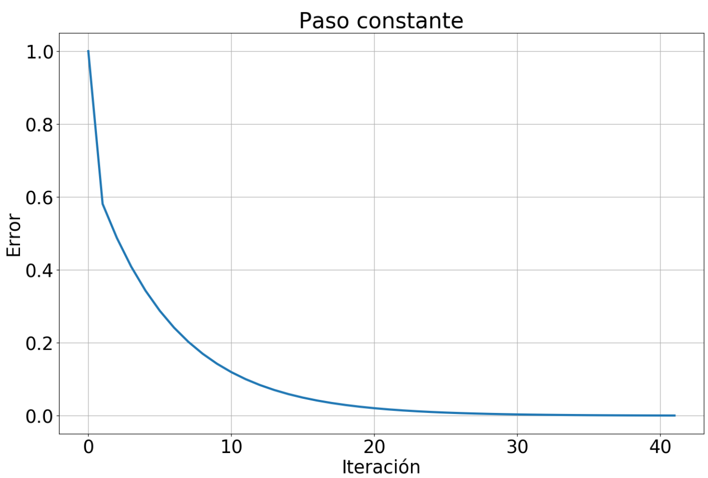
A continuación se muestra en una tabla la cantidad de iteraciones de cada método, el punto obtenido y el error final.

	Cantidad de Iteraciones	Punto obtenido	Error final
Paso Constante	41	(-0.13285,0.12783)	0.00052
Paso Decreciente	23	(-0.13285,0.12793)	0.00002
Line Search	6	(-0.13285,0.12793)	0.00000
Armijo	25	(-0.13285,0.12791)	0.00009

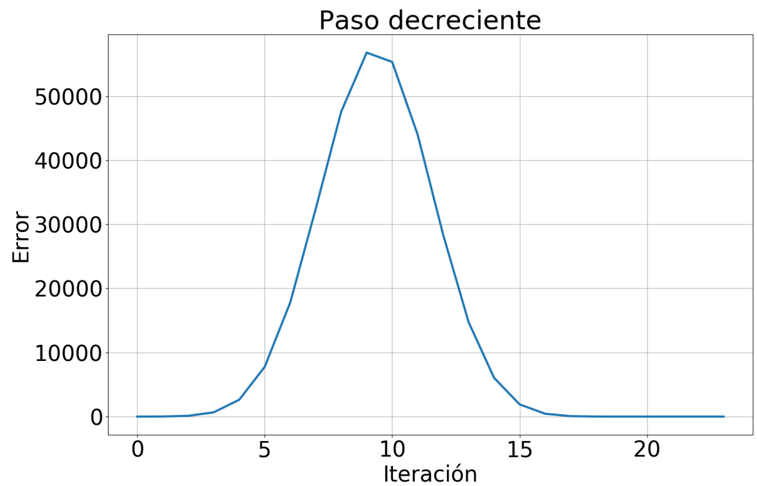
A partir de los resultados vemos que todos los métodos convergen al óptimo. El error en todos los casos es muy pequeño, destacándose que en Line Search es el menor debido a su búsqueda exhaustiva.

La gráfica del ”Paso Decreciente” tiene una forma que puede llamar la atención a primera vista. El error elevado se debe que para las primeras iteraciones los valores de los pasos son grandes, lo que lleva a que si bien nos movemos en una dirección de descenso, x_{k+1} este muy lejos de x_k y por tanto es posible moverse a un punto ”mucho” peor, esto se debe a que no es un método de descenso. Lo anterior se puede ver con más detalle en la Figura 3 adjunta al final de la sección, en la cual se observa la trayectoria.

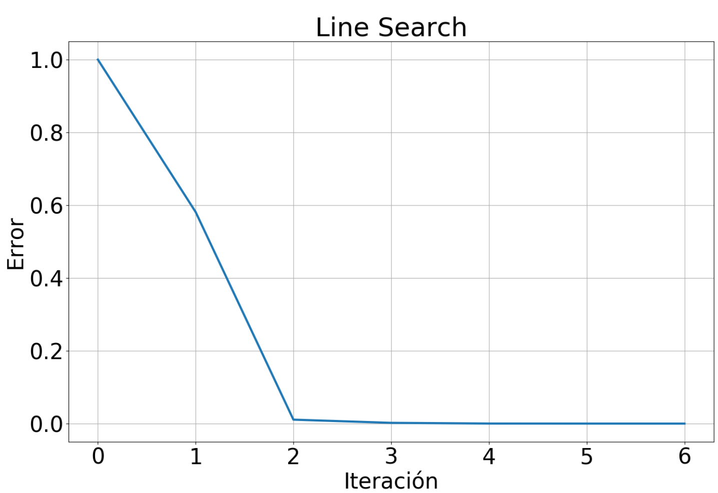
La velocidad de convergencia del ”Line Search” es la más alta, esto es porque que busca en una grilla el mejor punto, por lo que se garantiza que paso a paso se va hacia un ”muy buen” punto. Esta velocidad en iteraciones no significa menor costo ya que se debe evaluar la función una gran cantidad de veces por



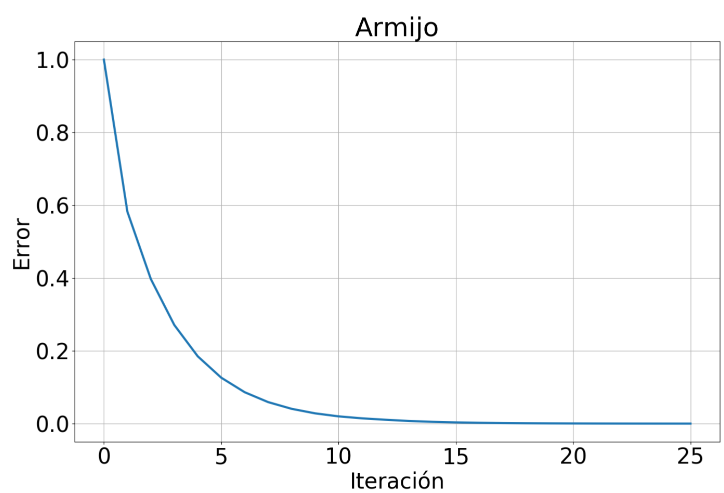
(a) $\alpha = \frac{1}{2\|\mathbf{A}\|_2^2}$



(b) $\alpha = \frac{0{,}001}{k}$. Nota: No se usa $\alpha = \frac{0{,}01}{k}$ debido a que el método no converge en las 1000 iteraciones propuestas.



(c) Grilla de 100 puntos con un largo de $0,001x$.



(d) $\sigma = 0,1$, $\beta = 0,5$, $\alpha = 0,1$

Figure 1: Evolución del error: $\frac{\|\bar{\mathbf{x}} - \mathbf{x_k}\|}{\|\bar{\mathbf{x}}\|}$

iteración. También, era de esperar que el paso constante, debido a la sencillez de este método, fuera quien tuviera una mayor cantidad de iteraciones.

En cuanto a la velocidad de convergencia al usar Armijo, se puede decir que los hiperparametros elegidos no son los mejores ya que en la teoría es un método que busca aproximarse al "Line Search" sin la necesidad de evaluar muchos puntos. En este aspecto, para mejorar el resultado, se realiza una búsqueda en una grilla de hiperparametros, obteniéndose para $\sigma = 0.17$, $\beta = 0.55$, $\alpha = 0.02$ una cantidad de 7 iteraciones y un error de 0.00002, lo cual confirma que existe un conjunto de parámetros para el cual el método es mucho más veloz. En la figura 2 se ve la gráfica del error correspondiente a estos parámetros.

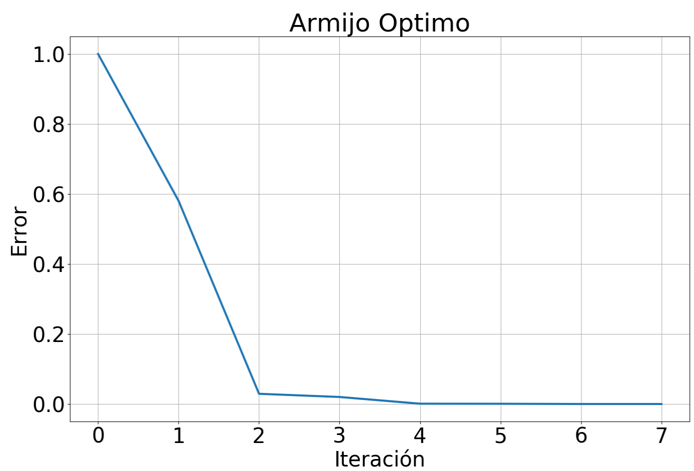
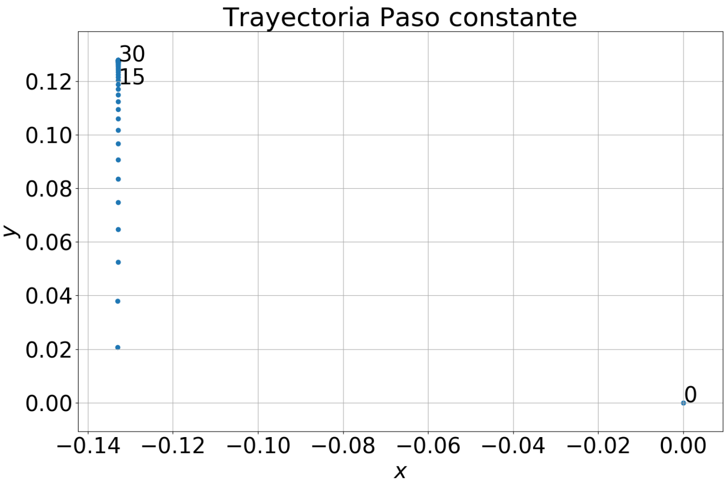


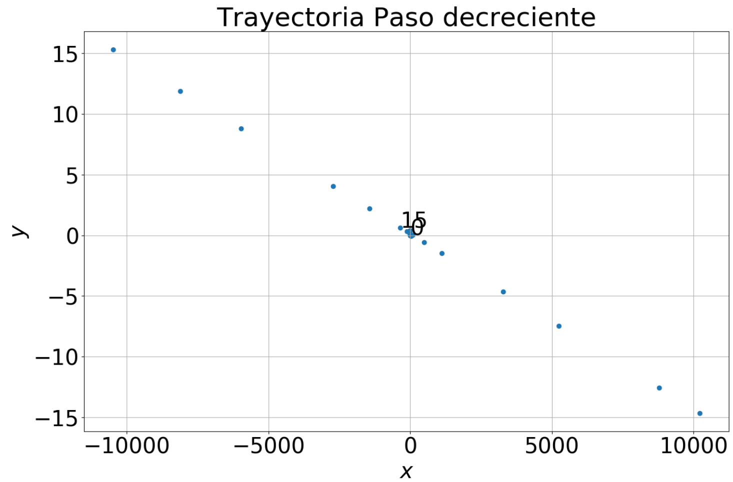
Figure 2: $\sigma = 0.17$, $\beta = 0.55$, $\alpha = 0.02$

El grid search realizado bien podría realizarse con los otros 3 métodos, pero debido a que esto esta fuera de alcance no se realiza en este caso.

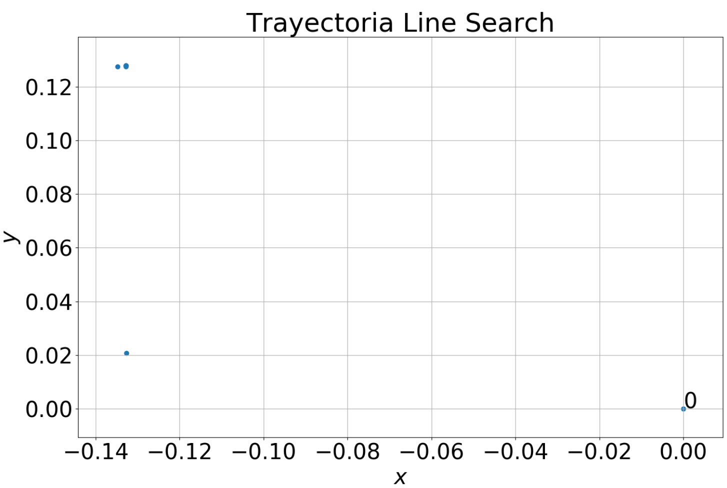
En la Figura 3 se puede observar las gráficas de trayectorias de los métodos.



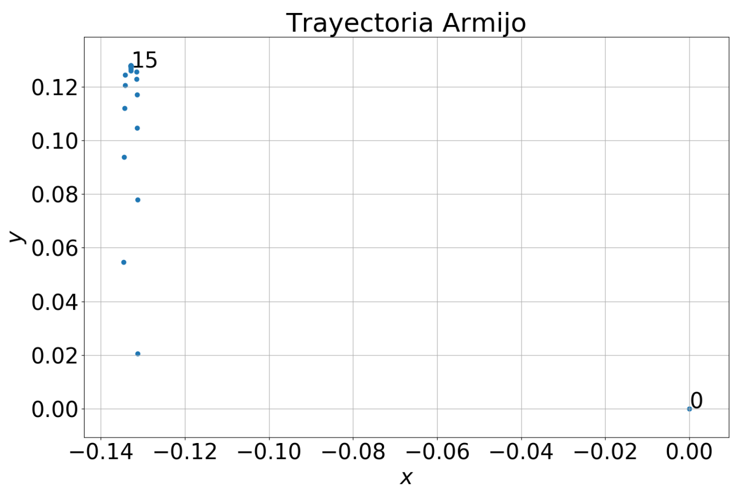
(a) $\alpha = \frac{1}{2\|\mathbf{A}\|_2^2}$



(b) $\alpha = \frac{0{,}001}{k}$.



(c) Grilla de 100 puntos con un largo de $0,001$.



(d) $\sigma = 0,1$, $\beta = 0,5$, $\alpha = 0,1$

Figure 3: Trayectoria de $\mathbf{x_k}$

Ejercicio 5

$$(PO) \min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax-b\|^2 + \underbrace{\sum_{1}^{n}|x_i|}_{\|x\|_1} = f(x)$$

a) * Para probar que $f(x)$ es convexa primero veamos que todas las normas son convexas, es decir, $g(x)=\|x\|$ es convexa.

Sea $x,y\in\mathbb{R}^n$ y $t\in[0,1]$

$$\Rightarrow \|tx+(1-t)y\| \underset{\substack{\uparrow \\ \text{desigualdad triangular}}}{\leq} \|tx\|+\|(1-t)y\| = t\|x\|+(1-t)\|y\|$$

$$\Rightarrow g(tx+(1-t)y)\leq t\,g(x)+(1-t)\,g(y)$$

$\Rightarrow$ Las normas son convexas

* Luego veamos que si $h(x)$ es convexa y no negativa $\Rightarrow h^2(x)$ es convexa

Sabemos que:

$$h(tx+(1-t)y)\leq t\,h(x)+(1-t)\,h(y)$$

Elevamos al cuadrado a ambos lados y como $h(x)\geq 0\ \forall x \Rightarrow$ se mantiene la desigualdad

$$h^2(tx+(1-t)y)\leq \underbrace{t^2h^2(x)+(1-t)^2h^2(y)+2t(1-t)h(x)h(y)}_{\circledast}$$

Trabajemos sobre $\circledast$, sumamos y restamo terminos

$$\circledast = t^2h^2(x)+(1-t)^2h^2(y)+2t(1-t)h(x)h(y)-t\,h^2(x)-(1-t)h(y)$$
$$+t\,h^2(x)+(1-t)\,h(y)$$

Agrupando los 5 primeros terminos:

$$\circledast = \underbrace{-t(1-t)\left[h^2(x)+h^2(y)+2h(x)h(y)\right]}_{\leq 0} + t h^2(x) + (1-t)h^2(y)$$

$$\Rightarrow \circledast \leq t h^2(x) + (1-t) h^2(y)$$

$$\Rightarrow h^2(tx+(1-t)y) \leq t h^2(x) + (1-t) h^2(y) \Rightarrow h^2 \text{ convexa}$$

Ahora con estos dos resultados, veamos como aplicarlos a (P0).

- $\|x\|_1$ es convexa por ser una norma
- $\|Ax+b\|$ es convexa por ser una norma compuesta con una función lineal (ver ej 1c)
- $\|Ax+b\|^2$ es convexa porque $\|Ax+b\|$ es convexa y no negativa
- $\frac{1}{2}\|Ax+b\|^2 + \|x\|_1$ es convexa por ser la suma positiva de funciones convexas

b) (QP) $\min\limits_{\substack{x \in \mathbb{R}^n \\ t \in \mathbb{R}^n}} \frac{1}{2} \|Ax - b\|^2 + \sum_{i=1}^{n} t_i = f(x,t)$

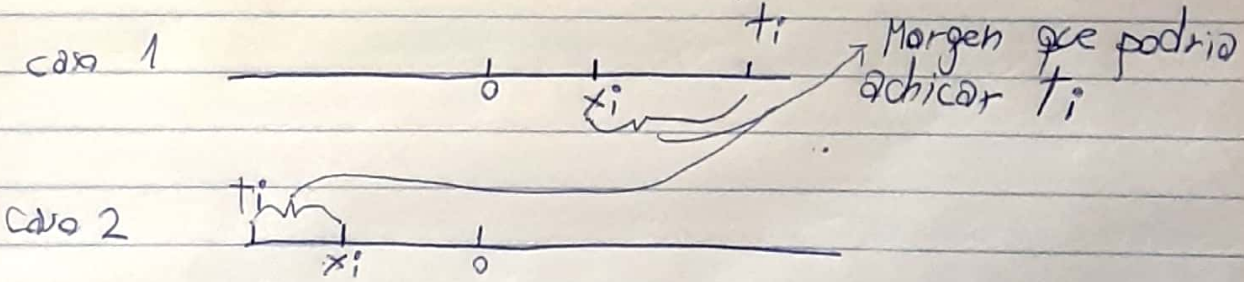
$\text{s.t: } x_i \leq t_i$

$-x_i \leq t_i$

Supongamos que $(\bar{x}, \bar{t})$ son solucion de QP
Sabemos que $t_i \geq |x_i| \ \forall i$ por cumplir las restricciones.

Vamos a probar que $\bar{t} = |\bar{x}|$, o lo que es lo mismo: $\bar{t}_i = |\bar{x}_i| \ \forall i$.

Supongamos por absurdo que $\bar{t}_i > |\bar{x}_i|$



$$f(\bar{x}, \bar{t}) = \frac{1}{2} \|A\bar{x} - b\|^2 + \sum_{i=1}^{n} \bar{t}_i \leq \frac{1}{2} \|A\bar{x} - b\|^2 + \sum_{i=1}^{n} |x_i|$$

Entonces si $\bar{t}_i > |x_i| \Rightarrow \bar{t}_i$ no seria el minimo
$\Rightarrow \bar{t}_i = |\bar{x}_i|$

Sabemos que en el minimo se cumple la siguiente relacion entre variables

$$\boxed{\bar{t}_i = |x_i|}$$

Escaneado con CamScanner

Entoces sabemos que podemos restringir la funcion $f(x,t)$ en $f(x,|x|)$ y buscar el minimo aqui y sera el mismo.

Buscar el minimo es $f(x,|x|)$ no es más que resolver el problema PQ

Ejercicio 5

Parte c

Como vimos en la parte a, $\frac{1}{2}\|A\mathbf{x} - b\|^2$.
Por otro lado $\sum_{i=1}^{n} t_i$ es convexo tratarse de la suma positiva de variables, o bien podría verse por medio del Hessiano, $\nabla^2 f(\mathbf{x})$, el cual es idénticamente igual a 0, $\nabla^2 f(\mathbf{x}) = 0$, lo que indica que es una función convexa.
Entonces, tenemos que $f(x,t)$ es convexa por ser la suma positiva de funciones convexas.

Para que (QP) sea convexo resta ver que la región factible también lo sea.

$$\begin{cases} x_i & \leq t_i \\ -x_i & \leq t_i \end{cases}$$

Esta región es un semiplano de R^{2n}, por lo que, como ya se menciono en otros ejercicio, es convexo. De esta forma (QP) también es convexo.

¿Puede un problema convexo ser equivalente a uno no convexo?
La respuesta es si.
A modo de ejemplo $f_1(x) = x^2$ y $f_2(x) = x^2(x^2 - 4x + 4.3)$ son problemas equivalentes ya que ambos tienen el minimo en el mismo punto $\bar{x} = 0$. Pero f_1 es convexa, mientras que f_2 no lo es. Se presenta una figura ilustrativa de lo anterior.

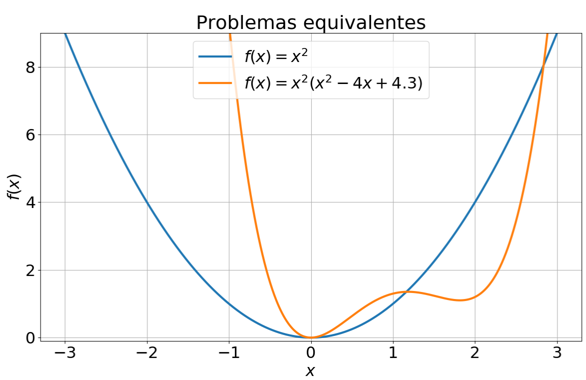


Figure 9: Problema convexo y no convexo equivalente

Parte d

Se utiliza el código presente en el archivo adjunto "ej6.py". La solución obtenida es

$$\bar{\mathbf{x}} = (-0.132746390.12730003)$$

$$\bar{\mathbf{t}} = (0.132746390.12730003)$$

Como era de esperar, los valores obtenidos con el solver mantienen la relación $\bar{t_i} = |\bar{x_i}| \ \forall_{i=1,2}$

Ejercicio 6

a) $\min_{x,y} \; 5x^2+5y^2+5x-3y-6xy+5/4 = f(x,y)$ (P)

S.t: $x^2+y^2 \leq R^2 \quad R\in\mathbb{R}^+$

Para probar que f es convexa calculemos $\nabla^2 f$.

$\nabla f = (10x+5-6y,\ 10y-3-6x)$

$\nabla^2 f = \begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix}$ cuyos valores propios son $\lambda_1 = 4 \quad \lambda_2 = 16$

Como los valores propios son todos mayores a 0 $(\lambda_i > 0 \ \forall i) \Rightarrow \nabla^2 f$ es definda positiva $\forall x,y$
$\Rightarrow$ f es convexa

La region factible es convexa por ser $B((0,0),R)$ y se probo su convexidad en 1)e) $\Rightarrow$ (P) es convexo

b) Como f es convexa, $\nabla f(\bar{x}) = 0$ es una condición suficiente y necesario de optimalidad

$\nabla f(\bar{x},\bar{y}) = 0 \Leftrightarrow \begin{cases} 10\bar{x} - 6\bar{y} = -5 \\ -6\bar{x} + 10\bar{y} = 3 \end{cases} \Rightarrow (\bar{x},\bar{y}) = \left(\frac{-1}{2}, 0\right)$

Sustituyendo en la restriccion; con $R<1/2$

$(-1/2)^2 + 0^2 \leq R^2 < \frac{1}{2}$

Lo cual no cumple la desigualdad $\Rightarrow$ la restricción esta activa

Ejercicio 6

Parte c

En primer lugar observemos la curvas de nivel en conjunto con la región factible.

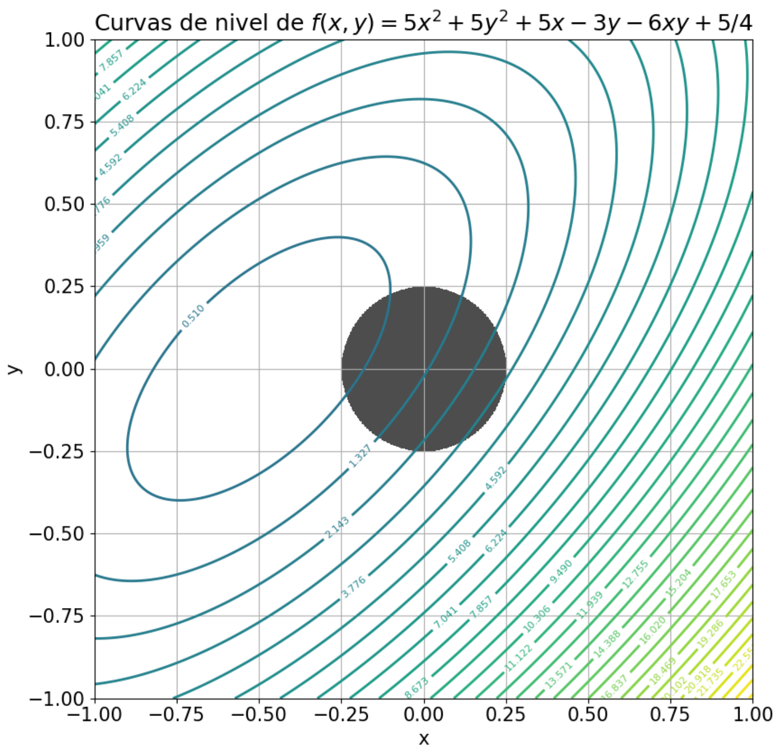
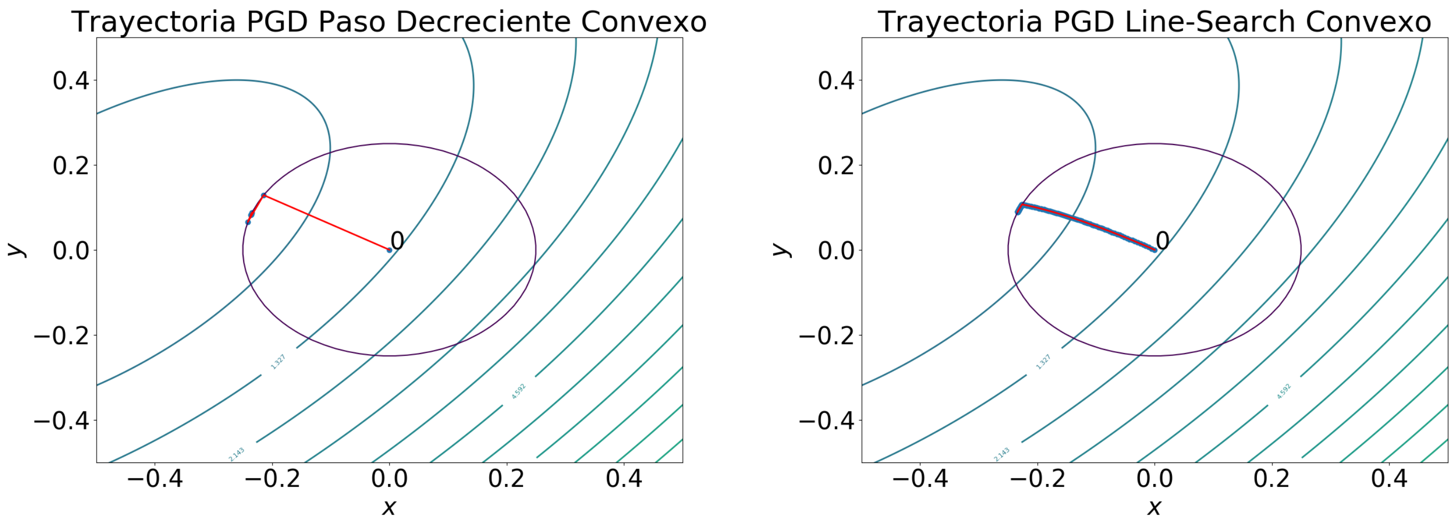


Figure 10: Curvas de nivel de f junto con la región factible.

Para resolver el problema se implemento el método del gradiente proyectado para el caso particular $\alpha^k = 1$, utilizando paso decreciente y line-search. En el caso de paso decreciente se utilizo $s_k = \frac{1}{k}$, mientras que en line-search se busco en una grilla de largo 0.001 con 100 puntos.

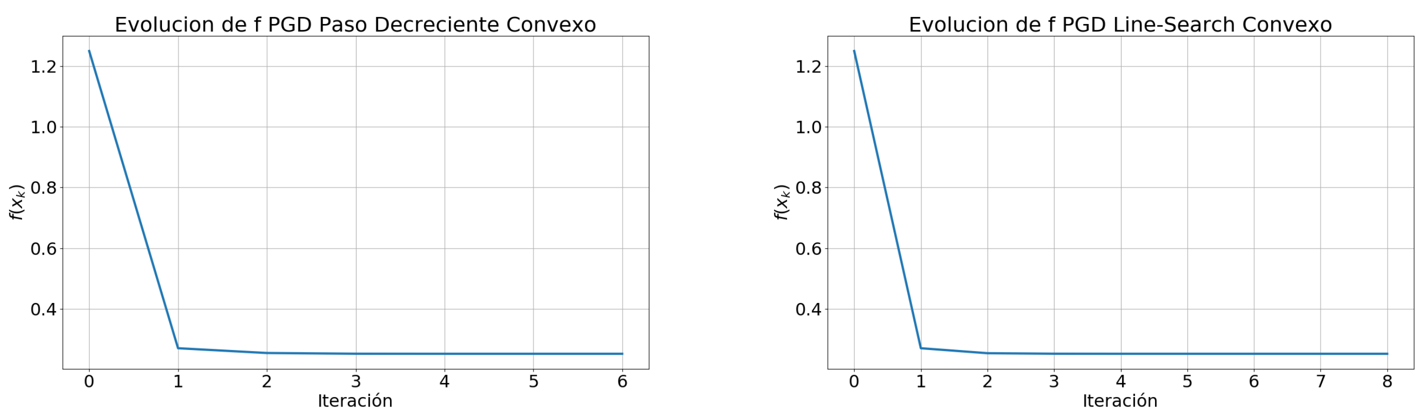
En ambos casos se partió desde el punto $(0, 0)$. Como criterio de parada se tomo $\frac{\|\mathbf{x_k} - \mathbf{x_{k-1}}\|}{\|\mathbf{x_{k-1}}\|} < \epsilon = 0.0001$.

Los resultados obtenidos se observan en la Figuras 11, 12, 13.



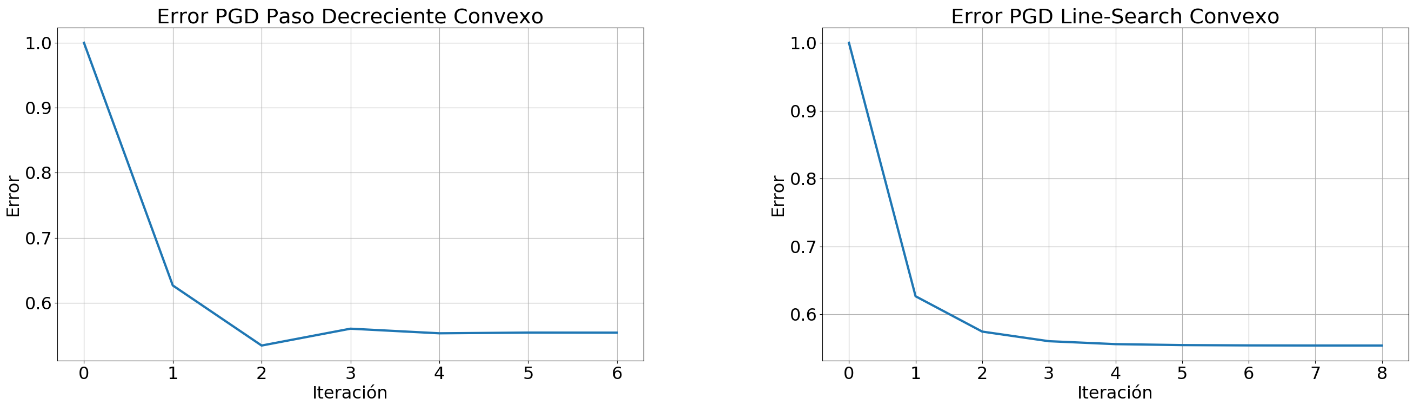
(a) (b)

Figure 11: Trayectoria de $\mathbf{x_k}$ junto a las curvas de nivel. En rojo se indica la trayectoria, y los puntos azules representan cada x_k



(a) (b)

Figure 12: Evolución del valor de la función de costo



(a) (b)

Figure 13: Evolución de $\left\|\left(x^k, y^k\right) - \left(x^{k-1}, y^{k-1}\right)\right\|$.

En la siguiente tabla se pueden ver los resultados numéricos obtenidos:

	Paso Decreciente	Line Search
Cantidad de Iteraciones	6	8
Punto obtenido	(-0.23581,0.08303)	(-0.23580,0.08305)

Se puede comprobar que el punto obtenido es correcto, ya que la sucesión se mueve hacia curvas de nivel de menor nivel y gráficamente el parece ser el oprimo

Para el caso del paso decreciente, se tiene un paso inicial alto, lo cual lleva a que inmediatamente, a partir de $k = 1$, x_k se comience a mover por el borde de la región factible.

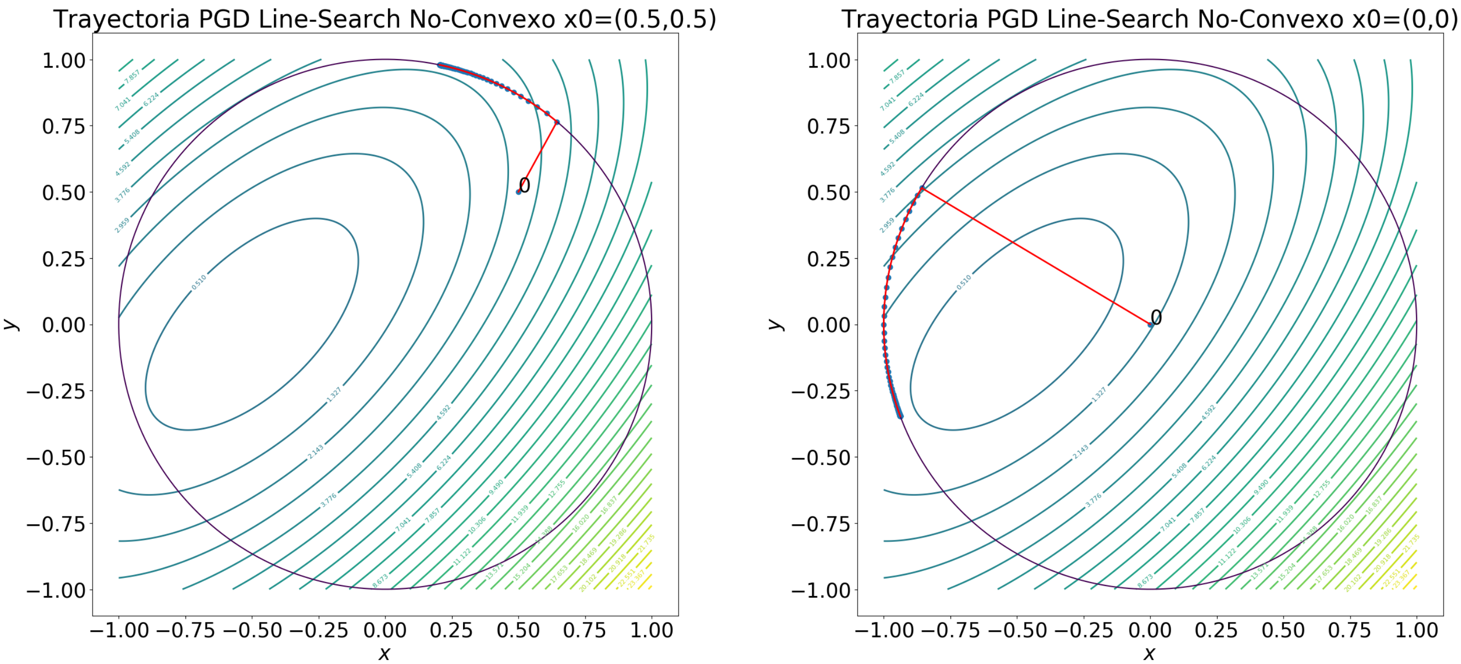
Parte d

$$\begin{gathered}\min_{(x,y)} 5x^2 + 5y^2 + 5x - 3y - 6xy + 5/4 \\ \text{s.t. } x^2 + y^2 \geq 1\end{gathered}$$

En este caso $f(x, y)$ es la misma, así que aun es convexa.

Lo que no es convexo es la región factible (C). Por ejemplo $(1, 0) \in C$ y $(0, 1) \in C$ pero sin embargo $\frac{1}{2}(1, 0) + \frac{1}{2}(0, 1) = (\frac{1}{2}, \frac{1}{2}) \notin C$, entonces C no es convexo.

Con el algoritmo con Line-Search, en este se obtiene la trayectoria de la Figura 14



(a) $x_0 = (\frac{1}{2}, \frac{1}{2})$

(b) $x_0 = (0, 0)$

Figure 14: Trayectoria en rojo, en conjunto con las lineas de nivel y el borde de la región factible en violeta.

Para ambas condiciones iniciales se observa que la sucesión intenta ingresar a la región factible, pero la proyección los envía hacia el borde, haciendo que todos x_k se encuentren sobre esta salvo para $k = 0$.

Además, para cada uno de los x_0 se tienen soluciones distintos. Esto se debe a que tenemos un problema no convexo, lo cual lleva a no tener garantía de convergencia al óptimo.

Los puntos óptimos obtenidos son

	Punto
$x_0 = (0, 0)$	(0.20433,0.97890)
$x_0 = (0.5, 0.5)$	(-0.93769,-0.34746)

Para el caso de paso decreciente sucede el mismo fenómeno, convergiendo a los mismos puntos.