
**MÉTODOS NUMÉRICOS DE OPTIMIZACIÓN
CON RESTRICCIONES (O0384)**

Contents

1 Algunos ejemplos de motivación	3
2 Minimización sin restricciones de funciones escalares de varias variables	6
2.1 Notación y nociones elementales	6
2.2 Nociones básicas, óptimos locales y globales	10
2.3 Condiciones necesarias y suficientes de primer o segundo orden para funciones diferenciables	11
2.4 Convexidad	15
2.5 Ejercicios.	18
3 Metodologías numéricas	20
3.1 Direcciones de descenso	20
3.2 Búsqueda lineal en una dirección	23
3.3 Convergencia	27
3.4 Método del gradiente o de Cauchy	31
3.5 Método de Newton	33
3.6 Métodos de quasi-Newton	40
3.7 Función de Rosenbrock	48
3.8 Método de direcciones conjugadas	53
3.9 Método de región de confianza	61
3.10 Ejercicios.	68
4 Teoría básica de Optimización con restricciones	71
4.1 Condiciones necesarias de primer orden para problemas con restricciones de igualdad	71
4.2 Condiciones necesarias de primer orden para problemas con restricciones de desigualdad	81
4.3 El abordaje de conos	87
4.4 Condiciones de calidad	95
4.5 Condiciones de optimalidad de segundo orden	97
4.6 Ejercicios.	103
5 Problemas con restricciones especiales	106
5.1 Método del gradiente proyectado	110
5.2 Problemas con restricciones lineales	114
5.3 Método de Conjuntos activos para programación cuadrática con restricciones lineales de desigualdad	118
5.4 Ejercicios.	125
6 Métodos numéricos para problemas generales de optimización	126
6.1 Métodos de penalidad	126
6.2 Método de Lagrangiano Aumentado	136
6.3 Método de programación cuadrática sucesiva	153
6.4 Métodos de barrera	159
6.5 Ejercicios.	163

1 Algunos ejemplos de motivación

- Ejemplo 1: un problema de planificación del personal. [1]

Se considera una sala de un hospital que trabaja las 24 horas del día. En distintos momentos del día, varía la cantidad de personal profesional que se necesita. La siguiente tabla muestra la demanda de personal durante 6 turnos de trabajo:

Turnos	1	2	3	4	5	6
Horas	0-4	4-8	8-12	12-16	16-20	20-24
Demanda	8	10	12	10	8	6

Cada profesional debe trabajar 2 turnos consecutivos. El objetivo es minimizar el número de profesionales cumpliendo la demanda.

Si llamamos

x_j := número de profesionales que comienzan a trabajar en el turno $j, j = 1, 2, \dots, 6$.

El número total de profesionales está dado por $x_1 + x_2 + \dots + x_6$.

Se tiene el siguiente calendario:

Turnos	x_1	x_2	x_3	x_4	x_5	x_6	Req
1	✓					✓	8
2	✓	✓					10
3		✓	✓				12
4			✓	✓			10
5				✓	✓		8
6					✓	✓	6

El número total de profesionales que trabajan en el primer turno está dado por $x_1 + x_6$ y debe ser $x_1 + x_6 \geq 8$.

Luego, el problema se puede formular como:

$$\begin{array}{ll} \text{Minimizar} & x_1 + x_2 + \dots + x_6 \\ \text{sujeto a} & \end{array}$$

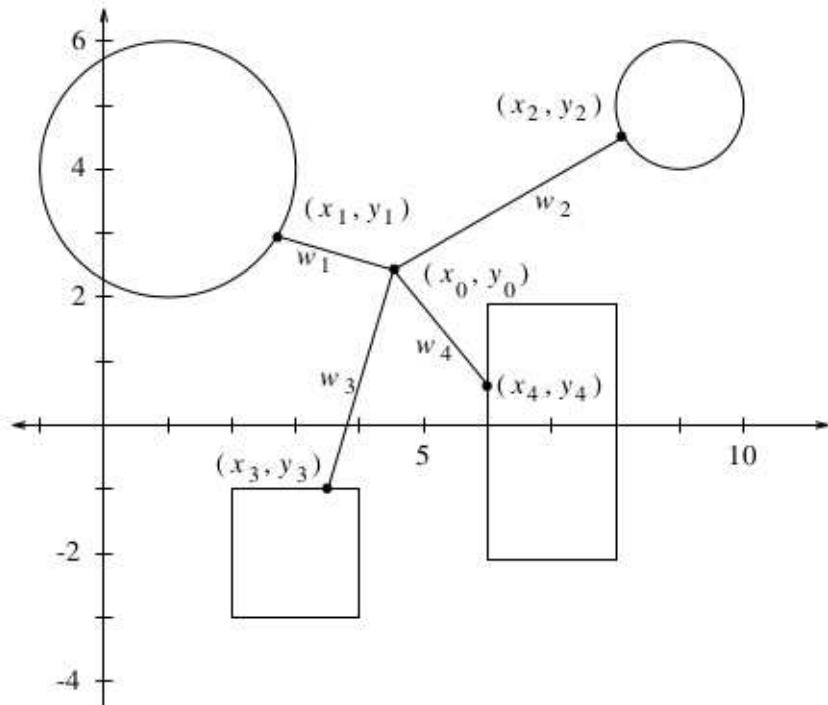
$$\begin{aligned} x_1 + x_6 &\geq 8 \\ x_1 + x_2 &\geq 10 \\ x_2 + x_3 &\geq 12 \\ x_3 + x_4 &\geq 10 \\ x_4 + x_5 &\geq 8 \\ x_5 + x_6 &\geq 6 \\ x_i &\geq 0, i = 1, \dots, 6. \end{aligned}$$

Este problema es un problema de programación lineal (se minimiza una función lineal sujeto a restricciones lineales) y tiene 2 soluciones: $x^* = (4, 6, 6, 4, 4, 4)$ y $x^{**} = (8, 2, 10, 0, 8, 0)$ con valor objetivo: $f = 28$.

- Ejemplo 2. [2]

Muchos problemas de optimización no lineal surgen a menudo en la ciencia y la ingeniería.

Suponga que se deben conectar cuatro edificios mediante cables eléctricos. Las posiciones y forma de los edificios se ilustran en la siguiente figura:



Los dos primeros edificios son circulares: uno con centro en $(1, 4)$ y radio 2, el otro con centro en $(9, 5)$ y radio 1. El tercer edificio es cuadrado con lados de longitud 2 centrado en $(3, -2)$. El cuarto edificio es rectangular con altura 4 y ancho 2 centrado en $(7, 0)$. Los cables eléctricos se unirán en algún punto central a determinar (x_0, y_0) y se conectará a cada edificio i en una posición a determinar (x_i, y_i) . El objetivo es minimizar la cantidad de cable a utilizar. Denominamos w_i a la longitud del cable que conecta el edificio i con (x_0, y_0) . Luego, un modelo para este problema es

$$\begin{aligned}
 & \text{Minimizar} && w_1 + w_2 + w_3 + w_4 \\
 & \text{sujeto a} && \\
 & && w_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}, i = 1, \dots, 4 \\
 & && (x_1 - 1)^2 + (y_1 - 4)^2 \leq 4 \\
 & && (x_2 - 9)^2 + (y_2 - 5)^2 \leq 1 \\
 & && 2 \leq x_3 \leq 4 \\
 & && -3 \leq y_3 \leq -1 \\
 & && 6 \leq x_4 \leq 8 \\
 & && -2 \leq y_4 \leq 2.
 \end{aligned}$$

Este problema se puede reformular como un problema de programación no lineal diferenciable con restricciones de igualdad y desigualdad, tema del presente curso.

- Ejemplo 3. [3]

Supongamos que se quiere comprar un auto nuevo entre cuatro modelos diferentes: un VW Golf, un Opel Astra, un Ford Focus y un Toyota Corolla. La decisión se tomará de acuerdo con el precio, el consumo de gasolina y la potencia. El cliente prefiere un coche barato, con alta potencia y bajo consumo de gasolina. En este caso, nos enfrentamos a un problema de decisión con cuatro alternativas (los cuatro modelos de autos) y tres criterios (las preferencias del cliente).

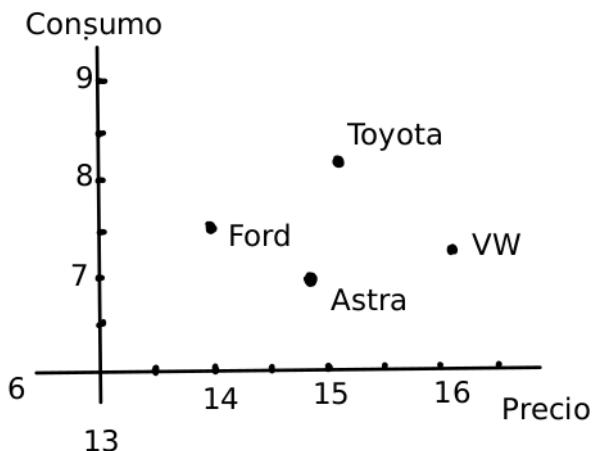
Supongamos que contamos con las siguientes características de los cuatro modelos de auto (los datos son inventados):

Criterio	VW	Astra	Ford	Toyota
Precio (100.000 pesos)	16.2	14.9	14.0	15.2
Consumo ($\frac{l}{100km}$)	7.2	7.0	7.5	8.2
Potencia (kW)	66.0	62.0	55.0	71.0

¿Cómo decidimos, cuál de los cuatro coches es la "mejor" alternativa, cuando el coche más potente es también el que consume más combustible, por lo que no podemos comprar un coche que sea barato, potente y que ahorre combustible.

Este problema es un ejemplo de lo que se conoce como problema multiobjetivo, en los cuales hay varios objetivos en conflicto que deben minimizarse simultáneamente.

Observar que si consideramos Precio versus Consumo se tiene el siguiente gráfico:



Se puede ver que Astra y Ford son dos opciones eficientes en el siguiente sentido: para ambos no hay otra alternativa que sea a la vez más barata y que consuma menos gasolina. Además, tanto Toyota como VW son más caros y consumen más gasolina que el Astra.

En este caso, la solución al problema es un elemento del conjunto discreto

$\Omega = \{ \text{VW Golf, Opel Astra, Ford Focus, Toyota Corolla} \}$. Si denotamos la función precio con f_1 y la función consumo con f_2 se tiene que las funciones $f_1, f_2 : \Omega \rightarrow \mathbb{R}$ son las funciones que se deben minimizar para encontrar una solución. El problema se modeliza de la siguiente forma:

$$\begin{aligned} & \text{Minimizar } (f_1(x), f_2(x)) \\ & \text{sujeto a } x \in \Omega \end{aligned}$$

problema que aborda la optimización multiobjetivo.

2 Minimización sin restricciones de funciones escalares de varias variables

2.1 Notación y nociones elementales

Para una función escalar de una variable $f : \mathbb{R} \rightarrow \mathbb{R}$ derivable hasta orden 2 sabemos que:

- Condición necesaria optimalidad de primer orden: $f'(x) = 0$. (punto estacionario):

$$\boxed{\text{Si } x^* \text{ es minimizador entonces } f'(x^*) = 0}$$

- Condición necesaria optimalidad de segundo orden: $f''(x) \geq 0$ y $f'(x) = 0$:

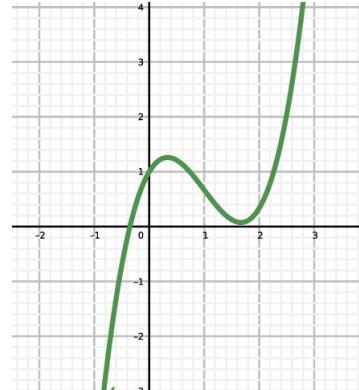
$$\boxed{\text{Si } x^* \text{ es minimizador entonces } f'(x^*) = 0 \text{ y } f''(x^*) \geq 0.}$$

- Condición suficiente de optimalidad: $f''(x) > 0$ y $f'(x) = 0$:

$$\boxed{\text{Si } f'(x^*) = 0 \text{ y } f''(x^*) > 0 \text{ entonces } x^* \text{ es minimizador local.}}$$

Ejemplo: $f(x) = x^3 - 3x^2 + 5/3x + 1$.

- Puntos estacionarios: $x_1 = 5/3, x_2 = 1/3$
- $f''(5/3) = 4$
- $f''(1/3) = -4$

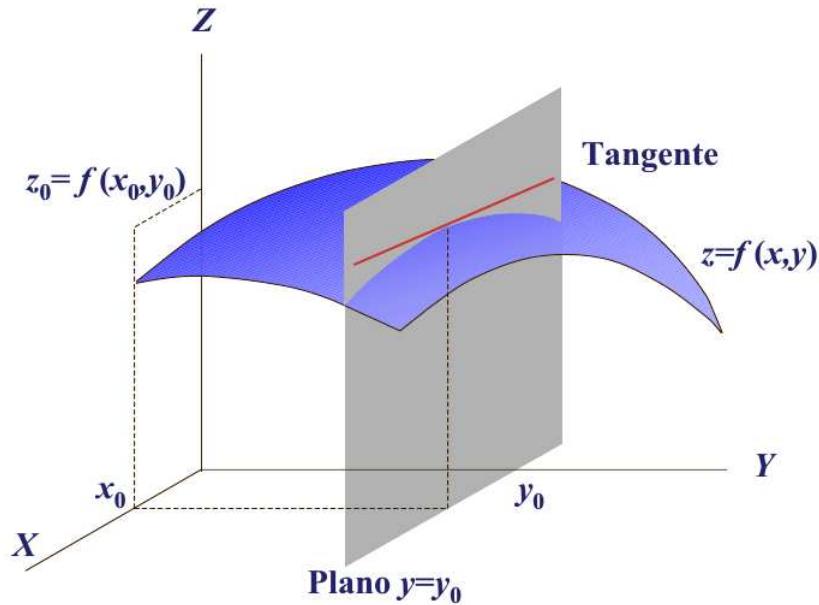


Para el caso de varias variables, denotamos a los puntos en \mathbb{R}^n como $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

De manera análoga, los vectores se denotan como $d = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$. Decimos que d_i es la i -ésima componente del vector d .

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable (en general vamos a asumir que todas las derivadas parciales de f existen y son continuas, es decir que f es C^1), el *vector gradiente de f* se define como

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}.$$



Si existen las derivadas de orden superior de f y son continuas, las derivadas cruzadas son iguales y se define la *matriz Hessiana* o matriz de las derivadas segundas de f a la matriz:

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f(x)}{\partial^2 x_n} \end{pmatrix}.$$

Si se considera la función vectorial: $F = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, su derivada primera se llama Jacobiano y se define como

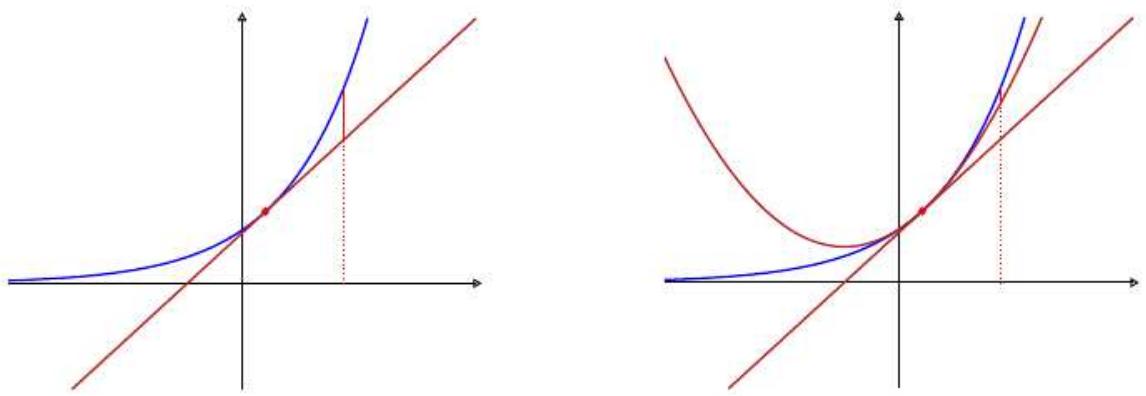
$$JF(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{pmatrix}.$$

Si f tiene derivadas segundas continuas se obtiene su desarrollo de Taylor alrededor de x_0 :

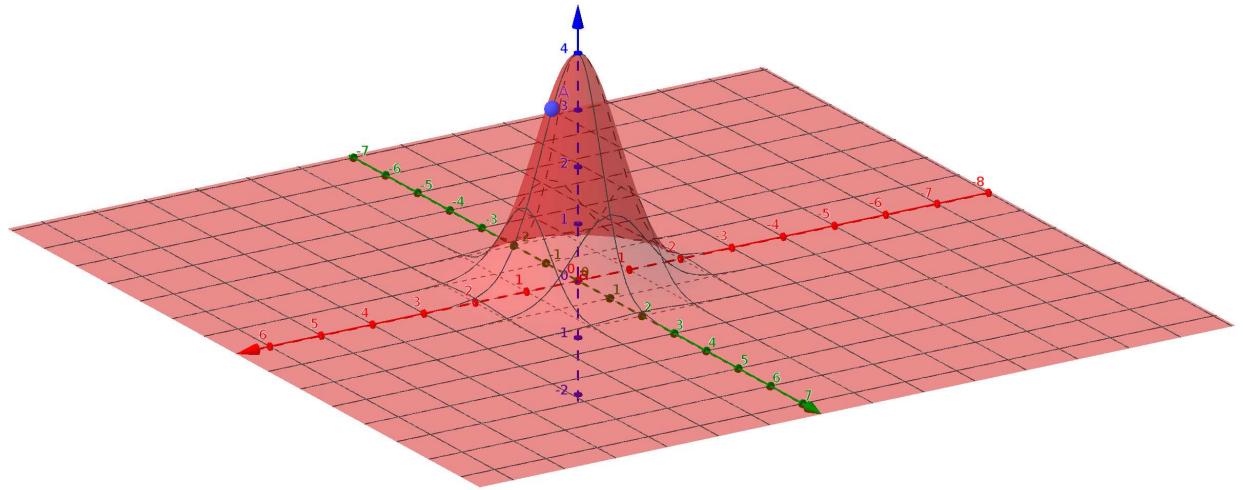
$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0) + o(\|x - x_0\|^2)$$

donde $\frac{o(\|x - x_0\|^2)}{\|x - x_0\|^2} \rightarrow 0$.

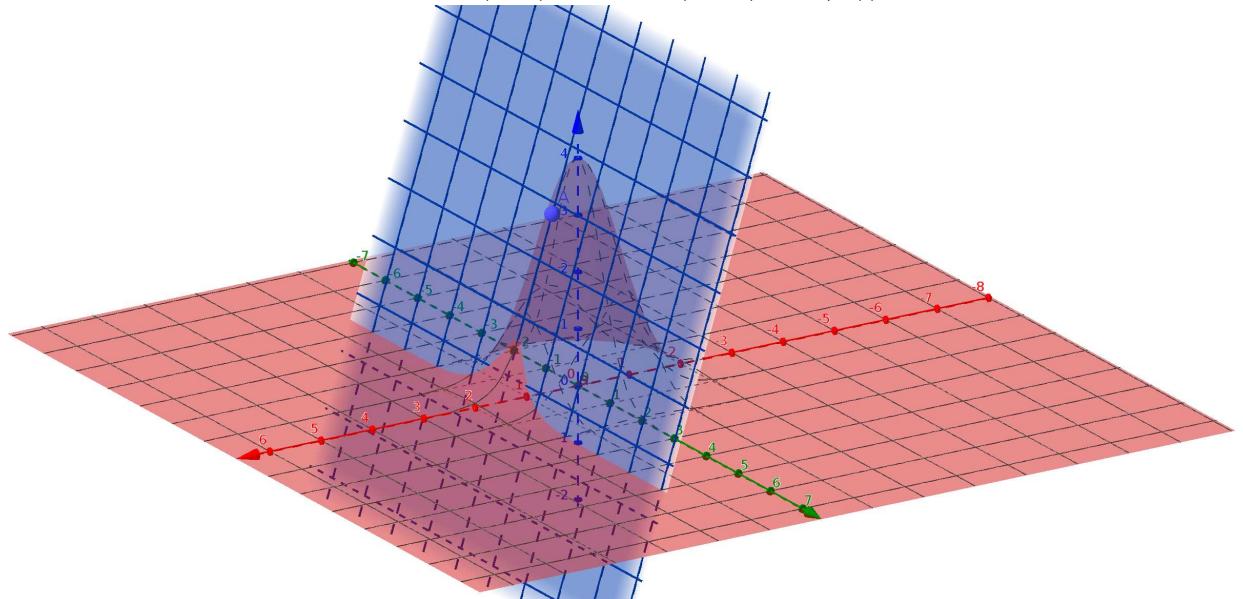
En las siguientes figuras se pueden ver las aproximaciones de Taylor de orden 1 y 2 de la curva en azul.



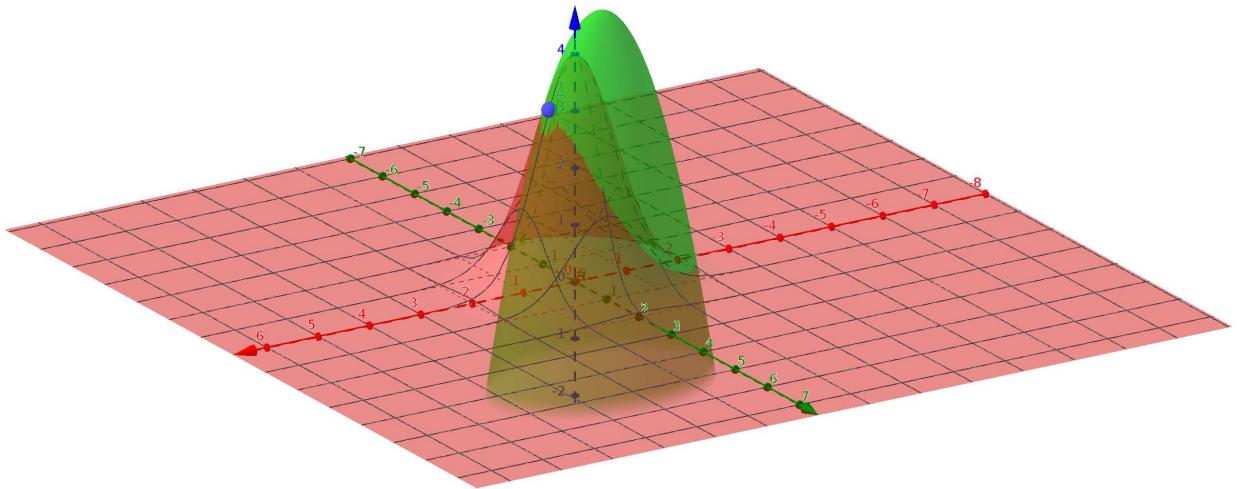
Para la función $f(x, y) = 4e^{-(x^2+y^2)}$, alrededor de $(1/2, 0)$ se tiene:



Polinomio de Taylor de orden 1: $P_1(x, y) = 4e^{-1/4}(1 - (x - 1/2))$



Polinomio de Taylor de orden 2: $P_2(x, y) = 4e^{-1/4}(1 - (x - 1/2) - 1/2(x - 1/2)^2 - y^2)$



Dada una matriz $A \in \mathbb{R}^{n \times n}$, $A = [a_1 \dots a_n]$, se tiene:

- Si $\det(A) = 0$ entonces $\{a_1, \dots, a_n\}$ son linealmente dependientes (LD) y A es singular.
- Si $\det(A) \neq 0$ entonces $\{a_1, \dots, a_n\}$ son linealmente independientes (LI) y A es no singular.

Además, si existe $x \neq 0, x \in \mathbb{R}^n$ tal que $Ax = 0$ entonces A es singular.

Si, para $x \neq 0, x \in \mathbb{R}^n$, $Ax = \lambda x$ entonces $(A - \lambda I)x = 0$ y $\det(A - \lambda I) = 0$. En este caso se tiene que λ es *autovalor* de A y x es *autovector* de A .

Una matriz que cumple $A = A^\top$ es simétrica y tiene n autovalores reales λ_i . Además, A posee una base ortonormal de autovectores y es diagonalizable.

Lema 2.1. *Sea $A \in \mathbb{R}^{n \times n}$ una matriz simétrica tal que λ_1 y λ_n son el menor y el mayor autovalor respectivamente. Entonces, para todo $x \in \mathbb{R}^n$:*

$$\lambda_1 \|x\|^2 \leq x^T Ax \leq \lambda_n \|x\|^2.$$

Definición 2.1. *Sea $A \in \mathbb{R}^{n \times n}$ una matriz simétrica. Decimos que A es definida positiva si $x^T Ax > 0$ para todo $x \neq 0$. Se denota con la expresión $A > 0$.*

Decimos que A es semidefinida positiva si $x^T Ax \geq 0$ para todo $x \in \mathbb{R}^n$. Se denota con la expresión $A \geq 0$.

Lema 2.2. *Sea $A \in \mathbb{R}^{n \times n}$ una matriz simétrica definida positiva. Entonces, existe $B \in \mathbb{R}^{n \times n}$ simétrica definida positiva tal que $A = B^T B$.*

2.2 Nociones básicas, óptimos locales y globales

Consideramos el problema general de optimización no lineal:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } x \in \Omega \subset \mathbb{R}^n \end{aligned} \tag{1}$$

para $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^1 en un conjunto abierto que contiene a Ω . La función f será llamada función objetivo. El conjunto Ω se denomina conjunto factible.

Consideramos al conjunto \mathbb{R} con el orden usual, de esta manera tiene sentido hablar de minimizador, aunque sabemos que no siempre puede garantizarse existencia de un minimizador cuando $\Omega = \mathbb{R}^n$.

Pueden considerarse diferentes representaciones del conjunto Ω :

1. $\Omega = \mathbb{R}^n$ representa el problema de minimización irrestricta.

2. $\Omega = \{x \in \mathbb{R}^n; Ax = b\}, A \in \mathbb{R}^{m \times n}$, problema de minimización con restricciones lineales de igualdad.

Si $\Omega = \{x \in \mathbb{R}^n; Ax \leq b\}, A \in \mathbb{R}^{m \times n}$, problema de minimización con restricciones lineales de desigualdad.

Si $\Omega = \{x \in \mathbb{R}^n; Ax = b, Cx \leq d\}, A \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{p \times n}$, problema de minimización con restricciones lineales.

Si además, $f(x) = e^T x$ el problema se llama problema de programación lineal.

Si, $f(x) = x^T H x + e^T x$ con H matriz simétrica, el problema se llama problema de programación cuadrática.

3. Si $\Omega = \{x \in \mathbb{R}^n; h(x) = 0\}$ con $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ el problema es de programación no lineal con restricciones de igualdad.

4. Si $\Omega = \{x \in \mathbb{R}^n; h(x) = 0, g(x) \leq 0\}$ con $h : \mathbb{R}^n \rightarrow \mathbb{R}^m, g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ el problema es de programación no lineal con restricciones de igualdad y desigualdad o problema general de programación no lineal.

Definición 2.2. Decimos que $x^* \in \Omega$ es un minimizador global de $f(x)$ en Ω si

$$f(x^*) \leq f(x), \forall x \in \Omega.$$

Definición 2.3. Decimos que $x^* \in \Omega$ es un minimizador global estricto de $f(x)$ en Ω si

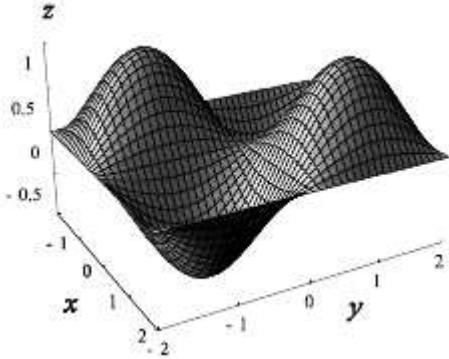
$$f(x^*) < f(x), \forall x \in \Omega, x \neq x^*.$$

Definición 2.4. Decimos que x^* es un minimizador local de $f(x)$ en Ω si

$$f(x^*) \leq f(x), \forall x \in \Omega \cap B(x^*, \varepsilon).$$

Definición 2.5. Decimos que x^* es un minimizador local estricto de $f(x)$ en Ω si

$$f(x^*) < f(x), \forall x \in \Omega \cap B(x^*, \varepsilon), x \neq x^*.$$



2.3 Condiciones necesarias y suficientes de primer o segundo orden para funciones diferenciables

Consideramos el problema de optimización (1) para el caso $\Omega = \mathbb{R}^n$.

Teorema 2.6. Condición necesaria de primer orden. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^1 . Si x^* es un minimizador local de f en \mathbb{R}^n entonces $\nabla f(x^*) = 0$.*

Demostración.

Sea $d \in \mathbb{R}^n$, $d \neq 0$ arbitraria.

Como x^* es un minimizador local de f entonces existe $\delta > 0$ tal que

$$f(x^*) \leq f(x^* + td) \quad \forall t \in [0, \delta]. \quad (2)$$

Por Taylor,

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^T d + r(t)$$

donde $\frac{r(t)}{t} \rightarrow 0$.

Usando (2) y dividiendo por $t > 0$ en la última expresión obtenemos

$$0 \leq \nabla f(x^*)^T d + \frac{r(t)}{t}.$$

Tomando límite cuando $t \rightarrow 0^+$ se obtiene

$$\nabla f(x^*)^T d \geq 0.$$

Luego, si fuera $\nabla f(x^*) \neq 0$ podríamos elegir $d = -\nabla f(x^*)$ y tendríamos $\|\nabla f(x^*)\|^2 \leq 0$. Absurdo.

Por lo tanto $\nabla f(x^*) = 0$. \square

Definición 2.7. *Un punto x^* que cumple $\nabla f(x^*) = 0$ se denomina punto estacionario de f .*

Observación: Alcanza con que f sea diferenciable.

Como sucede con funciones de una variable, el Teorema previo nos proporciona una condición necesaria: que no basta para que el punto sea un minimizador local. Por ejemplo, si consideramos la función $f(x, y) = x^3 + y^3$ diferenciable, vemos que el gradiente de f se anula en el origen de coordenadas, sin embargo ese punto no es ni minimizador ni maximizador local:

- $f(0, 0) = 0$,

- Para h pequeño positivo, $f(h, 0) = h^3 > 0$,
- Para h pequeño negativo, $f(h, 0) = h^3 < 0$.

El punto $(0, 0)$ es un punto silla o de ensilladura.

Definición 2.8. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función diferenciable en x_0 punto estacionario. Se dice que x_0 es un punto silla o de ensilladura de f si para todo $r > 0$ existen puntos $x \in B(x_0, r), x \neq x_0$ tal que $f(x) > f(x_0)$ y puntos $y \in B(x_0, r), y \neq x_0$ tal que $f(y) < f(x_0)$.

Ejemplos: 1. Consideremos la función $f(x, y) = x^2 + 3y^2 - 2x + 6y + 14$. Como f es diferenciable buscamos los puntos estacionarios. Luego

$$\nabla f(x, y) = (0, 0) \text{ si y solo si } x = 1, y = -1.$$

Observar que, completando cuadrados, tenemos que

$$f(x, y) = (x - 1)^2 + 3(y + 1)^2 + 10$$

y la gráfica de $z = f(x, y)$ representa un paraboloide elíptico con ramas hacia arriba que tiene su vértice en $(1, -1, 10)$. Luego, $f(x, y) \geq f(1, -1) = 10$ para cualquier punto (x, y) y $(1, -1)$ es un minimizador local y global de f .

2. Consideremos la función $f(x, y) = x^2 - y^2$. Como f es diferenciable buscamos los puntos estacionarios. Luego

$$\nabla f(x, y) = (0, 0) \text{ si y solo si } x = 0, y = 0.$$

Pero, sabemos que la gráfica de $z = f(x, y)$ es una paraboloide hiperbólico y se tiene que

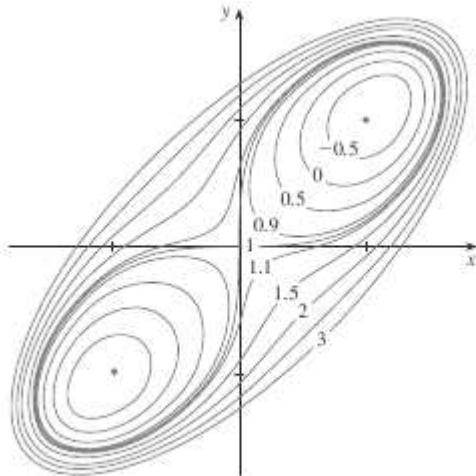
- $f(0, 0) = 0$,
- Para h pequeño positivo, $f(h, 0) = h^2 > 0$,
- Para h pequeño positivo, $f(0, h) = -h^2 < 0$

y concluimos que $(0, 0)$ es un punto silla de f .

3. Consideremos la función $f(x, y) = x^4 + y^4 - 4xy + 1$. Como f es diferenciable buscamos los puntos estacionarios. Luego

$$\nabla f(x, y) = (0, 0) \text{ si y solo si } (x, y) = (0, 0), (x, y) = (1, 1), (x, y) = (-1, -1).$$

Observar las curvas de nivel de f y concluir sobre la naturaleza de los puntos estacionarios.



El último teorema nos proporciona una condición necesaria para buscar los posibles extremos locales de una función diferenciable, pero necesitamos criterios que nos permitan afirmar cuando un punto estacionario de una función es un extremo local. En el caso de funciones reales de variable real, la derivada segunda desempeña un papel importante para saber si una función alcanza o no un extremo local en un punto estacionario. En el caso de funciones de varias variables reales, este papel corresponde a la matriz de las derivadas parciales segundas de la función.

Teorema 2.9. Condición necesaria de segundo orden. *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^2 . Si x^* es un minimizador local de f entonces $\nabla f(x^*) = 0$ y $\nabla^2 f(x^*)$ es semidefinida positiva (esto es $d^T \nabla^2 f(x^*) d \geq 0$ para todo $d \in \mathbb{R}^n$).*

Demostración.

La primera parte se obtiene del teorema anterior. Para probar la segunda parte, para $d \neq 0$ consideramos el desarrollo de Taylor de orden dos de f alrededor de x^* :

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^T d + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + r(t)$$

donde $\frac{r(t)}{t^2} \rightarrow 0$.

Como x^* es minimizador local, $\nabla f(x^*) = 0$ entonces para t suficientemente pequeño ($0 < t < \delta$) se tiene

$$0 \leq f(x^* + td) - f(x^*) = \frac{t^2}{2} d^T \nabla^2 f(x^*) d + r(t).$$

Dividiendo por t^2 y tomando límite cuando $t \rightarrow 0^+$ obtenemos

$$d^T \nabla^2 f(x^*) d \geq 0.$$

□

Observación 1: Se usa en el Teorema que: Si x^* es un minimizador local de f entonces

$$\forall d \in \mathbb{R}^n : f(x^*) \leq f(x^* + td) \text{ para todo } t \text{ pequeño.}$$

La recíproca no vale:

- Para $f(x, y) = (x - y^2)(x - \frac{1}{2}y^2)$ se tiene que $(0, 0)$ es el único punto estacionario que NO es minimizador. Sin embargo vale la condición.

Observación 2: La condición del teorema anterior es necesaria pero no suficiente, considerar por ejemplo, la función $f(x, y) = x^3 + y^3$ en el origen. Vemos que $\nabla^2 f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$ que se anula en el origen y por lo tanto cumple ser semidefinida positiva mientras que $(0, 0)$ no es minimizador local.

El próximo Teorema nos proporcionará una condición suficiente de optimalidad.

Teorema 2.10. Condición suficiente. *Sean f una función C^2 y x^* un punto estacionario de f . Si $\nabla^2 f(x^*)$ es definida positiva entonces x^* es un minimizador local de f .*

Demostración.

Como $\nabla^2 f(x^*)$ es una matriz definida positiva entonces todos los autovalores de $\nabla^2 f(x^*)$ son estrictamente positivos. Sea $\lambda > 0$ el menor autovalor de $\nabla^2 f(x^*)$.

Por Taylor, sea $d \neq 0$:

$$f(x^* + d) = f(x^*) + \nabla f(x^*)^T d + \frac{1}{2} d^T \nabla^2 f(x^*) d + r(d)$$

donde $\frac{r(d)}{\|d\|^2} \rightarrow 0$. Entonces:

$$f(x^* + d) = f(x^*) + \frac{1}{2} d^T \nabla^2 f(x^*) d + r(d)$$

$$\frac{f(x^* + d) - f(x^*)}{\|d\|^2} = \frac{1}{2} \frac{d^T \nabla^2 f(x^*) d}{\|d\|^2} + \frac{r(d)}{\|d\|^2} \geq \frac{\lambda}{2} + \frac{r(d)}{\|d\|^2}.$$

Como $\lim_{d \rightarrow 0} \frac{\lambda}{2} + \frac{r(d)}{\|d\|^2} = \frac{\lambda}{2} > 0$ existe $\delta > 0$ tal que $\forall d \neq 0 : \|d\| < \delta$

$$\frac{f(x^* + d) - f(x^*)}{\|d\|^2} > 0.$$

Por lo tanto:

$$f(x^* + d) > f(x^*)$$

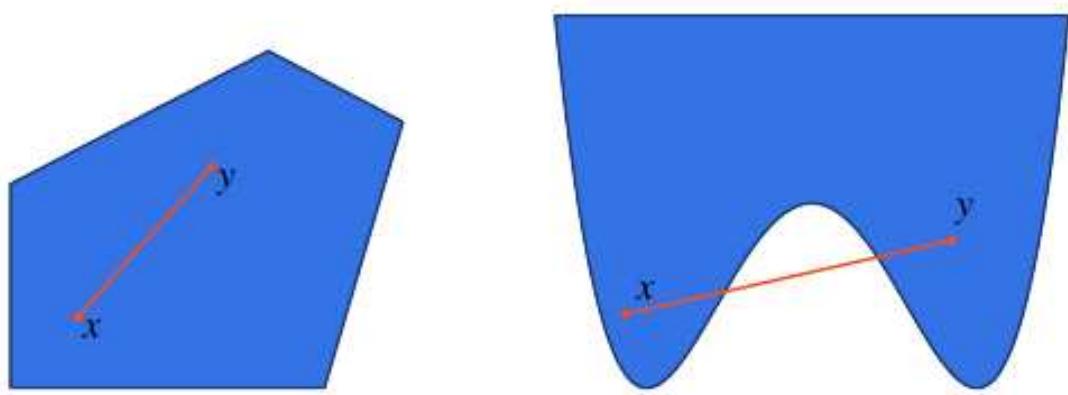
y x^* es un minimizador local estricto de f . □

2.4 Convexidad

En esta sección presentaremos algunos resultados elementales relacionados con las nociones de conjuntos convexos y funciones convexas, que son esenciales para la teoría de optimización no lineal.

Definición 2.11. Decimos que un conjunto $\Omega \subset \mathbb{R}^n$ es convexo si el segmento entre dos puntos cualesquiera de Ω está íntegramente contenido en el conjunto: $\alpha x + (1 - \alpha)y \in \Omega$ para todo $x, y \in \Omega, \alpha \in [0, 1]$.

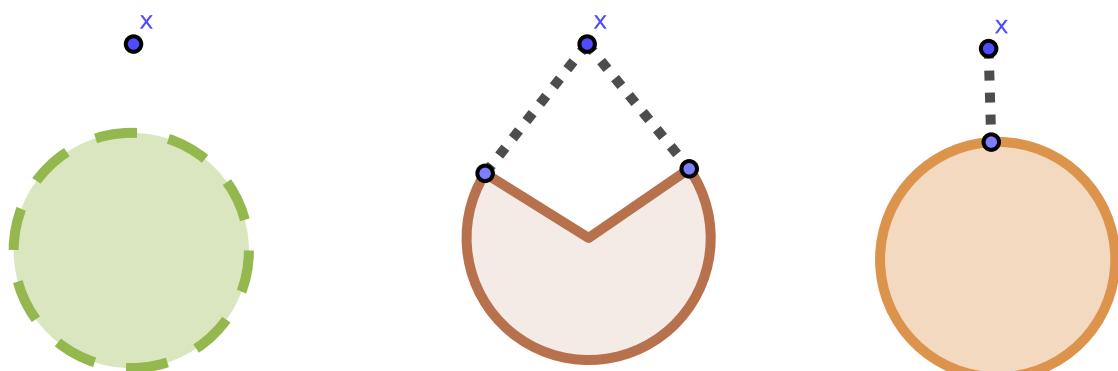
Ejemplos de un conjunto convexo y otro no convexo.



Definición 2.12. La proyección ortogonal de un punto $x \in \mathbb{R}^n$ sobre un conjunto $D \subset \mathbb{R}^n$ es el punto de D más cercano al punto x :

$$y^* = P_D(x) \Leftrightarrow y^* \text{ es solución global del problema } \text{Min}_{y \in D} \|y - x\|$$

El problema puede no tener solución y si tiene, no siempre se puede garantizar la unicidad.



Teorema 2.13. 1. Si $D \neq \emptyset$ es cerrado entonces existe $P_D(x), \forall x \in \mathbb{R}^n$.

2. Si $D \neq \emptyset$ es cerrado y convexo entonces existe $P_D(x) \forall x \in \mathbb{R}^n$ y es única.

Demostración.

1. Como $D \neq \emptyset$ existe $z \in D$. Sea $x \notin D$. Consideramos el conjunto

$$A = \underbrace{D}_{\text{cerrado}} \cap \underbrace{\{w \in \mathbb{R}^n : \|w - x\| \leq \|z - x\|\}}_{\text{cerrado y acotado}}.$$

A es un conjunto compacto.

Ahora consideramos el problema $\min_{y \in A} \|y - x\|$. La función $f(y) = \|y - x\|$ es continua en A compacto entonces el problema tiene solución $y^* \in A$. Luego, $y^* = P_D(x)$ ya que $\|y^* - x\| \leq \|y - x\|, \forall y \in D$.

2. Suponemos que $P_D(x)$ no es único. Entonces existe $z \in D : z \neq P_D(x)$ y $\|z - x\| = \|P_D(x) - x\|$.

Definimos $v = \frac{1}{2}(z + P_D(x))$. Por ser D un conjunto convexo, $v \in D$. Entonces

$$\|P_D(x) - x\|^2 = \|v - x\|^2 + \|v - P_D(x)\|^2 > \|v - x\|^2$$

que es un absurdo porque $P_D(x)$ es el minimizador. □

Proposición 2.1. Sea $D \neq \emptyset$ un conjunto cerrado y convexo. Sea $x \in \mathbb{R}^n$. Entonces:

1. Para todo $y \in D$ se tiene que $(x - P_D(x))^T(y - P_D(x)) \leq 0$.
2. Si $\bar{z} \in D$ cumple que $(x - \bar{z})^T(y - \bar{z}) \leq 0, \forall y \in D$ entonces $\bar{z} = P_D(x)$.
3. La función $P : \mathbb{R}^n \rightarrow D$ definida como $P(x) = P_D(x)$ cumple la siguiente propiedad:

$$\|P(x) - P(y)\| \leq \|x - y\|, \forall x, y \in \mathbb{R}^n$$

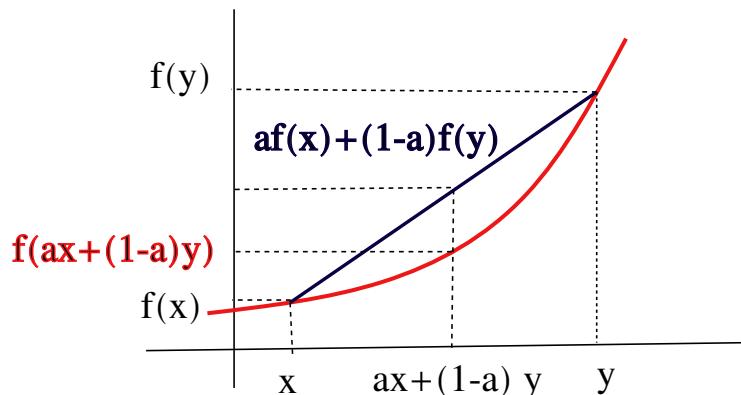
lo que implica que la proyección P es una función continua.

Definición 2.14. Sea $\Omega \subset \mathbb{R}^n$ un conjunto convexo. Un función escalar $f : \Omega \rightarrow \mathbb{R}$ es convexa en Ω si

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

para todo $x, y \in \Omega, \alpha \in [0, 1]$.

Como vimos para el caso $n = 1$, quiere decir que el segmento de línea que conecta $(x, f(x))$ con $(y, f(y))$ está por arriba del gráfico de la función. Observar la siguiente figura.



Problema de optimización convexa:

$$(PC) \quad \begin{array}{l} \text{minimizar } f(x) \\ \text{sujeto a } x \in \Omega \end{array}$$

siendo f una función escalar convexa en Ω y Ω un conjunto convexo.

El siguiente teorema muestra que cualquier solución local del problema de optimización convexa (PC) es una solución global:

Teorema 2.15. *Si x^* es un minimizador local de (PC) entonces x^* es un minimizador global de f .*

Queda como ejercicio demostrar el teorema anterior.

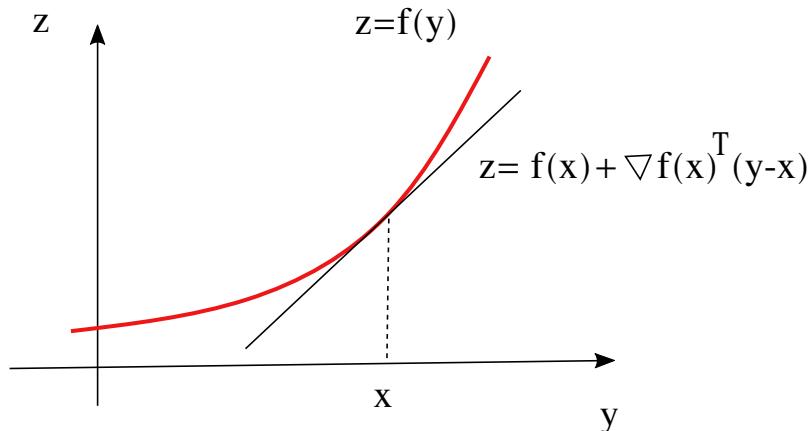
Para un problema general de optimización puede ser igual de difícil determinar si el problema es convexo (utilizando la definición de función convexa) como encontrar los minimizadores globales. Sin embargo hay problemas particulares importantes, como son los problemas lineales, donde la convexidad está fácilmente garantizada.

Cuando hay diferenciabilidad, la convexidad de una función escalar se puede caracterizar de forma más simple.

Proposición 2.2. *Si $f \in C^1(\Omega)$, Ω convexo. Entonces*

$$f \text{ es convexa en } \Omega \text{ si y solo si } f(y) \geq f(x) + \nabla f(x)^T(y - x), \forall x, y \in \Omega.$$

Queda como ejercicio para el lector demostrar la proposición anterior. Observar la siguiente figura.



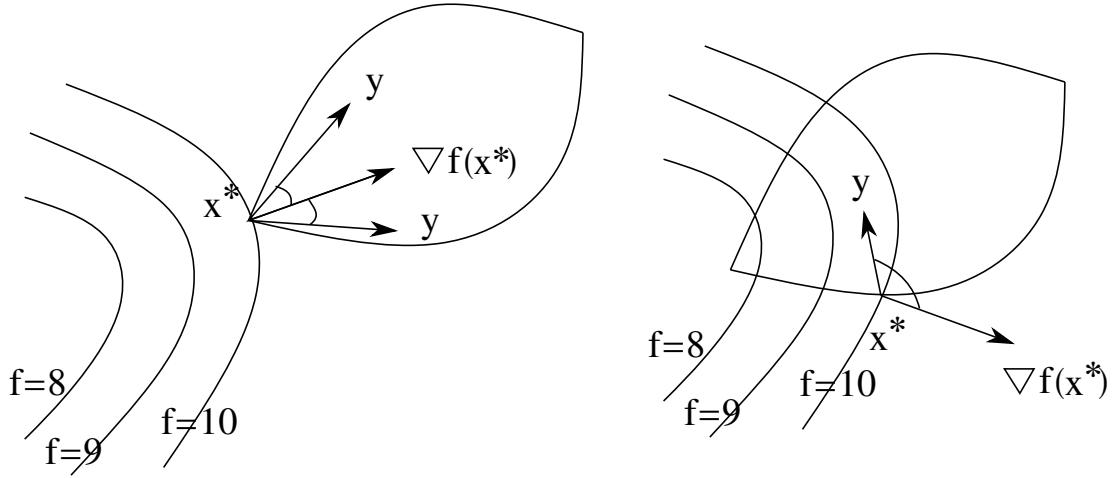
f queda por arriba de sus rectas tangentes.

Proposición 2.3. *Sea $f \in C^1(\Omega)$ una función convexa y Ω convexo. Si $x^* \in \Omega$ y $\nabla f(x^*)^T(y - x^*) \geq 0 \forall y \in \Omega$ entonces x^* es un minimizador global de f en Ω .*

En particular, todo punto estacionario de (PC) es minimizador global de (PC).

Demostración.

Si $\nabla f(x^*)^T(y - x^*) \geq 0 \forall y \in \Omega$ entonces por la propiedad anterior se tiene que $f(y) \geq f(x^*) \forall y \in \Omega$. Entonces x^* es minimizador global de f en Ω . \square



Para funciones escalares que tienen además derivadas hasta orden 2 continuas se tiene la siguiente caracterización muy útil para funciones convexas:

Proposición 2.4. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2$ y $\Omega \subset \mathbb{R}^n$ convexo. Entonces:

1. Si $\nabla^2 f(x)$ es semidefinida positiva $\forall x \in \Omega$ entonces f es convexa en Ω .
2. Si f es convexa en Ω e $\text{Int}(\Omega) \neq \emptyset$ entonces $\nabla^2 f(x)$ es semidefinida positiva $\forall x \in \Omega$.

Demostración.

1. Dados $x \in \Omega, d \in \mathbb{R}^n$ tal que $x + d \in \Omega$. Por el desarrollo de Taylor tenemos que

$$f(x + d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x + \xi d) d$$

para $\xi \in (0, 1)$. Como $\nabla^2 f(x + \xi d)$ es semidefinida positiva se tiene que

$$f(x + d) \geq f(x) + \nabla f(x)^T d$$

luego, usando la proposición anterior, f es convexa.

2. Sea $x \in \text{Int}(\Omega)$. Sea $d \in \mathbb{R}^n$ tal que $x + td \in \Omega$ para t suficientemente pequeño. Entonces, por la convexidad de f

$$0 \leq f(x + td) - f(x) - t \nabla f(x)^T d = t^2 / 2 d^T \nabla^2 f(x) d + r(t)$$

$$\text{con } \lim_{t \rightarrow 0} \frac{r(t)}{t^2} = 0.$$

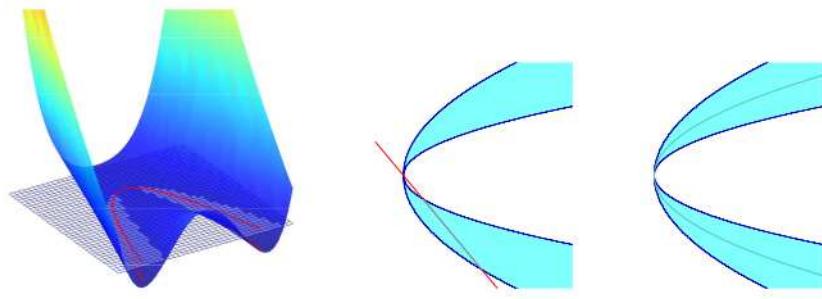
Dividiendo por t^2 y tomando límite cuando t tiende a 0 se tiene que

$$1/2 d^T \nabla^2 f(x) d \geq 0.$$

□

2.5 Ejercicios.

1. Dada la función $f(x, y) = (x - y^2)(x - \frac{1}{2}y^2)$. Mostrar que $(0, 0)$ es el único punto estacionario que NO es minimizador sin embargo $\forall d \in \mathbb{R}^n : f(x^*) \leq f(x^* + td)$ para todo t pequeño.



2. Sea $f(x_1, x_2) = (x_2 - x_1^2)^2 + x_1^5$. Mostrar que f tiene un único punto estacionario que no es maximizador ni minimizador de f .
3. Demostrar el Teorema 2.15.
4. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en Ω convexo. Probar que f es convexa si y solo si
 - (a) $f(y) \geq f(x) + (y - x)^T \nabla f(x)$ $\forall x, y \in \Omega$.
 - (b) $\nabla f(x)^T (y - x) \leq \nabla f(y)^T (y - x)$ $\forall x, y \in \Omega$.

Para (b):

\Rightarrow usar (a) con x ($f(x) \geq f(y) + \dots$) y con y ($f(y) \geq f(x) + \dots$) y sumar las dos desigualdades.
 \Leftarrow tratar de demostrar que f cumple (a). Para eso, usar el teorema del valor medio:

$$f(y) - f(x) = \nabla f(\xi)^T (y - x)$$

con $\xi = \lambda x + (1 - \lambda)y$ y usar la hipótesis con x y ξ .

5. Demostrar la Proposición 2.1.

Idea para la demostración:

1. Tomar un punto arbitrario $z \in D$ y considerar el punto $y = (1 - t)P_D(x) + tz \in D$. Usando que $\|x - P_D(x)\|^2 \leq \|x - y\|^2$ mostrar que $2(x - P_D(x))^T(z - P_D(x)) \leq t\|P_D(x) - z\|^2$ y tomar límite cuando $t \rightarrow 0$.
2. Demostrar que si, $\bar{z} \in D$ entonces

$$\|x - \bar{z}\|^2 - \|x - y\|^2 = (y - \bar{z})^T(2(x - \bar{z}) - (y - \bar{z})) \leq 0.$$

3. Usar (1.) para:

- x y $P_D(x)$ con $P_D(y)$,
- y y $P_D(y)$ con $P_D(x)$,

luego sumar las desigualdades, reordenar y usar la desigualdad de Cauchy-Schwartz.

3 Metodologías numéricas

Presentamos ahora algunos métodos o procesos algorítmicos para resolver problemas de optimización sin restricciones. Los métodos presentados generan, a partir de un punto inicial una sucesión de puntos $\{x_k\}$ según el siguiente paso iterativo:

$$x_{k+1} = x_k + t_k d_k$$

donde d_k es una dirección llamada de descenso para f en x_k y t_k es una longitud de paso en la dirección que hace bajar el valor de f .

Las diferentes opciones que existen para la elección de d_k y de t_k satisfaciendo algunas condiciones que mencionaremos generan diferentes métodos, llamados *métodos de direcciones de descenso*. En principio la elección de d_k y de t_k se realiza utilizando información sobre la función y su gradiente en un entorno de x_k .

Definidos los diferentes métodos de descenso es necesario estudiar las propiedades de convergencia de los mismos. El mejor resultado de convergencia posible para estos métodos consiste en probar que la sucesión generada converge a un minimizador global de f . Este tipo de resultados son difíciles de ser obtenidos en forma general, ya que valen con seguridad en el caso en que f es convexa y tiene minimizadores. En el caso no convexo lo mejor que se puede establecer (bajo ciertas hipótesis) es que la sucesión tiene puntos de acumulación y que todos ellos son puntos estacionarios (condición necesaria de primer orden).

3.1 Direcciones de descenso

Vamos a considerar el problema de minimización sin restricciones

$$\begin{aligned} (\text{PI}) \quad & \text{Minimizar } f(x) \\ & \text{sujeto a } x \in \mathbb{R}^n \end{aligned} \tag{3}$$

con la hipótesis de que f tiene derivadas parciales continuas (C^1).

Por el teorema que nos da condiciones necesarias de primer orden sabemos que si $\nabla f(x) \neq 0$ entonces x no es minimizador local de f y existe $z \in \mathbb{R}^n$ tal que $f(z) \leq f(x)$. Luego, para buscar un minimizador local, la estrategia más natural es, dado $x_k \in \mathbb{R}^n$, buscar x_{k+1} tal que $f(x_{k+1}) < f(x_k)$. Una manera de realizar esto es tomar una dirección d_k a partir de x_k en la que sepamos que f desciende y moverse un cierto paso t_k en esa dirección para obtener x_{k+1} de la forma $x_{k+1} = x_k + t_k d_k$ de modo que $f(x_{k+1}) < f(x_k)$. Los métodos que realizan este procedimiento sucesivamente se denominan *métodos de descenso*.

Definición 3.1. Una dirección $d \in \mathbb{R}^n$ es una dirección de descenso de f en x si existe $\varepsilon > 0$ tal que $f(x + td) < f(x), \forall t \in (0, \varepsilon]$.

En la dirección d se obtiene un decrecimiento de f , si $t > \varepsilon$ no hay más decrecimiento.

Lema 3.1. Sea f diferenciable en x .

1. Si $\nabla f(x)^T d < 0$ entonces d es una dirección de descenso de f en x .
2. Si d es una dirección de descenso de f en x entonces $\nabla f(x)^T d \leq 0$.

Demostración.

Para la parte 1., como f es diferenciable, tenemos que

$$\nabla f(x)^T d = \lim_{t \rightarrow 0} \frac{f(x + td) - f(x)}{t} < 0.$$

Entonces, para $t > 0$ suficientemente pequeño se tiene que $f(x + td) < f(x)$ como queríamos probar.

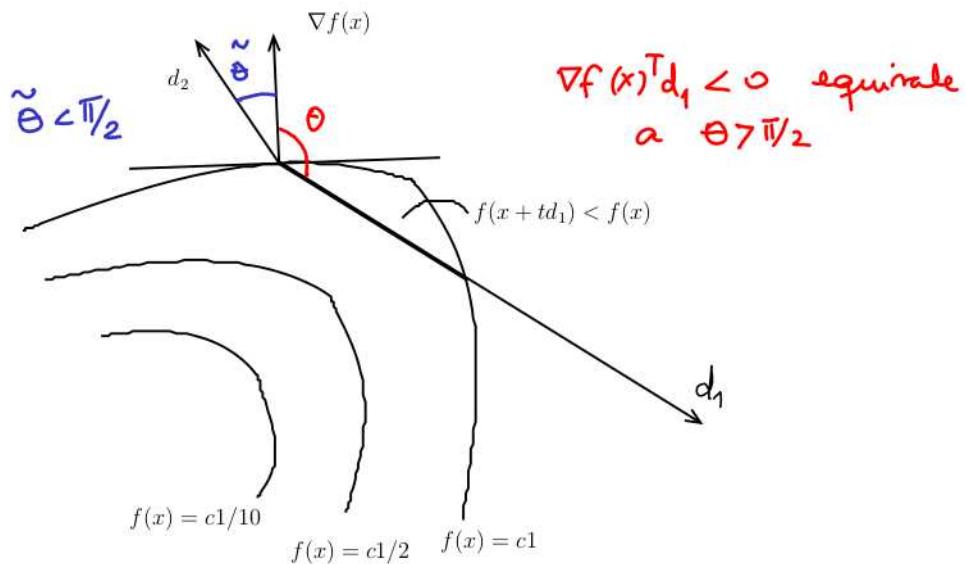
Para la parte 2., tenemos que

$$f(x + td) - f(x) = \nabla f(x)^T dt + o(t) < 0$$

luego, para $t > 0$,

$$\frac{f(x + td) - f(x)}{t} = \nabla f(x)^T d + \frac{o(t)}{t} < 0.$$

Haciendo $t \rightarrow 0^+$ obtenemos que $\nabla f(x)^T d \leq 0$. \square

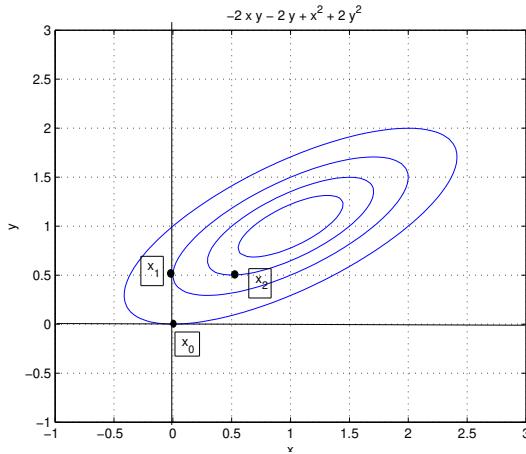


En el ejemplo, podemos ver las curvas de nivel de cierta función f . d_1 es dirección de descenso para f a partir de x pero d_2 no lo es.

Observación. Por un lado, si $\nabla f(x) \neq 0$ entonces el producto $\nabla f(x)^T d < 0$ quiere decir que en esa dirección podemos encontrar puntos donde el valor funcional de f es menor que en $f(x)$.

Por otro lado, $\nabla f(x)^T d < 0$ quiere decir que el ángulo que se forma entre d y $\nabla f(x)$ es mayor a 90° . Así, la dirección de máximo descenso se obtiene cuando el ángulo toma el valor de 180° . Esta dirección es $d = -\nabla f(x)$ y es llamada *dirección de Cauchy* o dirección de máximo descenso. Tiene la desventaja de que, cuando las curvas de nivel de la función objetivo son muy alargadas, la dirección es "casi" ortogonal a la dirección que hace llegar al minimizador rápidamente generando un proceso en zig-zag.

Ejemplo. Considerar la función $f(x, y) = -2xy - 2y + x^2 + 2y^2$. Mostrar que tiene un minimizador global en $(1, 1)$. Observar que $d = (d_1, d_2)$ es una dirección de descenso de f a partir de $(0, 0)$ si y solo si $d_2 > 0$. Ver las curvas de nivel de f :



El esquema iterativo o modelo de algoritmo general basado en direcciones de descenso consiste en los siguientes pasos:

Algoritmo general. Dado $x_0 \in \mathbb{R}^n$ inicial, en cada paso k :

Paso 1. Definir una dirección d_k de descenso para f en x_k ;

Paso 2. Determinar $t_k > 0$ (tamaño del paso) tal que $f(x_k + t_k d_k) < f(x_k)$;

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$;

Paso 4. Determinar si x_{k+1} es solución.

Como las direcciones son de descenso, siempre es posible encontrar una sucesión de pasos t_k en las condiciones planteadas anteriormente. En estas condiciones, la sucesión de valores funcionales $f(x_k)$ es decreciente, esto quiere decir que, en los diferentes pasos iterativos los puntos x_k se encuentran en curvas de nivel asociadas a valores cada vez menores. Esto justifica la esperanza de que la sucesión x_k pueda converger a una solución (aunque en general, solo podremos garantizar convergencia a puntos estacionarios.)

En lo que sigue de esta sección intentaremos definir los pasos que están en abierto en el esquema general.

El punto inicial x_0 debe ser un dato del problema.

[Paso 4.] El criterio de finalización para determinar si un nuevo punto es solución puede basarse en la pregunta de si $\|\nabla f(x_{k+1})\| \leq \varepsilon$ para algún parámetro ε inicial y fijo.

Nos resta estudiar cómo calcular posibles direcciones d_k **[Paso 1.]** y cómo calcular las longitudes de paso t_k **[Paso 2.]**.

Supongamos que d_k es una dirección de descenso de f a partir de x_k , existen diferentes reglas para determinar t_k , estas reglas se denominan reglas de búsqueda lineal.

3.2 Búsqueda lineal en una dirección

- Búsqueda lineal exacta.

Esta estrategia consiste en buscar el minimizador de f en la semirrecta $x_k + td_k$ para $t \geq 0$, es decir:

$$t_k = \operatorname{argmin}_{t \geq 0} f(x_k + td_k).$$

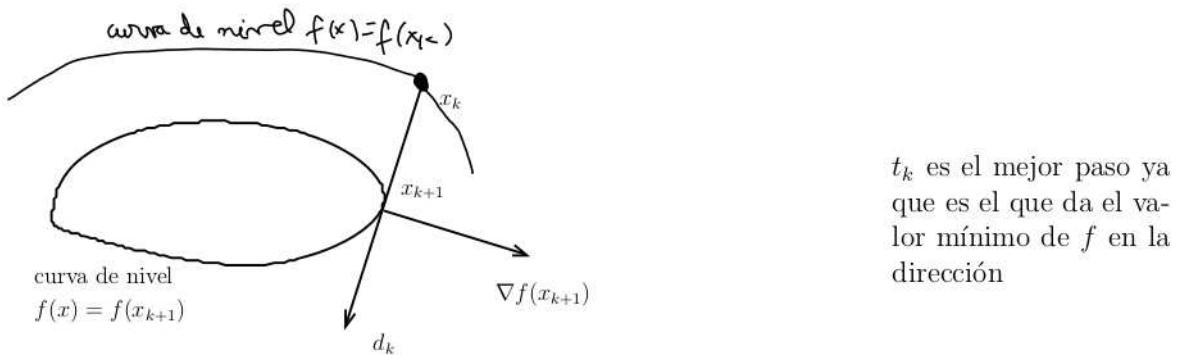
Definimos $\varphi(t) = f(x_k + td_k)$. Sabemos que, si f es diferenciable entonces, en la dirección d_k se cumple que

$$\varphi'(t_k) = 0.$$

Luego,

$$\nabla f(x_{k+1})^T d_k = 0.$$

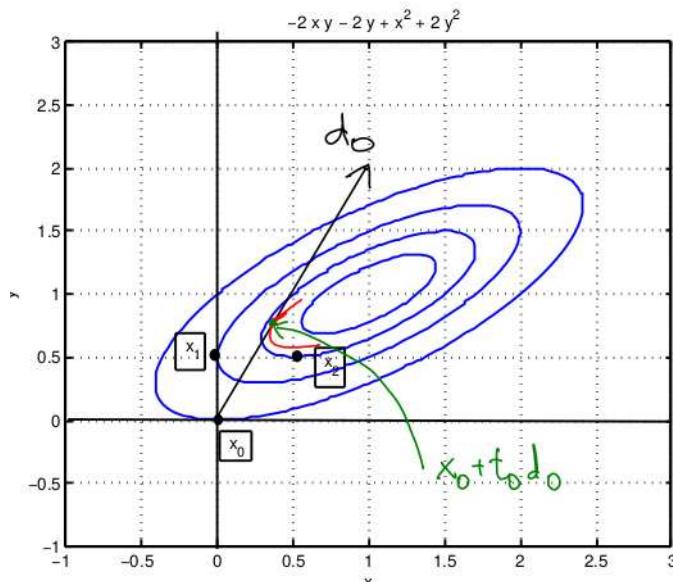
Si $\nabla f(x_{k+1}) \neq 0$ entonces x_{k+1} es el punto intersección de la semirrecta que parte de x_k en la dirección d_k con la curva de nivel de f que pasa por x_{k+1} . Si bien es la longitud de paso óptimo puede ser muy difícil de calcular en forma exacta.



Veamos la expresión de esta longitud de paso en el caso en que f es una función cuadrática. Sea $f(x) = \frac{1}{2}x^T Ax + b^T x$, A simétrica. Dada una dirección d , buscamos que valor t tal que $\nabla f(x+td)^T d = 0$. Como $\nabla f(x) = Ax + b$ tenemos que $(A(x + td) + b)^T d = 0$. Entonces $x^T Ad + td^T Ad + b^T d = 0$, luego $t = -\frac{\nabla f(x)^T d}{d^T Ad}$.

Ejemplo. Dada la función $f(x, y) = -2xy - 2y + x^2 + 2y^2$. Mostrar que $d_0 = (1, 2)$ es una dirección de descenso para f a partir de $x_0 = (0, 0)$ y calcular t_0 mediante búsqueda exacta.

A partir de x_0 se obtiene que $\varphi(t) = f(x_0 + td_0) = f(t, 2t) = -4t + 5t^2$ y tiene su minimizador en $t = \frac{2}{5}$. Luego, $x_1 = x_0 + td_0 = (\frac{2}{5}, \frac{4}{5})$.



- Regla de Armijo

Esta búsqueda es inexacta y se basa en calcular una longitud de paso que de un “descenso suficiente” de f en relación al valor $f(x_k)$.

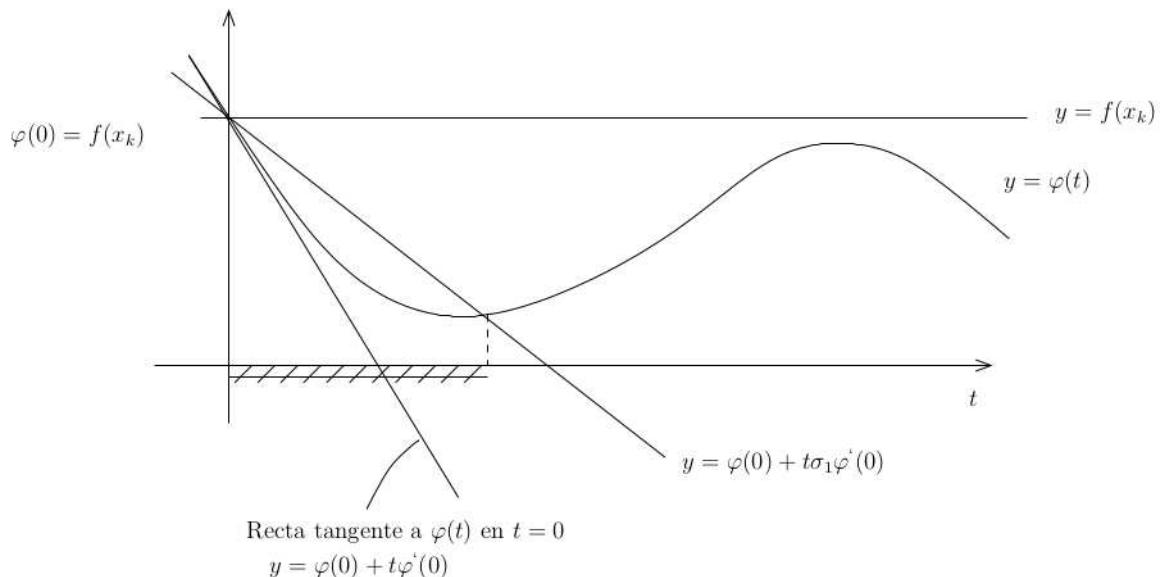
Consideramos f diferenciable, x_k, d_k dados y fijos. Definimos $\varphi(t) = f(x_k + td_k)$. Para $\sigma_1 \in (0, 1)$ fijo, consideramos los t tales que

$$\varphi(t) \leq \varphi(0) + \sigma_1 t \varphi'(0).$$

Luego, en términos de f se tiene la denominada *condición de Armijo*:

$$f(x_k + td_k) \leq f(x_k) + \sigma_1 t \nabla f(x_k)^T d_k. \quad (4)$$

La reducción que exige esta condición es proporcional a la longitud del paso t_k y al valor de la derivada direccional.



Como $\varphi'(0) < 0$ se tiene que $\varphi'(0) < \sigma_1 \varphi'(0)$ luego la pendiente $\sigma_1 \varphi'(0)$ genera una recta con menor inclinación que la recta $y = \varphi(0) + t \varphi'(0)$.

Lema 3.2. Sea f una función C^1 . Dados \bar{x}, d tales que $\nabla f(\bar{x})^T d < 0, \sigma_1 \in (0, 1)$ fijo entonces existe $\varepsilon = \varepsilon(\sigma_1)$ tal que (4) se cumple para todo $t \in (0, \varepsilon]$.

Demostración.

$$\text{Como } \nabla f(\bar{x})^T d = \lim_{t \rightarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t} < 0 \text{ entonces } \lim_{t \rightarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t \nabla f(\bar{x})^T d} = 1.$$

Por definición de límite, tenemos que existe $\varepsilon > 0$ tal que, para todo $t \in (0, \varepsilon]$

$$\frac{f(\bar{x} + td) - f(\bar{x})}{t \nabla f(\bar{x})^T d} \geq \sigma_1$$

como queríamos probar. \square

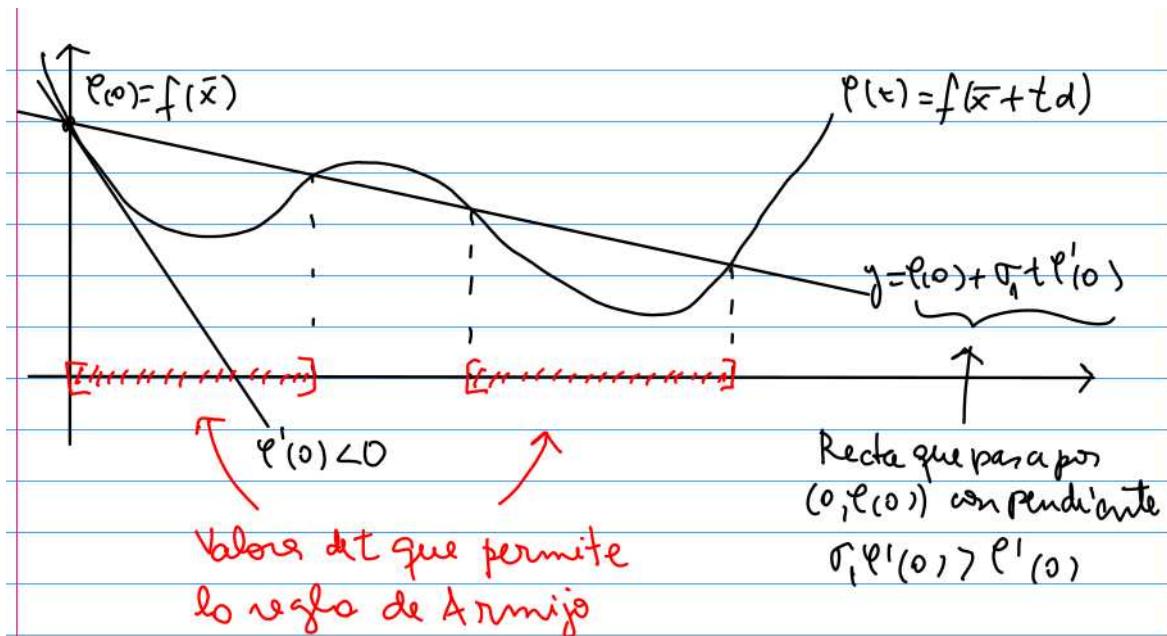
Ejemplo. Dada la función $f(x, y) = -2xy - 2y + x^2 + 2y^2$. Determinar el intervalo máximo de paso que se puede dar a partir de $x_0 = (0, 0)$ en la dirección $d_0 = (1, 2)$ con la regla de Armijo para $\sigma_1 = 1/2$.

La desigualdad (4) se traduce en:

$$-4t + 5t^2 \leq 0 + t/2(-4).$$

Luego se obtiene $t \in [0, \frac{2}{5}]$.

Observación: se puede dar la siguiente situación:



En general, en un proceso algorítmico que considere la búsqueda de Armijo, se procede de la siguiente manera:

1. Comenzar con $t = 1$.
2. Si $f(x_k + td_k) \leq f(x_k) + \sigma_1 t \nabla f(x_k)^T d_k$ entonces definir $t_k = t$;
3. si no, reducir t con algún criterio hasta que se cumpla la condición.

Posibles maneras de reducir t :

1. hacer $t = t/2$;
2. calcular $t \in [0.1t, 0.9t]$. (puede ser el punto medio del intervalo)

En el ejemplo de la función $f(x, y) = -2xy - 2y + x^2 + 2y^2$ se tiene que, para $\sigma_1 = 1/2$:

- $t = 1$ no vale,
- $t = 1/2$ no vale,
- $t = 1/4$ vale.

Si bien esta búsqueda inexacta es la más popular, puede no asegurar que se haga un progreso razonable ya que se pueden tomar pasos (valores de t_k) demasiado chicos lo que puede generar mucho esfuerzo para cambiar a un punto muy cercano al actual. Luego, en algunas implementaciones se considera la *Regla de Wolfe*.

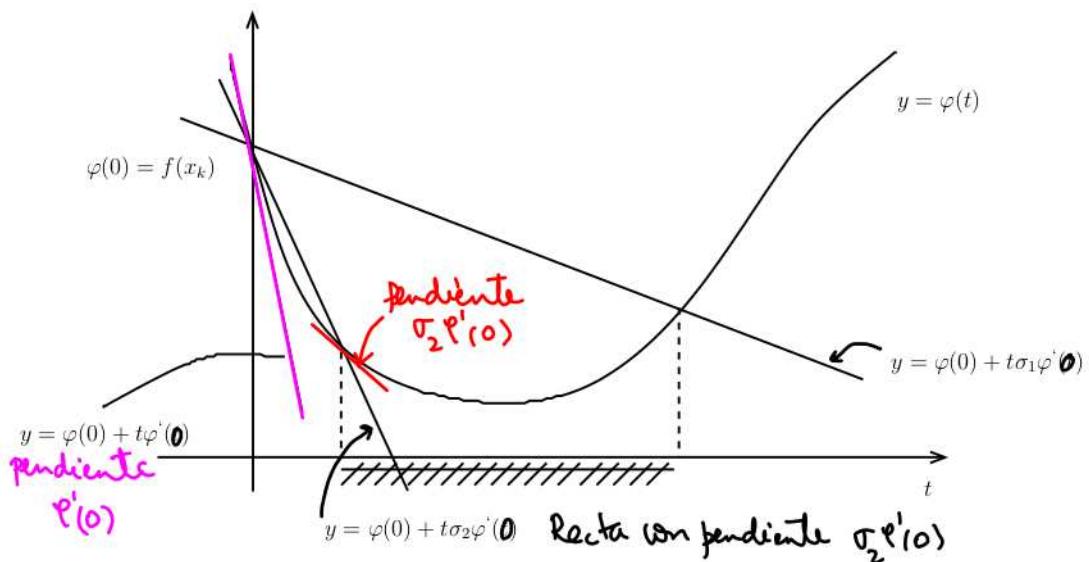
- Regla de Wolfe

Esta condición, además de pedir la condición de Armijo, exige otra condición para evitar valores de t_k muy pequeños, es decir, garantiza mayor desplazamiento, la posibilidad de dar pasos más largos.

Se pide que el valor de t deseado cumpla además la condición $\varphi'(t) \geq \sigma_2 \varphi'(0)$. En términos de f esta condición equivale a

$$\nabla f(x_k + td_k)^T d_k \geq \sigma_2 \nabla f(x_k)^T d_k. \quad (5)$$

Es decir, la pendiente de la recta tangente a $\varphi(t)$ en el nuevo t debe ser mayor o igual a una proporción de la pendiente a $\varphi(t)$ en $t = 0$. En general se consideran valores $0 < \sigma_1 < \sigma_2 < 1$ con $\sigma_1 < 1/2$.



Se considera la pendiente $\sigma_2 \varphi'(0)$ con $0 < \sigma_2 < 1$. Como $\varphi'(0) < 0$ se tiene que $\varphi'(0) < \sigma_2 \varphi'(0)$ luego la pendiente $\sigma_2 \varphi'(0)$ genera una recta con menor inclinación que la recta $y = \varphi(0) + t\varphi'(t)$.

Luego, cuando hablamos de *Regla de Wolfe* nos referimos a las condiciones (4) y (5) juntas.

Proposición 3.1. Sea f una función \mathcal{C}^1 y acotada inferiormente en la recta $x_k + td_k$ siendo d_k tal que $\nabla f(x_k)^T d_k < 0$ y $0 < \sigma_1 < \sigma_2 < 1, \sigma_1 < 1/2$. Entonces existe t_k que verifica la regla de Wolfe.

Demostración.

Como la función $\varphi(t) = f(x_k + td_k)$ es acotada inferiormente existe t_f el primer valor para el cual la gráfica de $\varphi(t)$ intersecta a la recta $y = f(x_k) + t\sigma_1 \nabla f(x_k)^T d_k$. Luego, $f(x_k + t_f d_k) = f(x_k) + t_f \sigma_1 \nabla f(x_k)^T d_k$.

Usando el Teorema del Valor Medio tenemos que existe t_c entre 0 y t_f tal que

$$f(x_k + t_f d_k) - f(x_k) = t_f \nabla f(x_k + t_c d_k)^T d_k$$

y entonces

$$\nabla f(x_k + t_c d_k)^T d_k = \sigma_1 \nabla f(x_k)^T d_k > \sigma_2 \nabla f(x_k)^T d_k.$$

Luego, existe $t_c < t_f$ tal que

$$\nabla f(x_k + t_c d_k)^T d_k > \sigma_2 \nabla f(x_k)^T d_k$$

y además, por definición de t_f tenemos que t_c cumple también la condición de Armijo. \square

Procedimiento práctico para verificar la regla de Wolfe.

Consideramos fijos $d_k, \nabla f(x_k), x_k, 0 < \sigma_1 < \sigma_2 < 1, \sigma_1 < 1/2$. Definimos un intervalo $[t_{min}, t_{max}]$. Tomamos un $t_{inicial} \in (t_{min}, t_{max})$.

En un paso iterativo iterativo hacemos lo siguiente:

1. Si t cumple (4) entonces ir a [2].

Si no vale (4), definir $t_{max} = t$ y calcular un $t_{nuevo} \in (t_{min}, t_{max})$ e ir a [1.] con $t = t_{nuevo}$.

2. Si t cumple (5) terminar.

Si no vale (5), definir $t_{min} = t$ y calcular un $t_{nuevo} \in (t_{min}, t_{max})$ e ir a [1.] con $t = t_{nuevo}$.

El no cumplimiento de (4) significa que el actual valor de t es demasiado grande, por esta razón se modifica el t_{max} . El no cumplimiento de (5) significa que el actual valor de t es demasiado pequeño, por esta razón se modifica el t_{min} .

3.3 Convergencia

Recordemos el esquema que planteamos como Algoritmo general:

Algoritmo general. Dado $x_0 \in \mathbb{R}^n$ inicial, en cada paso k :

Paso 1. Definir una dirección d_k de descenso para f en x_k ;

Paso 2. Determinar $t_k > 0$ (tamaño del paso) tal que $f(x_k + t_k d_k) < f(x_k)$;

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$;

Paso 4. Determinar si x_{k+1} es solución.

Supongamos que en el paso 4 consideramos:

Paso 4. Si $\nabla f(x_{k+1}) = 0$ parar.

En ese caso podemos decir que el proceso encuentra un punto estacionario en un número finito de pasos o bien genera una sucesión infinita de iterados de modo que:

$$f(x_{k+1}) < f(x_k), \forall k.$$

El último caso es el que nos interesa estudiar en esta sección. Nos preguntamos cómo saber si la sucesión tiene algún punto de acumulación y, si lo tiene, si el punto límite es estacionario.

Definición 3.2. *Un algoritmo se dice globalmente convergente cuando para cualquier sucesión $\{x_k\}$ generada por el algoritmo y para cualquier punto de acumulación \bar{x} de la sucesión se tiene que \bar{x} es un punto estacionario de (PI).*

Para garantizar la convergencia global de un algoritmo es importante la elección de la sucesión de direcciones d_k y de longitudes de paso t_k .

Consideramos una función continua $H : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ que asigna a cada $x \in \mathbb{R}^n$ una matriz definida positiva $H(x)$. Luego, si $\nabla f(x) \neq 0$ tenemos que la dirección:

$$d = -H(x)\nabla f(x)$$

es una dirección de descenso para f en x .

Consideramos el siguiente algoritmo:

Algoritmo de descenso A. Dado $x_0 \in \mathbb{R}^n$ inicial. Mientras $\nabla f(x_k) \neq 0$ repetir:

Paso 1. Calcular $d_k = -H(x_k)\nabla f(x_k)$.

Paso 2. Calcular $t_k > 0$ tal que $f(x_k + t_k d_k) < f(x_k)$;

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$;

Paso 4. Hacer $k = k + 1$ y repetir.

Convergencia del algoritmo de descenso A con búsqueda lineal exacta

Teorema 3.3. *Sea f una función C^1 . El algoritmo de descenso A con paso t_k calculado mediante búsqueda exacta es globalmente convergente.*

Demostración.

Sea $\{x_k\}$ una sucesión generada por el algoritmo y sea \bar{x} un punto de acumulación de la sucesión. Entonces, existe $K \subset \mathbb{N}$ tal que $x_k \rightarrow_{k \in K} \bar{x}$. Por continuidad de f se tiene que $f(x_k) \rightarrow_{k \in K} f(\bar{x})$ y como $\{f(x_k)\}_{k \in \mathbb{N}}$ es monótona decreciente se tiene que $\lim_{k \rightarrow \infty} f(x_k) = f(\bar{x})$ (Una sucesión monótona que tiene una subsucesión convergente, converge).

Suponemos por el absurdo que \bar{x} no es estacionario, entonces $\nabla f(\bar{x}) \neq 0$. Entonces $d = -H(\bar{x})\nabla f(\bar{x})$ es una dirección de descenso y existe $\bar{t} > 0$ tal que $f(\bar{x} + \bar{t}d) < f(\bar{x})$, es decir, $\beta = f(\bar{x}) - f(\bar{x} + \bar{t}d) > 0$.

Sea $h : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $h(x) = f(x) - f(x - \bar{t}H(x)\nabla f(x))$. Por la continuidad de h se tiene que

$$h(x_k) \rightarrow_{k \in K} h(\bar{x}) = \beta > 0$$

por lo tanto $h(x_k) \geq \beta/2$ para todo $k \in K, k \geq k_0$. Entonces

$$f(x_k) - f(x_k + \bar{t}d_k) = h(x_k) \geq \beta/2.$$

Pero t_k fue elegido por búsquedas exactas entonces

$$f(x_{k+1}) = f(x_k + t_k d_k) \leq f(x_k + \bar{t} d_k) \leq f(x_k) - \beta/2.$$

Es decir,

$$f(x_k) - f(x_{k+1}) \geq \beta/2, \forall k \in K, k \geq k_0. \quad (6)$$

Absurdo porque $\lim_{k \rightarrow \infty} f(x_k) = f(\bar{x})$.

Por lo tanto $\nabla f(\bar{x}) = 0$. \square

Convergencia del algoritmo de descenso A con búsquedas de Armijo

Teorema 3.4. *Sea f una función C^1 . El algoritmo de descenso A con paso t_k calculado usando la condición de Armijo es globalmente convergente.*

Demostración.

Sea $\{x_k\}$ una sucesión generada por el algoritmo y sea \bar{x} un punto de acumulación de la sucesión. Entonces, existe $K \subset \mathbb{N}$ tal que $x_k \rightarrow_{k \in K} \bar{x}$.

Nuevamente, por la continuidad de f se tiene que $f(x_k) \rightarrow_{k \in K} f(\bar{x})$ y como $\{f(x_k)\}_{k \in \mathbb{N}}$ es monótona decreciente se tiene que $\lim_{k \rightarrow \infty} f(x_k) = f(\bar{x})$.

Suponemos que \bar{x} no es estacionario, entonces $\nabla f(\bar{x}) \neq 0$.

Por la condición de Armijo se tiene que

$$f(x_{k+1}) = f(x_k + t_k d_k) \leq f(x_k) + \sigma_1 t_k \nabla f(x_k)^T d_k$$

entonces

$$f(x_k) - f(x_{k+1}) \geq -\sigma_1 t_k \nabla f(x_k)^T (-H(x_k) \nabla f(x_k)) = \sigma_1 t_k \nabla f(x_k)^T H(x_k) \nabla f(x_k) \geq 0.$$

Por lo tanto, como $f(x_k) - f(x_{k+1}) \rightarrow 0$ debe ser

$$t_k \nabla f(x_k)^T H(x_k) \nabla f(x_k) \rightarrow_{k \in K} 0.$$

Pero

$$\nabla f(x_k)^T H(x_k) \nabla f(x_k) \rightarrow_{k \in K} \nabla f(\bar{x})^T H(\bar{x}) \nabla f(\bar{x}). \quad (7)$$

Entonces $\lim_{k \in K} t_k = 0$.

Para cada $k \in K$, el valor inicial de la longitud de paso se tuvo que reducir en Armijo entonces existe $t'_k > 0$ para el cual Armijo no vale ($t_k = t'_k/2$). Como $t_k \rightarrow 0$, $t'_k \rightarrow 0$ y para t'_k Armijo no vale:

$$f(x_k + t'_k d_k) > f(x_k) + \sigma_1 t'_k \nabla f(x_k)^T d_k.$$

Como la función $\xi(t) = f(x_k + td_k) - f(x_k) - \sigma_1 t \nabla f(x_k)^T d_k$ es continua y

$$\xi(t_k) = f(x_k + t_k d_k) - f(x_k) - \sigma_1 t_k \nabla f(x_k)^T d_k < 0$$

$$\xi(t'_k) = f(x_k + t'_k d_k) - f(x_k) - \sigma_1 t'_k \nabla f(x_k)^T d_k > 0,$$

por el Teorema del valor intermedio existe $s_k \in [t_k, t'_k] : \xi(s_k) = 0$. Es decir

$$f(x_k + s_k d_k) - f(x_k) - \sigma_1 s_k \nabla f(x_k)^T d_k = 0,$$

entonces

$$f(x_k + s_k d_k) - f(x_k) = \sigma_1 s_k \nabla f(x_k)^T d_k.$$

Pero, por TVM, existe $\theta_k \in (0, 1)$ tal que

$$f(x_k + s_k d_k) - f(x_k) = s_k \nabla f(x_k + \theta_k s_k d_k)^T d_k$$

entonces

$$s_k \nabla f(x_k + \theta_k s_k d_k)^T d_k = \sigma_1 s_k \nabla f(x_k)^T d_k.$$

Por lo tanto

$$\nabla f(x_k + \theta_k s_k d_k)^T H(x_k) \nabla f(x_k) = \sigma_1 \nabla f(x_k)^T H(x_k) \nabla f(x_k).$$

Tomando límite para $k \in K$ y usando que $s_k \rightarrow 0$ porque $s_k \in [t_k, t'_k]$ y $t_k, t'_k \rightarrow 0$ se tiene que

$$\nabla f(\bar{x})^T H(\bar{x}) \nabla f(\bar{x}) = \sigma_1 \nabla f(\bar{x})^T H(\bar{x}) \nabla f(\bar{x}).$$

Esto es absurdo. Por lo tanto debe ser $\nabla f(\bar{x}) = 0$. \square

Convergencia del algoritmo de descenso A con búsqueda de Wolfe

Para el análisis de la convergencia utilizando la regla de Wolfe se debe tener en cuenta la calidad de la dirección de descenso. Esta medida de calidad se define en función del ángulo entre la dirección y el opuesto del gradiente. Es necesario considerar el siguiente teorema:

Teorema 3.5. *Sea f una función C^1 y acotada inferiormente en el conjunto $\mathcal{A} = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$. Suponemos que se genera una sucesión $\{x_k\}$ cumpliendo las condiciones (4) y (5) usando d_k de descenso. Si $\nabla f(x)$ es Lipschitz continuo en \mathcal{A} con constante L entonces*

$$\sum \cos^2(\theta_k) \|\nabla f(x_k)\|^2 < \infty \quad (8)$$

siendo θ_k el ángulo que forman d_k y $-\nabla f(x_k)$, es decir: $\cos(\theta_k) = \frac{-\nabla f(x_k)^T d_k}{\|\nabla f(x_k)\| \|d_k\|}$.

Demostración.

Recordemos que $\nabla f(x)$ es Lipschitz continuo con constante $L > 0$ entonces $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \forall x, y \in \mathbb{R}^n$.

Luego, usando esta condición, tenemos que

$$(\nabla f(x_{k+1}) - \nabla f(x_k))^T d_k \leq \|\nabla f(x_{k+1}) - \nabla f(x_k)\| \|d_k\| \leq Lt_k \|d_k\|^2.$$

Por otro lado, por la condición de Wolfe (5) tenemos que

$$(\nabla f(x_{k+1}) - \nabla f(x_k))^T d_k \geq (\sigma_2 - 1) \nabla f(x_k)^T d_k.$$

Combinando las dos desigualdades:

$$(\sigma_2 - 1) \nabla f(x_k)^T d_k \leq Lt_k \|d_k\|^2$$

y se tiene que

$$t_k \geq \left(\frac{\sigma_2 - 1}{L} \right) \frac{\nabla f(x_k)^T d_k}{\|d_k\|^2}.$$

Usando la primer condición de Wolfe y la definición de $\cos(\theta_k)$ se obtiene que

$$f(x_{k+1}) \leq f(x_k) + \sigma_1 \left(\frac{\sigma_2 - 1}{L} \right) \frac{(\nabla f(x_k)^T d_k)^2}{\|d_k\|^2} = f(x_k) + c(\cos(\theta_k))^2 \|\nabla f(x_k)\|^2$$

donde $c = \sigma_1 \left(\frac{\sigma_2 - 1}{L} \right)$.

Sumando desde 0 hasta k en la última desigualdad se tiene:

$$f(x_{k+1}) \leq f(x_0) + c \sum_{j=0}^k (\cos(\theta_j))^2 \|\nabla f(x_j)\|^2.$$

Como f es acotada inferiormente y $c < 0$ la sucesión de sumas parciales es acotada y se obtiene (8) como queríamos probar. \square

La condición (8) se llama *condición de Zoutendijk*.

Observemos que, si d_k es tal que $\cos(\theta_k) \geq \mu > 0$ entonces se tiene, por (8), que $\|\nabla f(x_k)\| \rightarrow 0$. Es decir, si en cada iteración k , la dirección de búsqueda no es ortogonal al gradiente en el punto entonces la sucesión de gradientes tiende a cero. Luego, si $\{x_k\}$ tiene una subsucesión convergente con punto límite x_* entonces $\nabla f(x_*) = 0$, es decir, ese punto límite es un punto estacionario de f .

Por ejemplo, si elegimos $d_k = -\nabla f(x_k)$, la dirección de máximo decrecimiento o dirección de Cauchy, se tiene un método globalmente convergente.

Si el algoritmo tiene una subsucesión de iterados $k \in K$ para los que $\cos(\theta_k) > \mu > 0$ entonces $\|\nabla f(x_k)\| \rightarrow_{k \in K} 0$. Por ejemplo, este caso aparece cuando se considera la dirección de Cauchy cada una cantidad fija de iteraciones.

Si d_k es calculada mediante una fórmula del tipo $d_k = -B_k^{-1}\nabla f(x_k)$ siendo B_k simétrica definida positiva tenemos que d_k es una dirección de descenso que además cumple la condición de que el coseno esta acotado. Se puede verificar que

$$\cos(\theta_k) = \frac{-\nabla f(x_k)^T d_k}{\|\nabla f(x_k)\| \|d_k\|} = \frac{\nabla f(x_k)^T B_k^{-1} \nabla f(x_k)}{\|\nabla f(x_k)\| \|B_k^{-1} \nabla f(x_k)\|} \geq \frac{\lambda_{\min}(B_k^{-1})}{\|B_k^{-1}\|} = \frac{1}{\|B_k\| \|B_k^{-1}\|} = \frac{1}{\mathcal{K}_{B_k}}.$$

\mathcal{K}_{B_k} es el número de condición de B_k . Luego, si este número esta acotado inferiormente se tiene una cota para el coseno lo que garantiza convergencia global si se utiliza la condición de Wolfe.

3.4 Método del gradiente o de Cauchy

Cuando consideramos el Algoritmo de descenso de la sección anterior usando $H(x) = Id$ se tiene el conocido método del gradiente o de Cauchy. Según los teoremas de la sección previa la convergencia global está asegurada cuando se considera búsqueda exacta, de Armijo o Wolfe.

Ejemplo. Realizar dos iteraciones del Método del gradiente con búsqueda exacta a partir del punto $x_0 = (0, 0)$ para la función $f(x, y) = -2xy - 2y + x^2 + 2y^2$.

Tenemos $\nabla f(x, y) = \begin{pmatrix} -2y + 2x \\ -2x - 2 + 4y \end{pmatrix}$.

$k = 0$

Para $x_0 = (0, 0)$, $d_0 = -\nabla f(0, 0) = (0 \ 2)^T$.

Luego, calculamos $\varphi(t) = f(x_0 + td_0) = -4t + 8t^2$ y tiene su minimizador en $t = \frac{1}{4}$. Luego,

$$x_1 = x_0 + td_0 = \left(0 \frac{1}{2}\right).$$

$$k = 1$$

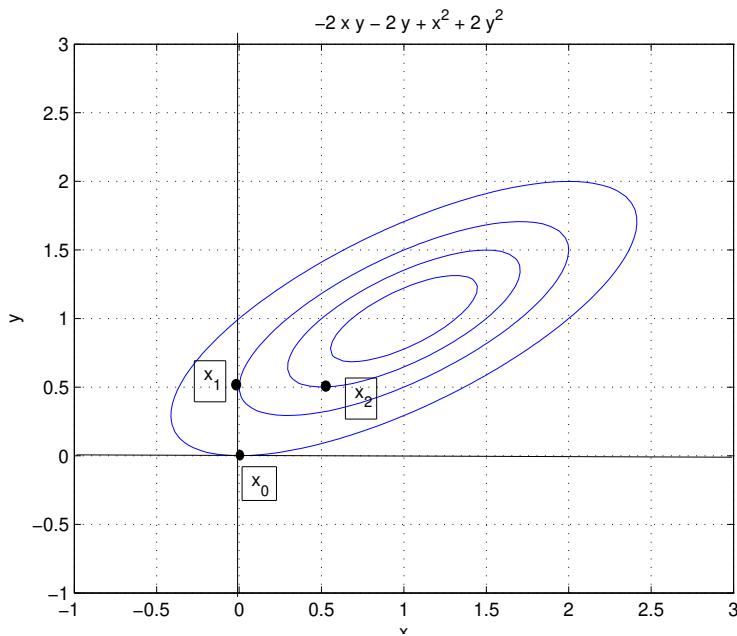
$$d_1 = -\nabla f\left(0 \frac{1}{2}\right) = (1 0)^T,$$

$\varphi(t) = f(x_1 + td_1) = -t + t^2 - 0.5$ y tiene su minimizador en $t = \frac{1}{2}$. Luego,

$$x_2 = x_1 + td_1 = \left(\frac{1}{2} \frac{1}{2}\right).$$

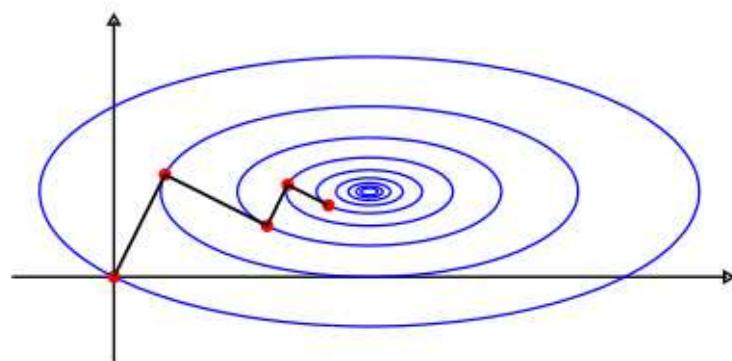
Observar los iterados obtenidos en la figura en la que se observan también las curvas de nivel de la función objetivo.

Mostrar que $x_3 = \left(\frac{1}{2} \frac{3}{4}\right)$.



Se puede ver que:

- Las direcciones que se generan son ortogonales ($d_{k+1}^T d_k = 0$).
- la velocidad de convergencia es lenta pese a que se trata de un función objetivo cuadrática convexa.



Para un estudio sobre velocidad de convergencia en el método de Cauchy ver [4].

3.5 Método de Newton

En su forma clásica, el método de Newton es un método para resolver sistemas de ecuaciones no lineales, es decir, no fue originariamente pensado como un método de minimización.

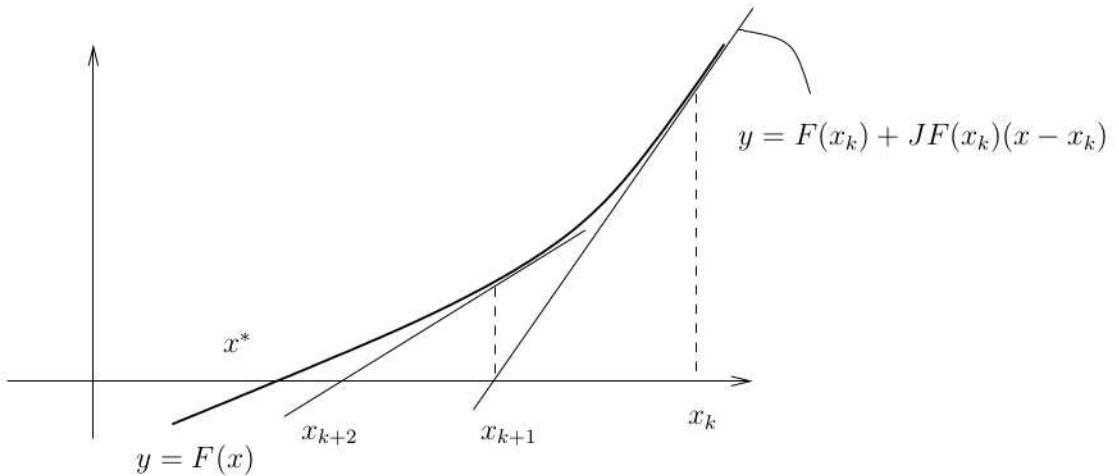
Originalmente se introdujo para resolver el problema de encontrar $x \in \mathbb{R}^n$ tal que $F(x) = 0$ donde $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Si x_k es una aproximación de la solución x^* del sistema, en un entorno de x_k podemos aproximar la ecuación $F(x) = 0$ por su linealización (polinomio de Taylor de orden 1 de F alrededor de x_k):

$$F(x_k) + JF(x_k)(x - x_k) = 0 \quad (9)$$

y buscar el punto que resuelve este sistema.

Método de Newton para sistemas de ecuaciones:

1. Dado x_0 inicial,
2. calcular x_{k+1} solución de (9),
3. hacer $k = k + 1$ y repetir.



El sistema (9) es un sistema lineal de ecuaciones con matriz $JF(x_k)$ que, para que tenga solución, debe ser $JF(x_k)$ no singular. Resolver este sistema lineal es una de las principales desventajas del método de Newton. Se puede realizar mediante factorizaciones de la matriz o usando métodos iterativos que resuelven sistemas lineales de ecuaciones.

Nos interesa ahora usar la filosofía del método de Newton para minimización irrestricta. Sea f una función \mathcal{C}^2 . La aproximación cuadrática de f en un punto x_k es

$$q_k(x) = f(x_k) + \nabla f(x_k)^T(x - x_k) + 1/2(x - x_k)^T \nabla^2 f(x_k)(x - x_k).$$

Podemos buscar la dirección d_k que nos lleve al minimizador de q_k . Sea $d = x - x_k$, luego $q_k(x) = q_k(x_k + d) = f(x_k) + \nabla f(x_k)^T d + 1/2 d^T \nabla^2 f(x_k) d$. Buscamos d_k tal que $\nabla q_k(x_k + d) = 0$, luego,

$$(\nabla^2 f(x_k))d_k = -\nabla f(x_k) \quad (10)$$

La dirección d_k solución del sistema (10) se denomina *dirección de Newton*.

Una vez obtenida la dirección de Newton, nos preguntamos si es de descenso. Infelizmente d_k puede no ser de descenso si $\nabla^2 f(x_k)$ no es definida positiva. Por ejemplo la función $f(x, y) = (1/2)(x^2 - y^2)$ en el punto $x_0 = (0, 1)$ verifica

$$\nabla f(x_0) = (0 \ -1)^T, \nabla^2 f(x_0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

En este caso la dirección de Newton es $d_0 = (0 \ -1)$ y $\nabla f(x_0)^T d_0 = 1 > 0$.

Ejemplo 2. Para la función $f(x, y) = -2xy - 2y + x^2 + 2y^2$, comenzando desde $x_0 = (0, 0)$ tenemos que $d_0 = -(\nabla^2 f(x_0))^{-1} \nabla f(x_0) = (1, 1)$ y $x_1 = x_0 + d_0 = (1, 1)$ que es la solución. Es de esperar que esto suceda ya que la función f es cuadrática.

La dirección de Newton (10) se encuentra resolviendo el sistema lineal que puede no estar bien definido si la matriz $\nabla^2 f(x_k)$ es singular. Además, nos tenemos que preguntar si, en el caso en el que el sistema tenga solución, la dirección resultante es de descenso. Observar que esto sucede siempre que $\nabla^2 f(x_k)$ es una matriz definida positiva.

Método de Newton puro. Dado $x_0 \in \mathbb{R}^n$ inicial. Mientras $\nabla f(x_k) \neq 0$ repetir:

Paso 1. Calcular $d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$.

Paso 2. Definir $x_{k+1} = x_k + d_k$.

Paso 3. Hacer $k = k + 1$ y repetir.

Teorema 3.6. *Convergencia del Método de Newton Puro.* Sea f una función C^2 tal que $\nabla^2 f(x)$ es Lipschitz continuo con constante $L > 0$. Supongamos que x_* es un punto estacionario de f tal que $\nabla^2 f(x_*) > 0$ y supongamos que x_0 esta cerca de x_* . Entonces $x_k \rightarrow x_*$.

Observaciones.

- Observar que el método no hace búsqueda lineal sino que dada la dirección de búsqueda se realiza el paso completo en esa dirección. De aquí el nombre del método.
- Considerando las hipótesis del teorema no se puede garantizar convergencia cuando
 - $\nabla^2 f(x_k)$ es singular o cercano a serlo ya que en este caso puede no ser posible encontrar la solución del sistema $\nabla^2 f(x_k)d = -\nabla f(x_k)$ para algún k .
 - Aún si $\nabla^2 f(x_k)$ es no singular, puede suceder que no haya decrecimiento en la función.

Estas situaciones ocurren, en general, cuando x_0 no está próximo de x_* y no se tiene $\nabla^2 f(x_*) > 0$.

Para demostrar el teorema se necesitan los siguientes lemas.

Lema 3.3. Si $\|A\| < 1$ entonces $I + A$ es no singular y

$$\frac{1}{1 + \|A\|} \leq \|(I + A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

Demostración. Como

$$\|A^n\| \leq \|A\|^n$$

se tiene que

$$\|A\|^n \rightarrow 0$$

ya que $\|A\| < 1$. Luego $A^n \rightarrow 0$.

Luego,

$$(I - A) \sum_{i=0}^n A^i = (I - A)(I + A + A^2 + \dots + A^n) = I + A + A^2 + \dots + A^n - A - A^2 - \dots - A^{n+1} = I - A^{n+1}.$$

Luego,

$$I - A^{n+1} = (I - A) \sum_{i=0}^n A^i$$

y tomando límite cuando $n \rightarrow \infty$ se tiene

$$I = (I - A) \sum_{n=0}^{\infty} A^n.$$

Luego, por un lado,

$$I = (I + A) \sum_{n=0}^{\infty} (-A)^n.$$

entonces

$$(I + A)^{-1} = \sum_{n=0}^{\infty} (-A)^n,$$

y se tiene que

$$\|(I + A)^{-1}\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|}.$$

Por otro lado,

$$1 = \|I\| = \|(I + A)^{-1}(I + A)\| \leq \|(I + A)^{-1}\| \|(I + A)\| \leq \|(I + A)^{-1}\| (1 + \|A\|)$$

y se tiene que

$$\|(I + A)^{-1}\| \geq \frac{1}{1 + \|A\|}$$

como queríamos demostrar. □

Lema 3.4. Si $\nabla^2 f(x^*)$ es no singular y

$$\|B - \nabla^2 f(x^*)\| \leq \frac{1}{2\|(\nabla^2 f(x^*))^{-1}\|}$$

entonces existe B^{-1} y $\|B^{-1}\| \leq 2\|(\nabla^2 f(x^*))^{-1}\|$.

Demostración. Definimos

$$A = B(\nabla^2 f(x^*))^{-1} - I = (B - \nabla^2 f(x^*))(\nabla^2 f(x^*))^{-1}.$$

Luego, por hipótesis,

$$\|A\| = \|(B - \nabla^2 f(x^*))(\nabla^2 f(x^*))^{-1}\| \leq \|(B - \nabla^2 f(x^*))\| \|(\nabla^2 f(x^*))^{-1}\| \leq \frac{1}{2} < 1.$$

Por el Lema previo, existe $(I + A)^{-1} = (B(\nabla^2 f(x^*))^{-1})^{-1}$ y se tiene que existe B^{-1} tal que

$$(I + A)^{-1} = (B(\nabla^2 f(x^*))^{-1})^{-1} = \nabla^2 f(x^*)B^{-1}$$

y

$$\|(I + A)^{-1}\| = \|\nabla^2 f(x^*)B^{-1}\| \leq \frac{1}{1 - \|B(\nabla^2 f(x^*))^{-1} - I\|} \leq 2.$$

Por lo tanto

$$\|B^{-1}\| = \|(\nabla^2 f(x^*))^{-1}\| \leq 2\|(\nabla^2 f(x^*))^{-1}\|$$

como queríamos demostrar. □

Teorema 3.7. Sea $f \in C^2$ y $\nabla^2 f(x)$ Lipschitz continuo con constante $L > 0$. Si x^* es un punto estacionario de f tal que $\nabla^2 f(x^*) > 0$ y si x_0 está suficientemente próximo de x^* entonces $x_k \rightarrow x^*$.

Demostración. Por continuidad suponemos que existe $(\nabla^2 f(x_k))^{-1}$. Tenemos que

$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k)$$

entonces

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - (\nabla^2 f(x_k))^{-1} \nabla f(x_k) = \\ &\quad \nabla^2 f(x_k)^{-1} (\nabla^2 f(x_k)(x_k - x^*) - \nabla f(x_k)) = \\ &\quad \nabla^2 f(x_k)^{-1} (\nabla^2 f(x_k)(x_k - x^*) - (\nabla f(x_k) - \nabla f(x^*))). \end{aligned}$$

Además, como

$$\nabla f(x_k) - \nabla f(x^*) = \int_0^1 \nabla^2 f(x^* + t(x_k - x^*)) (x_k - x^*) dt$$

se tiene que

$$\begin{aligned} \|\nabla^2 f(x_k)(x_k - x^*) - (\nabla f(x_k) - \nabla f(x^*))\| &= \left\| \int_0^1 (\nabla^2 f(x_k) - \nabla^2 f(x^* + t(x_k - x^*))) (x_k - x^*) dt \right\| \leq \\ &\leq \int_0^1 \|\nabla^2 f(x_k) - \nabla^2 f(x^* + t(x_k - x^*))\| \|x_k - x^*\| dt \leq \\ &\leq \int_0^1 L \|x_k - x^*\|^2 (1-t) dt = 1/2 L \|x_k - x^*\|^2. \end{aligned}$$

Por lo tanto

$$\|x_{k+1} - x^*\| \leq \|(\nabla^2 f(x_k))^{-1}\| \frac{L}{2} \|x_k - x^*\|^2. \quad (11)$$

A partir de esta desigualdad, si

$$\|x_0 - x^*\| \leq \frac{1}{2L \|(\nabla^2 f(x^*))^{-1}\|} \quad (12)$$

se tiene que

$$\|\nabla^2 f(x_0) - \nabla^2 f(x^*)\| \leq L \|x_0 - x^*\| \leq \frac{1}{2 \|(\nabla^2 f(x^*))^{-1}\|}$$

entonces, por el lema

$$\text{existe } (\nabla^2 f(x_0)^{-1}) \text{ y } \|\nabla^2 f(x_0)^{-1}\| \leq 2 \|\nabla^2 f(x^*)^{-1}\|$$

entonces x_1 está bien definido y además, usando (11) en la primer desigualdad y (12) en la última,

$$\|x_1 - x^*\| \leq \frac{L}{2} \|(\nabla^2 f(x_0))^{-1}\| \|x_0 - x^*\|^2 \leq L \|(\nabla^2 f(x^*))^{-1}\| \|x_0 - x^*\|^2 \leq 1/2 \|x_0 - x^*\|. \quad (13)$$

Luego,

$$\begin{aligned} \|\nabla^2 f(x_1) - \nabla^2 f(x^*)\| &\leq L \|x_1 - x^*\| \stackrel{(13)}{\leq} \frac{L}{2} \|x_0 - x^*\| \stackrel{(12)}{\leq} \frac{L}{2} \frac{1}{2L \|(\nabla^2 f(x^*))^{-1}\|} = \frac{1}{4 \|(\nabla^2 f(x^*))^{-1}\|} \\ &\leq \frac{1}{2 \|(\nabla^2 f(x^*))^{-1}\|} \end{aligned}$$

entonces, por el lema

$$\text{existe } (\nabla^2 f(x_1)^{-1}) \text{ y } \|\nabla^2 f(x_1)^{-1}\| \leq 2\|\nabla^2 f(x^*)^{-1}\|. \quad (14)$$

Luego, x_2 está bien definido. Además, por (11)

$$\|x_2 - x^*\| \leq \frac{L}{2} \|(\nabla^2 f(x_1))^{-1}\| \|x_1 - x^*\|^2$$

y usando la cota de (14)

$$\|x_2 - x^*\| \leq L \|(\nabla^2 f(x^*))^{-1}\| \|x_1 - x^*\|^2 = L \|(\nabla^2 f(x^*))^{-1}\| \|x_1 - x^*\| \|x_1 - x^*\|.$$

Luego, usando (13),

$$\|x_2 - x^*\| \leq L \|(\nabla^2 f(x^*))^{-1}\| \frac{1}{2} \|x_0 - x^*\| \|x_1 - x^*\|$$

y usando la cota de $\|x_0 - x^*\|$ (12):

$$\|x_2 - x^*\| \leq \frac{1}{4} \|x_1 - x^*\|.$$

Finalmente, por (13)

$$\|x_2 - x^*\| \leq \frac{1}{8} \|x_0 - x^*\|.$$

Repetiendo estos pasos se puede demostrar que x_{k+1} está bien definido y que

$$\|x_k - x^*\| \leq \frac{1}{2^{k+1}} \|x_0 - x^*\| \rightarrow 0 \text{ cuando } k \rightarrow \infty$$

y se tiene la convergencia. □

Observación. Se demuestra además que la velocidad de convergencia es cuadrática:

$$\|x_{k+1} - x^*\| \leq \gamma \|x_k - x^*\|^2.$$

El teorema anterior NO es un teorema de convergencia global ya que depende de la ubicación del punto inicial x_0 . El siguiente ejemplo muestra que el método de Newton puro puede fallar en converger globalmente.

Ejemplo. Considerar $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = -\frac{x^6}{6} + \frac{x^4}{4} + 2x^2$. f tiene un minimizador local en $x^* = 0$ y dos maximizadores globales en $x = \pm\sqrt{(1 + \sqrt{17})/2} \approx \pm 1.6$.

El método de Newton puro aplicado a f iniciando en $x_0 = 1$ tiene las siguientes iteraciones:

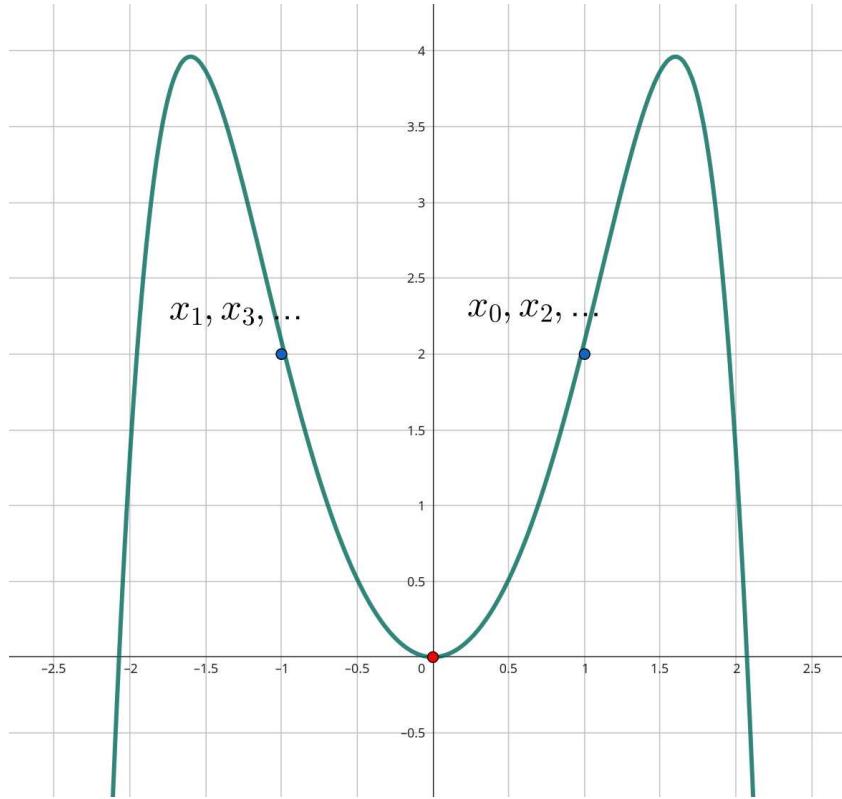
$$k = 0 \quad x_0 = 1 \quad d_0 = -2$$

$$k = 1 \quad x_1 = -1 \quad d_1 = 2$$

\vdots

$$x_{2k} = 1$$

$$x_{2k+1} = -1$$



Observar que las direcciones generadas son de descenso en cada iteración pero la sucesión no se acerca al minimizador local.

Analicemos a continuación la convergencia del siguiente método.

Método de Newton con búsqueda lineal de Armijo. Dado $x_0 \in \mathbb{R}^n$ inicial. Mientras $\nabla f(x_k) \neq 0$ repetir:

Paso 1. Calcular si es posible, $d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$.

Paso 2. Calcular t_k mediante Armijo en la dirección d_k .

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$.

Paso 4. Hacer $k = k + 1$ y repetir.

Para este proceso se puede probar convergencia cuando f es convexa utilizando los resultados de la sección anterior:

Teorema 3.8. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^2 y supongamos que $\nabla^2 f(x)$ es definida positiva para todo $x \in \mathbb{R}^n$. Entonces el Algoritmo anterior con búsqueda de Armijo (o exacta) es globalmente convergente.

Demostración. Considerar $H(x) = (\nabla^2 f(x))^{-1}$ en el Teorema 3.4. □

Un resultado más general es el siguiente:

Teorema 3.9. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^3 . Sea x_* es un minimizador local de f tal que $\nabla^2 f(x_*)$ es definida positiva. Entonces existe $\varepsilon > 0$ tal que, si $\|x_0 - x_*\| < \varepsilon$ y $t_k = 1$ para todo k en la búsqueda de Armijo entonces la sucesión $\{x_k\}$ generada por el método anterior verifica:

1. $\nabla^2 f(x_k)$ es definida positiva para todo k ,

2. $x_k \rightarrow x^*$.

Demostración. Para los interesados, ver por ejemplo Luenberguer, [5]. \square

Este es un teorema de convergencia local que nos dice que si empezamos con x_0 suficientemente cerca del minimizador local x^* entonces:

1. Los sistemas que definen el cálculo de las direcciones están bien definidos,
2. toda la sucesión converge al minimizador local.

El método de Newton se puede modificar cuando hay singularidad de la matriz Hessiana.

Newton con modificación de la diagonal. Dado $x_0 \in \mathbb{R}^n$ inicial. Mientras $\nabla f(x_k) \neq 0$ repetir:

Paso 1. Calcular la dirección d_k .

- 1a. Si $\nabla^2 f(x_k) > 0$ entonces $B_k = \nabla^2 f(x_k)$.
- 1b. Si no, $B_k = \nabla^2 f(x_k) + \mu_k I$ con μ_k elegido de modo que B_k sea definida positiva.
- 1c. Calcular $d_k = -(B_k)^{-1} \nabla f(x_k)$.

Paso 2. Calcular t_k mediante Wolfe en la dirección d_k .

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$.

Paso 4. Hacer $k = k + 1$ y repetir.

Este es un método práctico donde se modifica la diagonal de la matriz Hessiana de modo de obtener una nueva matriz semejante que sea definida positiva.

Teorema 3.10. *Sea f una función C^2 y supongamos que cada matriz B_k verifica $\|B_k\| \|B_k^{-1}\| \leq C$ para $C > 0$ y para todo k . Entonces todo punto de acumulación de x_k es punto estacionario de f .*

Demostración. El resultado se obtiene utilizando la condición de Zoutendijk. \square

Vimos que la dirección de Newton es calculada como solución de un sistema lineal de $n \times n$:

$$\nabla^2 f(x_k) d = -\nabla f(x_k).$$

Una manera de resolver este sistema puede ser mediante factorizaciones de $\nabla^2 f(x_k)$ aunque, en principio, no sepamos si esta matriz es definida positiva. Podríamos por ejemplo, intentar la factorización de Cholesky de la matriz. Si es posible, se tiene que $\nabla^2 f(x_k) = LDL^T$ y la dirección se encuentra mediante la solución de dos sistemas triangulares. Si no es posible obtener la factorización de Cholesky de la matriz Hessiana quiere decir que esta matriz no es definida positiva y en ese caso una posibilidad es aumentar la diagonal de $\nabla^2 f(x_k)$ como se propuso previamente. Esta nueva matriz es definida positiva y tiene factorización de Cholesky.

Otra manera es usar un método iterativo y terminar en alguna iteración con alguna solución inexacta pero aproximada. Muchas de estas reglas iterativas están basadas en la medición del residuo del sistema

$$r_k = \nabla^2 f(x_k) d_k + \nabla f(x_k)$$

donde d_k es la dirección buscada llamada *dirección de Newton inexacta*. Una posibilidad puede ser considerar una adaptación apropiada del método de gradientes conjugados.

3.6 Métodos de quasi-Newton

En el método del gradiente se utiliza la dirección de máximo descenso $d_k = -\nabla f(x_k)$ que puede ser muy lento si las curvas de nivel son muy alargadas.

En el método de Newton se utiliza la dirección $d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$. Este es un método ideal, por su teoría de convergencia, pero tiene sus inconvenientes:

1. puede ser muy difícil de calcular y evaluar la matriz Hessiana (y se necesita en cada iteración);
2. requiere resolver un sistema lineal de ecuaciones de $n \times n$ en cada iteración para calcular d_k :

$$(\nabla^2 f(x_k))d_k = -\nabla f(x_k);$$

3. Puede no generar direcciones de descenso (cuando $\nabla^2 f(x_k)$ no es definida positiva).

Sin embargo, su filosofía es acertada.

Los métodos de Quasi-Newton se definen para que se parezcan lo más posible al método de Newton pero de modo que sean más fáciles de aplicar iterativamente.

En lugar de considerar el modelo exacto (usando $\nabla^2 f(x_k)$ como en Newton) se considera un modelo cuadrático aproximado de la forma:

$$m_k(x) = f(x_k) + \nabla f(x_k)^T(x - x_k) + 1/2(x - x_k)^T B_k(x - x_k)$$

donde B_k es una matriz simétrica definida positiva y se desea que sea una “buena aproximación” del $\nabla^2 f(x_k)$.

Además, es deseable que al pasar de B_k a B_{k+1} no sea necesario introducir muchos cálculos.

Método general de Quasi-Newton. Dados $x_0 \in \mathbb{R}^n$ inicial, B_0 inicial simétrica definida positiva.

Paso 1. Calcular la dirección d_k tal que $B_k d_k = -\nabla f(x_k)$.

Paso 2. Calcular t_k mediante búsqueda lineal en la dirección d_k .

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$.

Paso 4. Calcular $\nabla f(x_{k+1})$, si $\|\nabla f(x_{k+1})\| < \varepsilon$ terminar. Si no, generar una nueva matriz B_{k+1} simétrica definida positiva.

Paso 5. Hacer $k = k + 1$ y repetir.

Observaciones:

1. Si B_k es DP entonces d_k es una dirección de descenso:

$$\nabla f(x_k)^T d_k = \nabla f(x_k)^T B_k^{-1}(-\nabla f(x_k)) = -\nabla f(x_k)^T B_k^{-1} \nabla f(x_k) < 0.$$

2. Si B_k es DP entonces x_{k+1} está bien definido y se tiene que el nuevo modelo m_{k+1} en x_{k+1} será de la forma:

$$m_{k+1}(x) = f(x_{k+1}) + \nabla f(x_{k+1})^T(x - x_{k+1}) + 1/2(x - x_{k+1})^T B_{k+1}(x - x_{k+1}).$$

Luego, los métodos de Quasi-Newton exigen que la nueva matriz B_{k+1} sea tal que $m_{k+1}(x)$ y $f(x)$ tengan gradientes iguales en x_k y x_{k+1} .

Como

$$\nabla m_{k+1}(x) = \nabla f(x_{k+1}) + B_{k+1}(x - x_{k+1}) \text{ entonces}$$

- $\nabla m_{k+1}(x_{k+1}) = \nabla f(x_{k+1}) \checkmark$

- $\nabla m_{k+1}(x_k) = \nabla f(x_k)$ si y solo si

$$\nabla f(x_{k+1}) + B_{k+1}(x_k - x_{k+1}) = \nabla f(x_k) \Leftrightarrow B_{k+1}(x_k - x_{k+1}) = \nabla f(x_k) - \nabla f(x_{k+1})$$

que da lugar a la conocida ecuación secante:

$$B_{k+1}(x_{k+1} - x_k) = \nabla f(x_{k+1}) - \nabla f(x_k), \text{ ecuación secante.}$$

Si definimos $s_k = x_{k+1} - x_k$, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ la ecuación secante se escribe como:

$$B_{k+1}s_k = y_k. \quad (15)$$

Observar que, al premultiplicar en (15) por s_k^T se tiene que:

$$s_k^T B_{k+1} s_k = s_k^T y_k = (x_{k+1} - x_k)^T (\nabla f(x_{k+1}) - \nabla f(x_k)) = (x_{k+1} - x_k)^T \nabla^2 f(\xi_k) (x_{k+1} - x_k)$$

para ξ entre x_k y x_{k+1} . Luego

$$s_k^T B_{k+1} s_k = s_k^T \nabla^2 f(\xi_k) s_k$$

Por lo tanto, B_{k+1} tiene la misma curvatura que $\nabla^2 f(x)$ en un punto intermedio entre x_k y x_{k+1} en la dirección s_k .

Luego, si buscamos matrices B_{k+1} que cumplan (15) tenemos un sistema de n ecuaciones con $n(n+1)/2$ incógnitas. Es decir, hay infinitas matrices que cumplen la ecuación secante. Como mencionamos que es deseable que al pasar de B_k a B_{k+1} no sea necesario realizar demasiados cálculos se propone considerar adaptadas secantes de rango 1 y 2.

- Adaptada secante de rango 2.

Se propone que

$$B_{k+1} = B_k + B'_k + B''_k$$

donde B'_k, B''_k son matrices de rango 1.

Como debe ser $B_{k+1}s_k = y_k$ entonces

$$B_k s_k + B'_k s_k + B''_k s_k = y_k \Rightarrow B'_k s_k + B''_k s_k = y_k - B_k s_k.$$

Existen muchas maneras de que se cumpla la igualdad. Una opción estudiada es la siguiente: proponer que se cumplan las siguientes igualdades

$$B'_k s_k = y_k \quad \text{y} \quad B''_k s_k = -B_k s_k.$$

Entonces, si queremos que B'_k, B''_k sean matrices de rango 1 (de la forma $B = uv^T$) se puede elegir:

$$B'_k = \frac{y_k y_k^T}{y_k^T s_k} \quad (\text{cumple } B'_k s_k = \frac{y_k y_k^T s_k}{y_k^T s_k})$$

$$B''_k = \frac{-B_k s_k s_k^T B_k}{s_k^T B_k s_k} \quad (\text{cumple } B''_k s_k = \frac{-B_k s_k s_k^T B_k s_k}{s_k^T B_k s_k})$$

entonces se obtiene la conocida fórmula BFGS debido a Broyden-Fletcher-Goldfarb-Shanno:

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}.$$

Es la actualización secante más popular y utilizada.

Teorema 3.11. Si B_k es simétrica definida positiva y se tiene que $y_k^T s_k > 0$ entonces B_{k+1} es simétrica definida positiva.

Demostración. Simetría ✓.

Sea $z \in \mathbb{R}^n, z \neq 0$,

$$z^T B_{k+1} z = z^T B_k z + \underbrace{\frac{y_k^T z}{y_k^T s_k}}_{z^T B_k s_k} - \underbrace{\frac{z^T B_k s_k s_k^T B_k z}{s_k^T B_k s_k}}_{(s_k^T B_k z)^2} = z^T B_k z + \underbrace{\frac{(y_k^T z)^2}{y_k^T s_k}}_{\geq 0} - \frac{(s_k^T B_k z)^2}{s_k^T B_k s_k}.$$

Consideremos la expresión

$$z^T B_k z - \frac{(s_k^T B_k z)^2}{s_k^T B_k s_k}. \quad (16)$$

Sean $u = B_k^{1/2} s_k, v = B_k^{1/2} z$. Por la desigualdad de Cauchy-Schwartz: $|u^T v|^2 \leq \|u\|^2 \|v\|^2$ entonces

$$|s_k^T B_k^{1/2} B_k^{1/2} z|^2 \leq \|B_k^{1/2} s_k\|^2 \|B_k^{1/2} z\|^2 \Rightarrow$$

$$(s_k^T B_k z)^2 \leq (s_k^T B_k s_k)(z^T B_k z) \Rightarrow z^T B_k z - \frac{(s_k^T B_k z)^2}{s_k^T B_k s_k} \geq 0.$$

Además, se tiene que $|u^T v|^2 = \|u\|^2 \|v\|^2$ si y solo si u y v son colineales, luego

$$\frac{(s_k^T B_k z)^2}{s_k^T B_k s_k} = z^T B_k z \Leftrightarrow \alpha B_k^{1/2} s_k = B_k^{1/2} z \Leftrightarrow \alpha s_k = z.$$

Pero en este caso $(y_k^T z)^2 = \alpha^2 (y_k^T s_k)^2 > 0$. Por lo tanto $z^T B_{k+1} z > 0$.

Luego, si B_k es definida positiva y $y_k^T s_k > 0$ entonces B_{k+1} es definida positiva. □

Observación. Bajo las consideraciones del teorema anterior las direcciones son de descenso.

Nos falta ver bajo qué condiciones podemos garantizar que $y_k^T s_k > 0$ para todo k .

Esta condición se puede garantizar teniendo en cuenta las reglas de búsqueda lineal que se quieran considerar:

$$\begin{aligned} y_k^T s_k > 0 &\Leftrightarrow (\nabla f(x_{k+1}) - \nabla f(x_k))^T (x_{k+1} - x_k) > 0 \Leftrightarrow \nabla f(x_{k+1})^T \underbrace{(x_{k+1} - x_k)}_{t_k d_k} > \nabla f(x_k)^T (x_{k+1} - x_k) \Leftrightarrow \\ &\nabla f(x_{k+1})^T d_k > \nabla f(x_k)^T d_k. \end{aligned} \quad (17)$$

- Si se utiliza búsqueda exacta, (17) se cumple:

$$\Phi'(t_k) = 0 \Rightarrow \nabla f(x_{k+1})^T d_k = 0 \Rightarrow \nabla f(x_k)^T d_k < 0 = \nabla f(x_{k+1})^T d_k.$$

- Si se utiliza búsqueda de Wolfe con $\sigma_2 \in (0, 1)$, (17) se cumple:

$$\nabla f(x_k)^T d_k \underbrace{<}_{\nabla f(x_k)^T d_k < 0} \sigma_2 \nabla f(x_k)^T d_k \leq \nabla f(x_{k+1})^T d_k.$$

Observación. No alcanza con Armijo para garantizar que $y_k^T s_k > 0$.

En cada paso del método se calcula d_k tal que $B_k d_k = -\nabla f(x_k)$, es decir, $d_k = -B_k^{-1} \nabla f(x_k)$. Dado que B_{k+1} es una adaptada de rango 2 de B_k , se puede calcular B_{k+1}^{-1} a partir de B_k^{-1} . Se utiliza 2 veces la fórmula de Sherman-Morrison para el cálculo de la inversa de una matriz modificada por medio de una adaptada de rango 1:

Sean $u, v \in \mathbb{R}^n$ no nulos y $A \in \mathbb{R}^n$ una matriz no singular. Entonces $A + uv^T$ es no singular si y solo si $\sigma = 1 + v^T A^{-1} u \neq 0$ y en este caso se tiene que $(A + uv^T)^{-1} = A^{-1} - \frac{1}{\sigma} A^{-1} u v^T A^{-1}$.

Luego, si llamamos $H_k = B_k^{-1}$, $H_{k+1} = B_{k+1}^{-1}$ entonces

$$H_{k+1} = H_k + \left(1 + \frac{y_k^T H_k y_k}{y_k^T s_k}\right) \frac{s_k s_k^T}{y_k^T s_k} - \frac{s_k y_k^T H_k + H_k y_k s_k^T}{y_k^T s_k}.$$

Otra forma de escribir H_{k+1} es:

$$H_{k+1} = (I - \rho_k y_k s_k^T)^T H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T \text{ con } \rho_k = \frac{1}{y_k^T s_k}.$$

Con esta fórmula se puede demostrar fácilmente el siguiente resultado:

Proposición 3.2. Si H_k es definida positiva y se tiene que $y_k^T s_k > 0$ entonces H_{k+1} es definida positiva.

Método de Quasi-Newton con adaptada BFGS (directa o inversa). Dados $x_0 \in \mathbb{R}^n$ inicial, H_0 (B_0) inicial simétrica DP. Mientras $\nabla f(x_k) \neq 0$ repetir:

Paso 1. Calcular la dirección $d_k = -H_k \nabla f(x_k)$ (inversa). ($B_k d_k = -\nabla f(x_k)$ (directa)..).

Paso 2. Calcular la longitud de paso t_k en la dirección d_k .

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$, $s_k = x_{k+1} - x_k$, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$.

Definir

$$\begin{aligned} H_{k+1} &= H_k + \left(1 + \frac{y_k^T H_k y_k}{y_k^T s_k}\right) \frac{s_k s_k^T}{y_k^T s_k} - \frac{s_k y_k^T H_k + H_k y_k s_k^T}{y_k^T s_k} \\ &\quad (B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}) \end{aligned}$$

Paso 4. Hacer $k = k + 1$ y repetir.

Ejemplo. Realizar algunas iteraciones del Método de Quasi-Newton con fórmula de la inversa y búsqueda exacta a partir del punto $x_0 = (0, 0)$ para la función $f(x, y) = -2xy - 2y + x^2 + 2y^2$ considerando $H_0 = I$.

$$x_0 = (0, 0), \nabla f(0, 0) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}.$$

$$d_0 = -H_0 \nabla f(0, 0) = -\begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}. \text{ Entonces, usando búsqueda exacta } t = 1/4.$$

$$x_1 = x_0 + 1/4 d_0 = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}.$$

Adaptar

$$H_1 = H_0 + \left(1 + \frac{y_0^T H_0 y_0}{y_0^T s_0}\right) \frac{s_0 s_0^T}{y_0^T s_0} - \frac{s_0 y_0^T H_0 + H_0 y_0 s_0^T}{y_0^T s_0}$$

$$s_0 = x_1 - x_0 = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}.$$

$$y_0 = \nabla f(x_1) - \nabla f(x_0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}. \text{ Entonces}$$

$$s_0 s_0^T = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} (0 \ 1/2) = \begin{pmatrix} 0 & 0 \\ 0 & 1/4 \end{pmatrix}. \quad y_0^T H_0 y_0 = y_0^T y_0 = \|y_0\|^2 = 5.$$

$$s_0 y_0^T H_0 = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} (-1 \ 2) = \begin{pmatrix} 0 & 0 \\ -1/2 & 1 \end{pmatrix}.$$

$$H_0 y_0 s_0^T = y_0 s_0^T = \begin{pmatrix} -1 \\ 2 \end{pmatrix} (0 \ 1/2) = \begin{pmatrix} 0 & -1/2 \\ 0 & 1 \end{pmatrix}.$$

Entonces:

$$H_1 = I + \left(1 + \frac{5}{1}\right) \begin{pmatrix} 0 & 0 \\ 0 & 1/4 \end{pmatrix} - \frac{\begin{pmatrix} 0 & 0 \\ -1/2 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1/2 \\ 0 & 1 \end{pmatrix}}{1} =$$

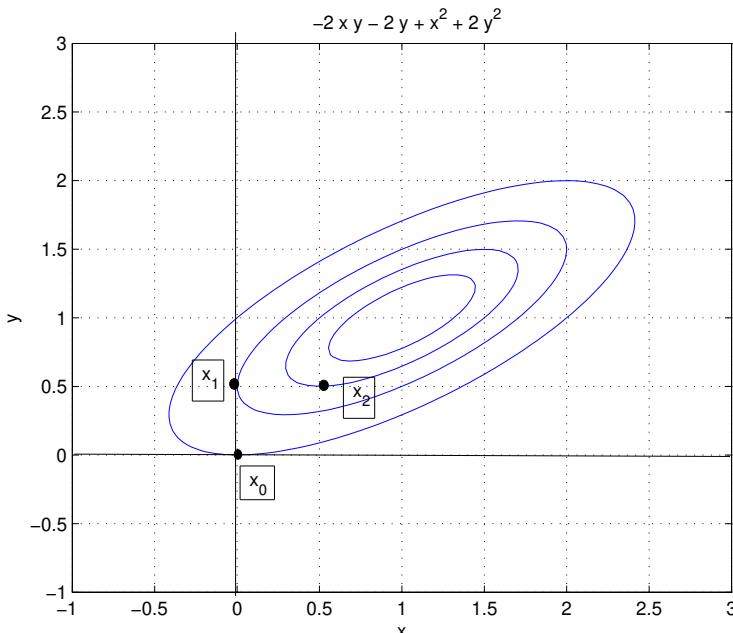
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 3/2 \end{pmatrix} - \begin{pmatrix} 0 & -1/2 \\ -1/2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

Entonces,

$$d_1 = -H_1 \nabla f(x_1) = -\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}.$$

$\varphi(t) = f(x_1 + td_1) = \dots = \frac{t^2}{2} - t - 0.5$ y tiene su minimizador en $t = 1$.

$$\text{Entonces, } x_2 = x_1 + td_1 = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = x^*.$$



Se puede demostrar que $H_2 = (\nabla^2 f(x, y))^{-1}$. Esto se debe a que se trata de una cuadrática convexa.

Teorema 3.12. *Supongamos que el método de Quasi-Newton se aplica con pasos t_k obtenidos mediante búsqueda lineal exacta al problema cuadrático:*

$$\text{Minimizar } q(x)$$

siendo $q(x) = \frac{1}{2}x^T Ax - b^T x$ siendo $A \in \mathbb{R}^{n \times n}$ simétrica definida positiva. Si ningún iterado x_0, x_1, \dots, x_{n-1} es óptimo entonces x_n es solución. Es decir, en n iteraciones, a lo sumo, se obtiene la solución del problema.

Interesados ver [4].

- Adaptada DFP (Davidon-Fletcher-Powell)

Surge de definir

$$H_{k+1} = H_k + H'_k + H''_k$$

siendo H'_k y H''_k matrices simétricas definidas positivas.

Se conoce como adaptada secante de rango 2 inversa (se propone directamente para la matriz inversa).

Se pide que cumpla la ecuación secante para la inversa:

$$H_{k+1}y_k = s_k. \quad (18)$$

Luego se tiene

$$H_{k+1}y_k = H_k y_k + H'_k y_k + H''_k y_k = s_k$$

entonces

$$H'_k y_k + H''_k y_k = s_k - H_k y_k.$$

Una posibilidad entonces es considerar las siguientes matrices de rango 1:

$$H'_k = \frac{s_k s_k^T}{s_k^T y_k}, \quad H''_k = -\frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}.$$

Luego, la adaptada DFP se define como

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}.$$

Es como la adaptada BFGS pero intercambiando los roles de y_k y s_k .

Proposición 3.3. Si H_k es simétrica definida positiva y se tiene que $s_k^T y_k > 0$ entonces H_{k+1} es simétrica definida positiva.

Método de Quasi-Newton con adaptada DFP. Dados $x_0 \in \mathbb{R}^n$ inicial, H_0 inicial simétrica DP. Mientras $\nabla f(x_k) \neq 0$ repetir:

Paso 1. Calcular la dirección $d_k = -H_k \nabla f(x_k)$ (inversa).

Paso 2. Calcular la longitud de paso t_k en la dirección d_k .

Paso 3. Definir $x_{k+1} = x_k + t_k d_k$, $s_k = x_{k+1} - x_k$, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$.

Definir

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}.$$

Paso 4. Hacer $k = k + 1$ y repetir.

Observaciones: 1. Se puede iniciar con $H_0 = I$.

2. De la misma manera que se demostró para BFGS se puede demostrar que el algoritmo está bien definido cuando se considera búsqueda exacta o de Wolfe siempre que se considere H_0 simétrica definida positiva.

Las actualizaciones BFGS para la inversa y DFP se engloban en la siguiente fórmula:

$$H_{k+1}^* = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \xi_k (y_k^T H_k y_k) w_k w_k^T$$

para $\xi_k \in [0, 1]$ con $w_k = \frac{s_k}{s_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k}$.

- Si $\xi_k = 0$ es DPF.
- Si $\xi_k = 1$ es BFGS para la inversa.

Se puede demostrar:

- Como las fórmulas BFGS para la inversa y DFP cumplen la ecuación secante inversa (18) entonces H_{k+1}^* cumple (18).
- Como DFP y BFGS para la inversa cumplen la Proposición 3.3 entonces H_{k+1}^* cumple (3.3) para $\xi_k \in [0, 1]$.

Teorema 3.13. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^1 acotada inferiormente. Sea $x_0 \in \mathbb{R}^n$ inicial y $\{x_k\}$ una sucesión generada por el método de Quasi-Newton para $d_k = -H_k \nabla f(x_k)$ calculada mediante la fórmula adaptada

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \xi_k (y_k^T H_k y_k) w_k w_k^T$$

para $\xi_k \in [0, 1]$ con $w_k = \frac{s_k}{s_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k}$ utilizando H_0 simétrica definida positiva y búsqueda de Wolfe.

Suponemos que cada matriz H_k posee sus autovalores contenidos en $[\lambda_{\min}, \lambda_{\max}]$, $0 < \lambda_{\min} < \lambda_{\max}$. Si x^* es un punto de acumulación de $\{x_k\}$ entonces $\nabla f(x^*) = 0$.

Demostración. Sea $K \subset \mathbb{N}$ tal que $\lim_{k \in K} x_k = x^*$. Suponemos que $\nabla f(x^*) \neq 0$.

Por continuidad sabemos que

$$\lim_{k \in K} f(x_k) = f(x^*).$$

Además, como vale la condición de Armijo y d_k es una dirección de descenso se tiene que $\{f(x_k)\}$ es monótona decreciente entonces debe ser

$$\lim_{k \in \mathbb{N}} f(x_k) = f(x^*).$$

Luego,

$$\lim_{k \in \mathbb{N}} f(x_{k+1}) - f(x_k) = 0. \quad (19)$$

Por la condición de Armijo se tiene que

$$f(x_{k+1}) - f(x_k) \leq \sigma_1 t_k \nabla f(x_k)^T d_k < 0.$$

Luego, por (19)

$$\lim_{k \in \mathbb{N}} t_k \nabla f(x_k)^T d_k = 0. \quad (20)$$

Además, como $d_k = -H_k \nabla f(x_k)$ se tiene que

$$\|d_k\|^2 = d_k^T d_k = \nabla f(x_k)^T H_k^2 \nabla f(x_k).$$

Como H_k simétrica y definida positiva, H_k^2 también es simétrica y definida positiva y los autovalores de H_k^2 son los cuadrados de los autovalores de H_k . Entonces

$$\|d_k\|^2 = d_k^T d_k = \nabla f(x_k)^T H_k^2 \nabla f(x_k) \leq (\lambda_{\max}(H_k))^2 \|\nabla f(x_k)\|^2 \leq \lambda_{\max}^2 \|\nabla f(x_k)\|^2. \quad (21)$$

Como $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^1 se tiene que

$$\lim_{k \in K} \nabla f(x_k) = \nabla f(x^*).$$

Luego, para $k \in K$ suficientemente grande

$$\|d_k\| \leq 2 \|\nabla f(x^*)\| \lambda_{\max}.$$

Es decir, $\{d_k\}_{k \in K}$ es acotada. Como toda sucesión acotada tiene una subsucesión convergente, existe $K_1 \subset K$ y $d \in \mathbb{R}^n$ tales que $\lim_{k \in K_1} d_k = d$.

Luego, por un lado, para $k \in K_1$ grande

$$\nabla f(x_k)^T d_k = -\nabla f(x_k)^T H_k \nabla f(x_k) \leq -\lambda_{\min} \|\nabla f(x_k)\|^2 \leq -\lambda_{\min} \frac{\|\nabla f(x^*)\|^2}{2}$$

y se tiene que

$$\lim_{k \in K_1} \nabla f(x_k)^T d_k \leq -\lambda_{\min} \frac{\|\nabla f(x^*)\|^2}{2} < 0. \quad (22)$$

Luego, por (20) debe ser

$$\lim_{k \in K_1} t_k = 0.$$

Luego, para cada $k \in K_1$ grande, Armijo no vale para una longitud de paso anterior. Es decir, existe $t'_k > 0, t'_k \rightarrow 0$ para el cual Armijo no vale:

$$f(x_k + t'_k d_k) - f(x_k) > \sigma_1 t'_k \nabla f(x_k)^T d_k.$$

Utilizando el Teorema del Valor Medio existe $0 < \xi_k < t'_k$ tal que

$$\nabla f(x_k + \xi_k d_k)^T d_k > \sigma_1 \nabla f(x_k)^T d_k.$$

Tomando límite para $k \in K_1$ obtenemos

$$\nabla f(x^*)^T d \geq \sigma_1 \nabla f(x^*)^T d$$

y debe ser $\nabla f(x^*)^T d \geq 0$ contradiciendo (22).

Por lo tanto debe ser $\nabla f(x^*) = 0$. □

3.7 Función de Rosenbrock

La función de Rosenbrock es una función no convexa utilizada como problema test para estudiar el rendimiento de algoritmos de optimización. Es también conocida como la función del plátano o banana por sus curvas de nivel.

La definición general en el plano está dada por:

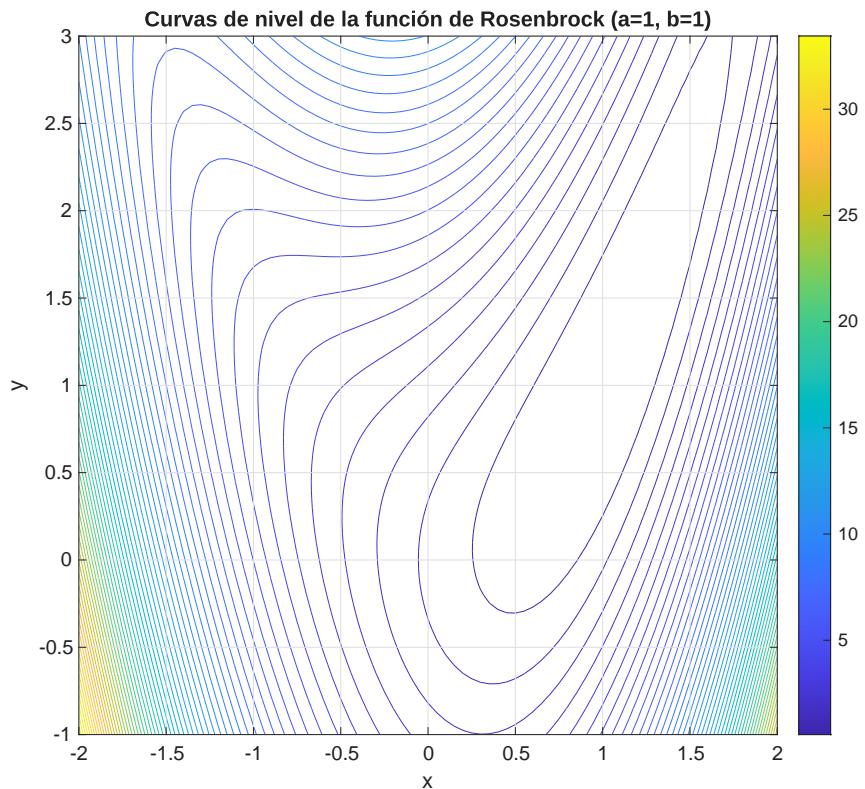
$$f(x, y) = (a - x)^2 + b(y - x^2)^2$$

Tiene un minimizador global en $(x, y) = (a, a^2)$, donde $f(x, y) = 0$.

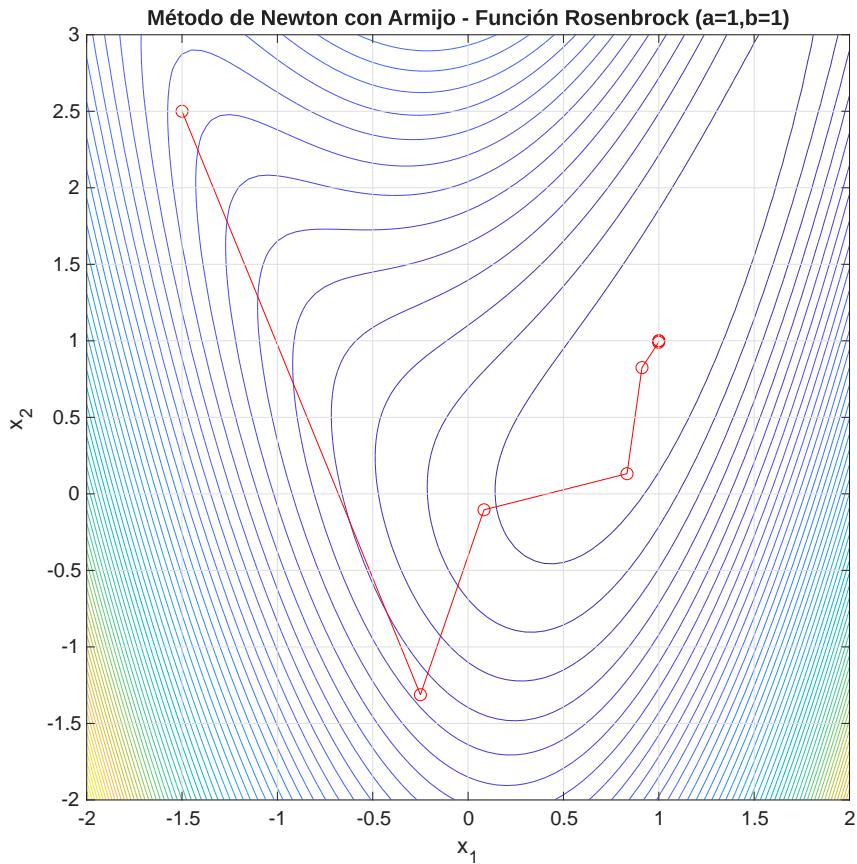
El minimizador global está dentro de un valle plano, largo, estrecho y de forma parabólica. Debido a esto, converger al minimizador puede ser difícil.

Veamos a continuación cómo se comportan los métodos de descenso cuando se aplican a la función de Rosenbrock para $a = b = 1$.

La siguiente figura muestra las curvas de nivel para la función de Rosenbrock con los parámetros $a = b = 1$.



Observemos que sucede con el método de Newton con búsqueda de Armijo comenzando en el punto $(-1.5, 2.5)$.



La secuencia de iterados es la siguiente:

Tabla con los iterados obtenidos

0	$(-1.5, 2.5)$
1	$(-0.25, -1.3125)$
2	$(0.0833, -0.1041)$
3	$(0.8333, 0.1319)$
4	$(0.9117, 0.8251)$
5	$(0.9989, 0.9902)$
6	$(0.9999, 0.9999)$
7	$(1, 1)$

Observemos que sucede con el método de quasi Newton BFGS con búsqueda de Wolfe comenzando en el punto $(-1.5, 2.5)$.

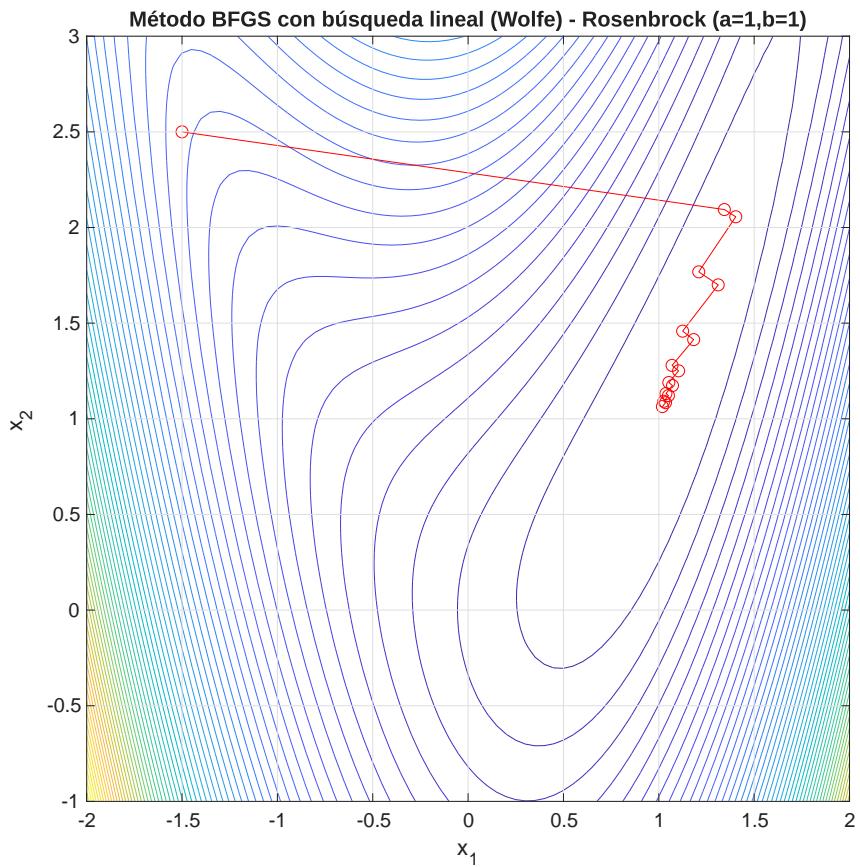


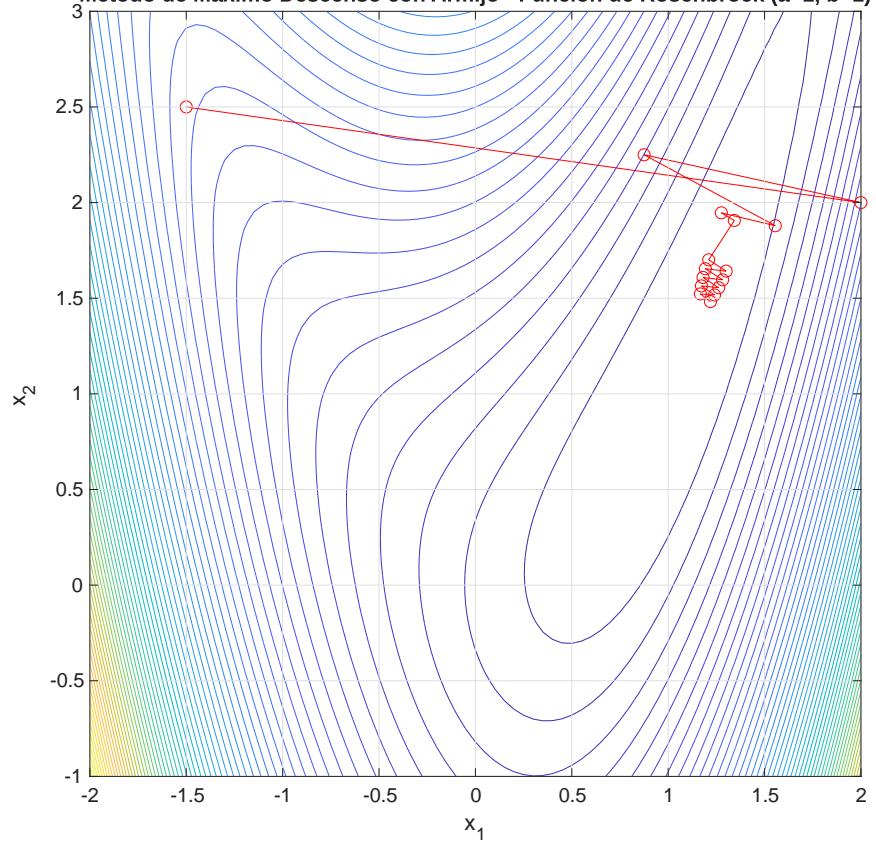
Tabla con los primeros 15 iterados obtenidos

0	(-1.5, 2.5)
1	(1.343750 , 2.093750)
2	(1.403457 , 2.056055)
3	(1.209364 , 1.768220)
4	(1.311523 , 1.700018)
5	(1.124942 , 1.458115)
6	(1.182827 , 1.413914)
7	(1.069399 , 1.279199)
8	(1.103699 , 1.250796)
9	(1.052615 , 1.189584)
10	(1.071750 , 1.173803)
11	(1.037453 , 1.131906)
12	(1.050449 , 1.121361)
13	(1.026428 , 1.091568)
14	(1.035193 , 1.084545)
15	(1.018810 , 1.063887)

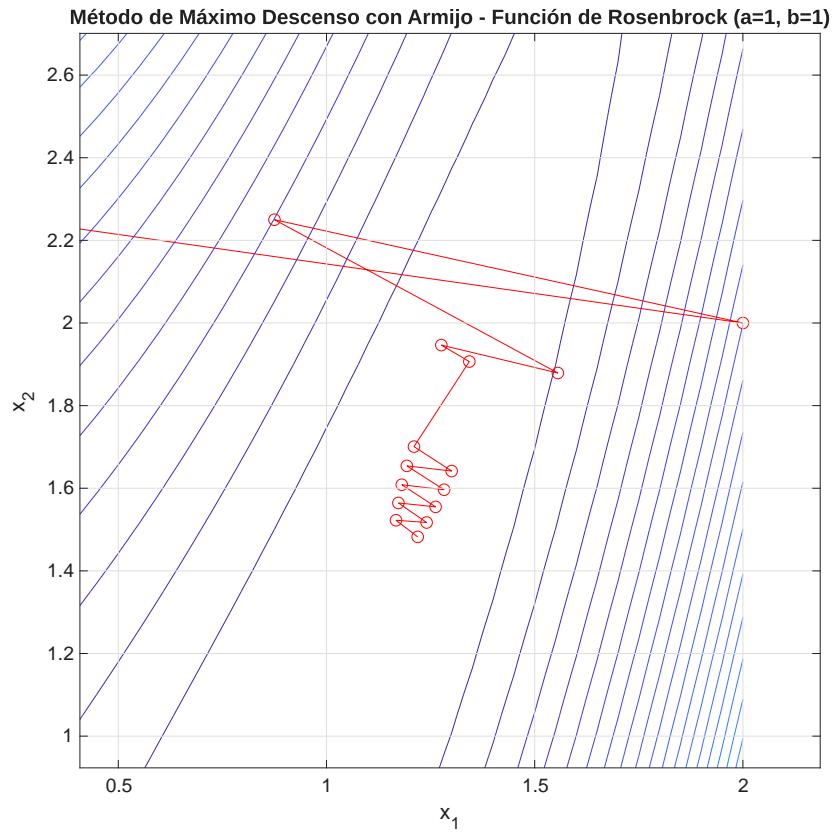
Observemos que sucede con el método de máximo descenso con búsqueda de Armijo comenzando en el punto $(-1.5, 2.5)$ y realizando 15 iteraciones.

Se tiene para $k = 20$ encuentra el punto $(1.219024, 1.482306)$.

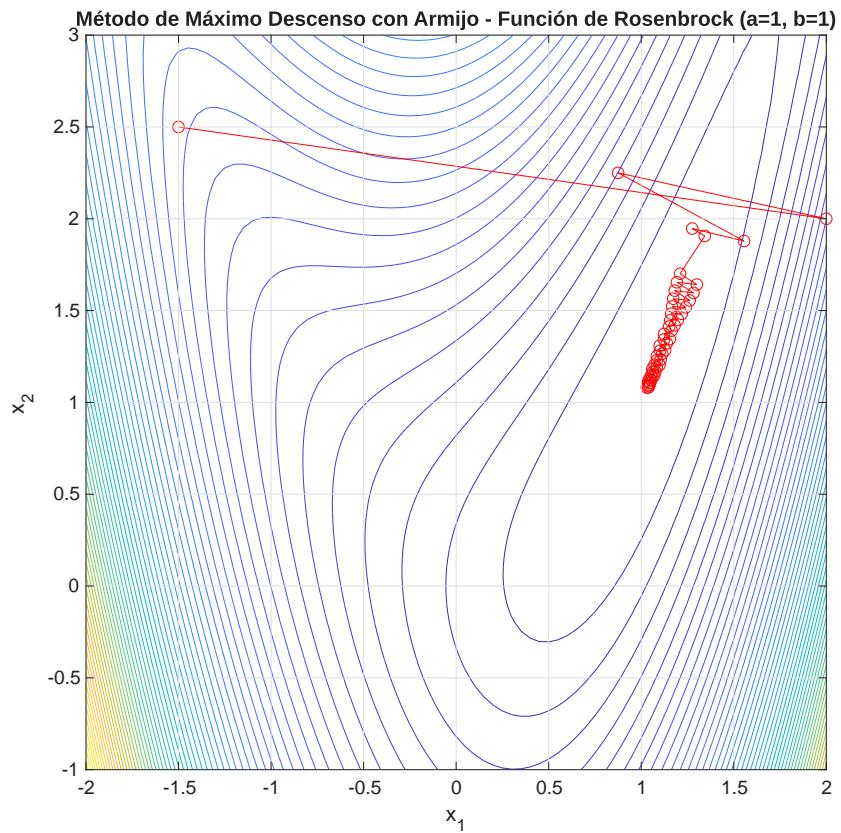
Método de Máximo Descenso con Armijo - Función de Rosenbrock ($a=1$, $b=1$)



Haciendo un zoom en los últimos iterados:



Si permitimos 50 iteraciones, se tiene la siguiente figura, en la cual, para $k = 50$ encuentra el punto $(1.030074, 1.081084)$.



3.8 Método de direcciones conjugadas

Los métodos de direcciones conjugadas son métodos de primer orden (usan información de la función objetivo y de su gradiente) que poseen convergencia mas rápida que el método de máximo descenso.

Definición. Sea $A \in \mathbb{R}^{n \times n}$, una matriz simétrica definida positiva. Decimos que las direcciones $d^0, d^1, \dots, d^k \in \mathbb{R}^n - \{0\}$ son A -conjugadas si

$$(d^i)^T A d^j = 0$$

para $i, j = 0, 1, \dots, k$ con $i \neq j$.

Observar que en el caso en que A es la identidad, las direcciones A -conjugadas son vectores ortogonales en el sentido usual. Se puede demostrar que las direcciones A -conjugadas son linealmente independientes, [4].

El uso de direcciones conjugadas da lugar al algoritmo de gradiente conjugado lineal para minimizar funciones cuadráticas convexas (o para resolver sistemas lineales de ecuaciones con matrices definidas positivas). Este algoritmo fue introducido por Hestenes y Stiefel (1952). Consideremos una función cuadrática en \mathbb{R}^n dada por

$$f(x) = \frac{1}{2}x^T A x + b^T x + c \quad (23)$$

con $A \in \mathbb{R}^{n \times n}$ simétrica y definida positiva, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$. La función f tiene un único minimizador global x^* que cumple

$$Ax^* + b = \nabla f(x^*) = 0. \quad (24)$$

Entonces, dado un conjunto cualquiera de direcciones A -conjugadas $\{d^0, d^1, \dots, d^{n-1}\}$, vamos a definir una sucesión finita de iterados de la siguiente forma:

Dado x^0 inicial, para $k = 0, 1, \dots, n-1$,

$$x^{k+1} = x^k + t_k d^k \quad (25)$$

donde

$$t_k = \operatorname{argmin}_{t>0} f(x^k + td^k).$$

Como f es cuadrática, el valor de t_k se puede calcular de forma exacta:

$$t_k = -\frac{\nabla f(x^k)^T d^k}{(d^k)^T A d^k}. \quad (26)$$

Teorema 3.14. *Considere la función cuadrática dada por (23) y sea x^* su minimizador dado por la ecuación (24). Dado $x^0 \in \mathbb{R}^n$, la sucesión finita definida en (25)-(26) cumple que $x^n = x^*$.*

Demostración.

Por definición de direcciones A -conjugadas, el conjunto $\{d^0, d^1, \dots, d^{n-1}\}$ es una base de \mathbb{R}^n . Por lo tanto, existen escalares $\alpha_i \in \mathbb{R}$, $i = 0, 1, \dots, n-1$, tales que

$$x^* - x^0 = \sum_{i=0}^{n-1} \alpha_i d^i. \quad (27)$$

Considere $k \in \{0, 1, \dots, n-1\}$ arbitrario. Multiplicando en (27) por $(d^k)^T A$ a izquierda y teniendo en cuenta que las direcciones son A -conjugadas, se obtiene que

$$(d^k)^T A(x^* - x^0) = \alpha_k (d^k)^T A d^k.$$

Entonces,

$$\alpha_k = \frac{(d^k)^T A(x^* - x^0)}{(d^k)^T A d^k}. \quad (28)$$

Por otro lado, por la definición de x^k en (25), tenemos

$$x^k = x^0 + t_0 d^0 + t_1 d^1 + \dots + t_{k-1} d^{k-1},$$

que al multiplicar por $(d^k)^T A$ a izquierda implica que

$$(d^k)^T A x^k = (d^k)^T A x^0,$$

ya que las direcciones son A -conjugadas. Sustituyendo esto en (28) y utilizando (24), obtenemos

$$\alpha_k = \frac{(d^k)^T A(x^* - x^0)}{(d^k)^T A d^k} = \frac{-(d^k)^T b - (d^k)^T A x^0}{(d^k)^T A d^k} = -\frac{(d^k)^T(b + Ax^k)}{(d^k)^T A d^k} = -\frac{(d^k)^T \nabla f(x^k)}{(d^k)^T A d^k} = t_k.$$

Por lo tanto, de (27) se sigue que

$$x^* = x^0 + \sum_{i=0}^{n-1} t_i d^i = x^n,$$

completando la demostración. \square

El siguiente resultado es importante para demostrar que el punto x^k minimiza la cuadrática no solo en la recta $x^k + td^k$, sino también en la variedad afín de dimensión k , dada por

$$x^0 + \text{span}\{d^0, d^1, \dots, d^{k-1}\}.$$

Lema 3.5. *Dado $x^0 \in \mathbb{R}^n$, considere la sucesión finita definida en (25)-(26). Entonces,*

$$\nabla f(x^k)^T d^j = 0,$$

para todo $j = 0, 1, \dots, k-1$.

Demostración.

Sabemos que, por el cálculo de t_k mediante búsquela lineal exacta:

$$\nabla f(x^k)^T d^{k-1} = 0,$$

y vale la afirmación para $j = k-1$.

Considere ahora $j < k-1$. Usando que

$$\nabla f(x^{k+1}) = A(x^k + t_k d^k) + b = \nabla f(x^k) + t_k A d^k.$$

y el hecho de que las direcciones son A -conjugadas, obtenemos

$$\nabla f(x^k)^T d^j = (\nabla f(x^{k-1}) + t_{k-1} A d^{k-1})^T d^j = \nabla f(x^{k-1})^T d^j$$

y el resultado se obtiene por el principio de inducción sobre k . \square

El siguiente teorema demuestra que una elección de direcciones conjugadas conduce a obtener no solo el minimizador en n pasos, sino que en cada iteración x^k es el minimizador en el espacio afín generado por las direcciones obtenidas hasta el momento.

Teorema 3.15. Dado $x^0 \in \mathbb{R}^n$, considere la sucesión finita definida en (25)-(26). Entonces, el punto x^k minimiza f sobre la variedad afín

$$C = x^0 + \text{span}\{d^0, d^1, \dots, d^{k-1}\}.$$

Demostración.

Observar que, por (25)-(26), se tiene que $x^k \in C$. Luego, para todo $y \in C$

$$y - x^k \in \text{span}\{d^0, d^1, \dots, d^{k-1}\}.$$

Por lo tanto, por el Lema anterior, se cumple que

$$\nabla f(x^k)^T(y - x^k) = 0.$$

Como f es convexa sabemos que

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) \quad \forall x, y$$

entonces

$$f(y) \geq f(x^k) \quad \forall y \in C$$

y x^k minimiza f en C . □

El abordaje que presentamos del uso de direcciones conjugadas considera minimización en una dirección y luego establece la equivalencia con la minimización en variedades afines de dimensión creciente, partiendo de 1 y llegando a n . Sin embargo, es posible invertir la presentación de estos temas, comenzando con variedades y después obteniendo la minimización unidireccional.

Método de gradientes conjugados para funciones cuadráticas

El algoritmo de Gradientes Conjugados no da un modo de generar direcciones conjugadas.

Dado $x^0 \in \mathbb{R}^n$, se define

$$d^0 = -\nabla f(x^0)$$

y, para $k = 0, 1, \dots, n-2$, la dirección d^{k+1} se obtiene como una combinación de la dirección previa d^k con menos el gradiente de la cuadrática en x^{k+1} ($-\nabla f(x^{k+1})$) es decir

$$d^{k+1} = -\nabla f(x^{k+1}) + \beta_k d^k \quad (29)$$

donde x^{k+1} se calcula como en (25) y el parámetro β_k de la combinación se define de modo que la dirección d^{k+1} sea A -conjugada con d^k . Es decir, queremos que

$$(d^{k+1})^T A d^k = 0.$$

Luego,

$$(d^k)^T A (-\nabla f(x^{k+1}) + \beta_k d^k) = (d^k)^T A d^{k+1} = 0$$

lo que nos proporciona

$$\beta_k = \frac{(d^k)^T A \nabla f(x^{k+1})}{(d^k)^T A d^k}. \quad (30)$$

Podemos ahora presentar el algoritmo de gradientes conjugados.

Algoritmo de gradientes conjugados.

Dados: $x^0 \in \mathbb{R}^n$, definir $d^0 = -\nabla f(x^0)$.

(*) Si $\nabla f(x^k) = 0$ parar. Sino:

Paso 1. Definir el paso en la dirección $d^k, t_k = -\frac{\nabla f(x^k)^T d^k}{(d^k)^T A d^k}$.

Paso 2. Definir el nuevo punto $x^{k+1} = x^k + t_k d^k$.

Paso 3. Calcular el nuevo gradiente $\nabla f(x^{k+1})$.

Paso 4. Definir el parámetro $\beta_k = \frac{(d^k)^T A \nabla f(x^{k+1})}{(d^k)^T A d^k}$ (hace que las direcciones sean conjugadas).

Paso 5. Definir la nueva dirección de búsqueda $d^{k+1} = -\nabla f(x^{k+1}) + \beta_k d^k$. Hacer $k = k + 1$ e ir a (*).

El Algoritmo está bien definido. Si se tiene que

$$\nabla f(x^k) \neq 0,$$

entonces

$$d^k \neq 0,$$

y es una dirección de descenso para f en x^k :

$$\nabla f(x^k)^T d^k = \nabla f(x^k)^T (-\nabla f(x^k) + \beta_{k-1} d^{k-1}) = -\|\nabla f(x^k)\|^2. \quad (31)$$

Luego, el nuevo punto puede ser calculado.

En el siguiente resultado mostramos que las direcciones generadas por el algoritmo son, de hecho, A -conjugadas y que los gradientes son ortogonales.

Teorema 3.16. *Si x^k y d^k son generados por el Algoritmo de gradientes conjugados, entonces*

$$\nabla f(x^k)^T \nabla f(x^j) = 0 \quad y \quad (d^k)^T A d^j = 0,$$

para todo $j = 0, 1, \dots, k - 1$.

Demostración.

Para simplificar la notación, escribamos $g_i = \nabla f(x^i)$. El resultado se demuestra por inducción en k .

Para $k = 1$, usando que $\nabla f(x^{k+1})^T d^k = 0$, obtenemos

$$g_1^T g_0 = -g_1^T d^0 = 0.$$

Además, la definición de β_0 implica que

$$(d^1)^T A d^0 = 0.$$

Supongamos ahora que el resultado vale hasta k . Vamos a probar que vale para $k + 1$.

Por la hipótesis inductiva, las direcciones

$$d^0, d^1, \dots, d^k$$

son A -conjugadas. Por lo tanto, podemos aplicar el Lema 3.5 y concluir que

$$g_{k+1}^T d^j = 0, \quad \text{para } j = 0, 1, \dots, k.$$

Así, usando (29), obtenemos

$$g_{k+1}^T g_j = g_{k+1}^T (-d^j + \beta_{j-1} d^{j-1}) = 0, \quad (32)$$

para $j = 0, 1, \dots, k$.

Finalmente, de la definición de β_k , tenemos que

$$(d^{k+1})^T A d^k = 0.$$

Además, para $j < k$, la hipótesis inductiva implica que

$$(d^{k+1})^T A d^j = (-g_{k+1} + \beta_k d^k)^T A d^j = -g_{k+1}^T A d^j + \beta_k (d^k)^T A d^j.$$

Usando que $\nabla f(x^{k+1}) = \nabla f(x^k) + t_k A d^k$ y (32), obtenemos

$$(d^{k+1})^T A d^j = -g_{k+1}^T \left(\frac{g_{j+1} - g_j}{t_j} \right) = 0,$$

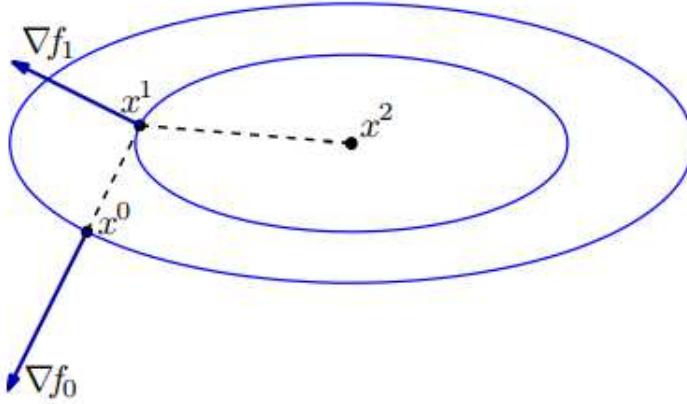
lo que demuestra que

$$(d^{k+1})^T A d^j = 0 \quad \text{para todo } j < k+1,$$

y por tanto las direcciones d^0, d^1, \dots, d^{k+1} son A -conjugadas. \square

La siguiente figura ilustra la aplicación del algoritmo de gradientes conjugados para la minimización de una función cuadrática en \mathbb{R}^2 .

Se puede observar la ortogonalidad de los gradientes en los iterandos y que la solución se obtiene en 2 pasos.



Ejemplo. Para la función cuadrática $f(x, y) = -2xy - 2y + x^2 + 2y^2$ iniciando en $x_0 = (0, 0)$, $d_0 = -\nabla f(0, 0) = (0 \ 2)^T$, $\varphi(t) = f(x_0 + td_0) = -4t + 8t^2$ y tiene su minimizador en $t_0 = \frac{1}{4}$. Luego,

$$x_1 = x_0 + t_0 d_0 = (0 \ \frac{1}{2}).$$

$$k = 1$$

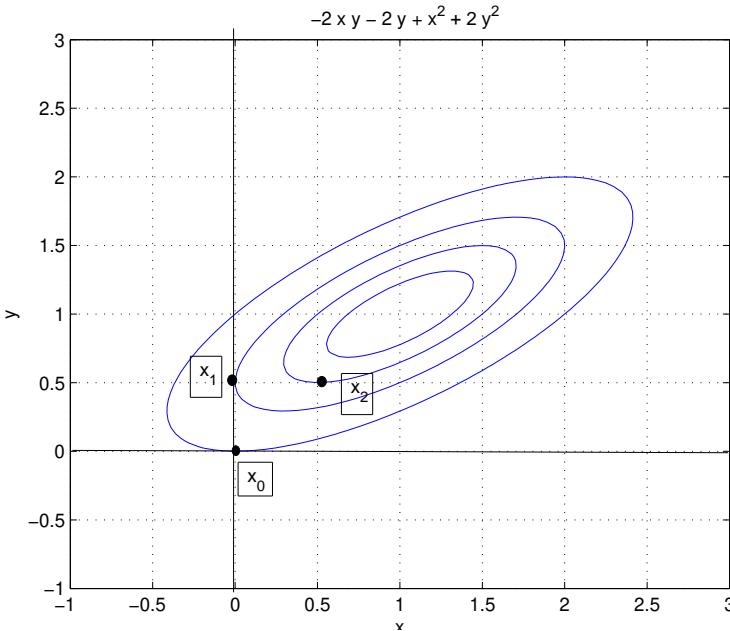
$$\nabla f(x_1) = \nabla f(0 \ \frac{1}{2}) = (-1 \ 0)^T$$

$$\beta_0 = \frac{(d_0)^T A \nabla f(x_1)}{(d_0)^T A d_0} = \frac{4}{16} = 1/4$$

$$d_1 = -\nabla f(x_1) + 1/4 d_0 = (1 \ 0)^T + 1/4(0 \ 2)^T = (1 \ 1/2)^T$$

$$t_1 = -\frac{\nabla f(x_1)^T d_1}{(d_1)^T A d_1} = 1$$

$$x_2 = x_1 + t_1 d_1 = \left(0 \frac{1}{2}\right) + 1 \cdot (1 \frac{1}{2}) = (1 1) = x^*.$$



Extensión del método de gradientes conjugados para funciones no cuadráticas

El método de gradientes conjugados posee adaptaciones para minimizar funciones no cuadráticas. Para eso, es necesario discutir cómo calcular el tamaño del paso t_k y el coeficiente β_k . La búsqueda lineal, que en el caso cuadrático se realizaba de forma cerrada teniendo en cuenta la forma de la función cuadrática, en el caso no cuadrático puede realizarse mediante los métodos de búsqueda inexactas, como Armijo o Wolfe. Para el cálculo de β_k , se pueden utilizar las expresiones más generales que recuperan la vista en el caso cuadrático (30). La fórmula en el caso cuadrático:

$$\beta_k = \frac{(d^k)^T A \nabla f(x^{k+1})}{(d^k)^T A d^k}. \quad (33)$$

da lugar a las siguientes:

$$\beta_k^{PR} = \frac{\nabla f(x^{k+1})^T (\nabla f(x^{k+1}) - \nabla f(x^k))}{\nabla f(x^k)^T \nabla f(x^k)}. \quad (34)$$

propuesta por Polak-Ribière y

$$\beta_k^{FR} = \frac{\nabla f(x^{k+1})^T \nabla f(x^{k+1})}{\nabla f(x^k)^T \nabla f(x^k)}. \quad (35)$$

propuesta por Fletcher-Reeves. Combinando estas opciones, se pueden obtener otras variantes del método.

Teorema 3.17. *Sea f una función cuadrática, entonces las expresiones (33), (34) y (35) coinciden, es decir,*

$$\beta_k = \beta_k^{PR} = \beta_k^{FR}.$$

Demostración.

Denotemos $g_i = \nabla f(x^i)$. Como $\nabla f(x^{k+1}) = \nabla f(x^k) + t_k Ad^k$, tenemos que

$$Ad^k = \frac{g_{k+1} - g_k}{t_k}.$$

Por lo tanto,

$$\beta_k = \frac{g_{k+1}^T Ad^k}{(d^k)^T Ad^k} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{(d^k)^T(g_{k+1} - g_k)}.$$

Usando el Lema 3.5 y (31), obtenemos

$$\beta_k = \frac{g_{k+1}^T(g_{k+1} - g_k)}{g_k^T g_k},$$

lo que prueba la primera igualdad. La otra expresión sigue del hecho de que $g_{k+1}^T g_k = 0$. \square

Cabe destacar que en el caso de minimización de funciones no cuadráticas no se puede asegurar la convergencia en n pasos.

Algoritmo de Gradientes conjugados para funciones no cuadráticas

Dado $x^0 \in \mathbb{R}^n$, haga $d^0 = -\nabla f(x^0)$.

Repetir mientras $\nabla f(x^k) \neq 0$:

Paso 1. Calcular la longitud del paso t_k .

Paso 2. Actualizar $x^{k+1} = x^k + t_k d^k$.

Paso 3. Calcular β_k usando (34) o (35).

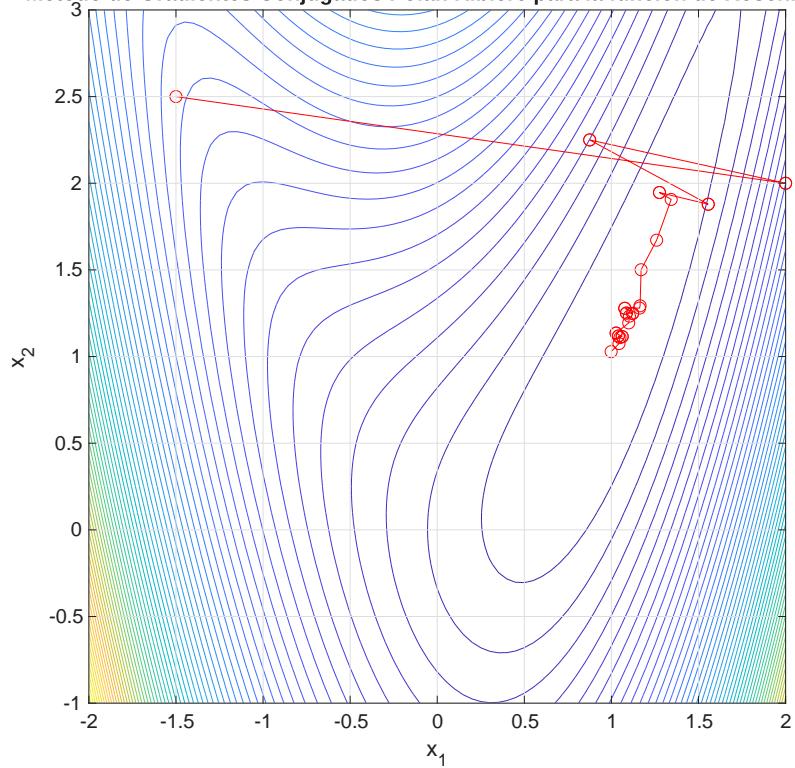
Paso 4. Definir $d^{k+1} = -\nabla f(x^{k+1}) + \beta_k d^k$.

Paso 5. hacer $k = k + 1$.

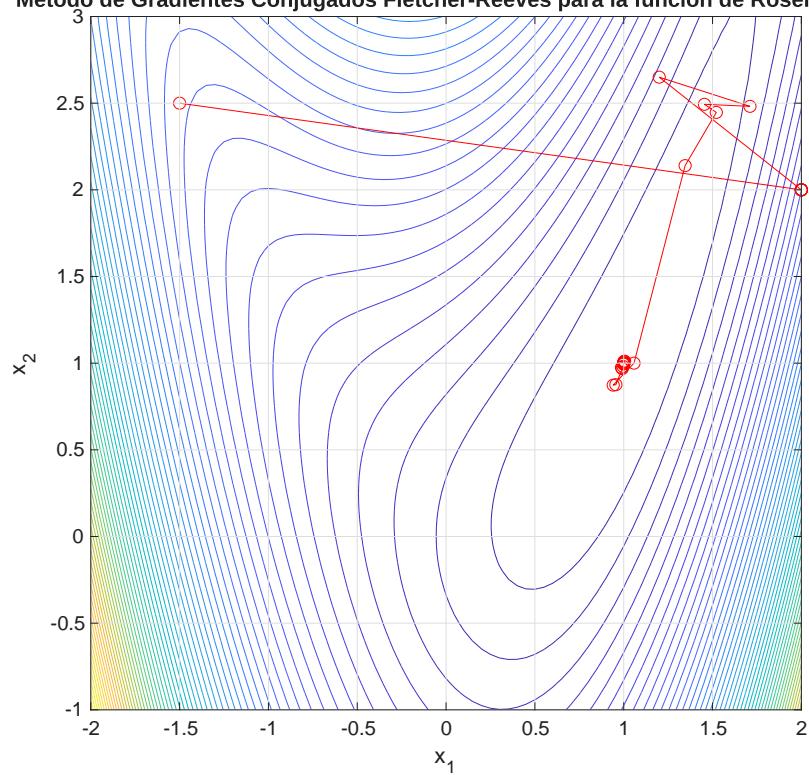
Observar que si t_k es calculado mediante minimización exacta en la dirección, entonces las direcciones generadas por el Algoritmo son de descenso, ya que la relación (31) también se verifica en este caso. Sin embargo, la búsqueda de Armijo no garantiza que se cumpla esta propiedad. Para sortear esta dificultad, existen salvaguardas que pueden ser encontradas con detalles en [6].

Observemos que sucede con el algoritmo de gradientes conjugados para la función de Rosenbrock con búsqueda de Armijo comenzando en el punto $(-1.5, 2.5)$, usando β_k^{PR} y usando β_k^{FR}

Método de Gradientes Conjugados Polak-Ribiere para la función de Rosenbrock



Método de Gradientes Conjugados Fletcher-Reeves para la función de Rosenbrock



3.9 Método de región de confianza

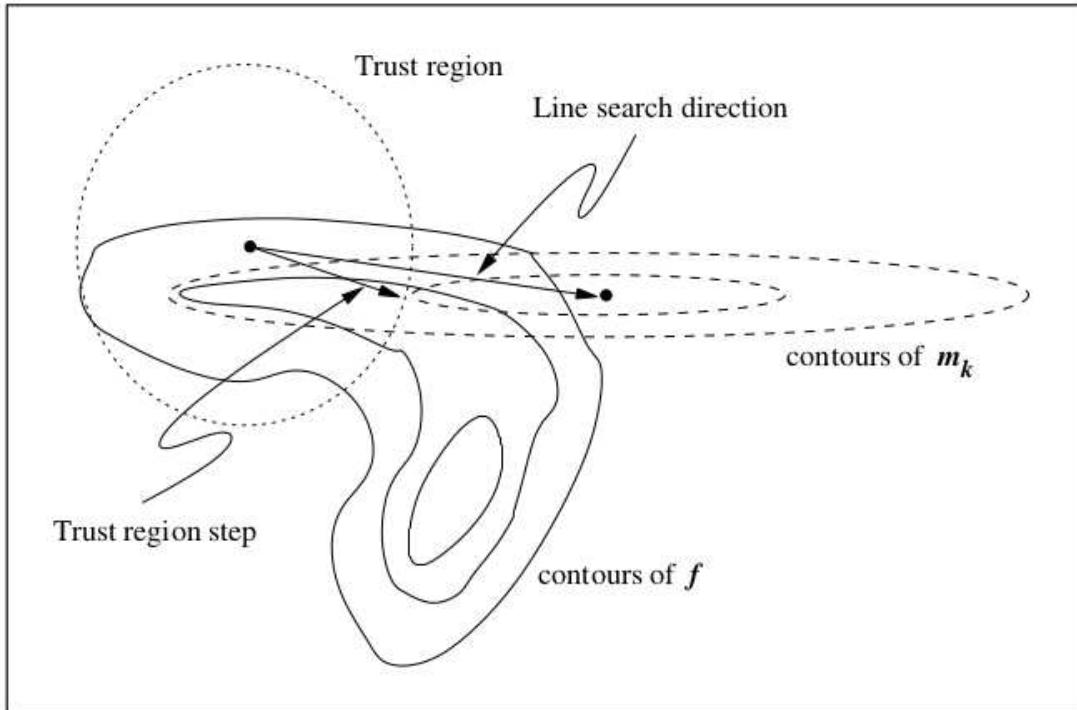
Los métodos basados en regiones de confianza, en cada iterado x_k , definen un modelo $m_k(x)$ de la función objetivo y una región entorno al punto actual en el cual confían en el modelo adoptado. Luego, se calcula un minimizador aproximado del modelo en la región de confianza y se analiza si el punto obtenido es aceptable o no para la función objetivo original.

Si el punto calculado con el modelo produce una reducción razonable en el valor de la función objetivo original el punto se acepta y se repite el procedimiento alrededor de ese nuevo punto. Si no, se piensa que el modelo no representa adecuadamente a la función en la región de confianza considerada. En este caso, se rechaza el punto calculado y se reduce el tamaño de la región de confianza para tratar de encontrar un nuevo punto.

Tanto los métodos de búsqueda lineal como los métodos de regiones de confianza generan sus pasos con la ayuda de un modelo cuadrático de la función objetivo, pero usan el modelo de manera diferente. Los métodos de búsqueda lineal los usan para generar una dirección de búsqueda y después enfocan su esfuerzo en encontrar una longitud de paso apropiado t en esa dirección. Los métodos de regiones de confianza definen una región alrededor del iterado actual dentro de la cual confían en el modelo como una representación adecuada de la función objetivo y eligen el paso como un minimizador aproximado del modelo en la región de confianza. Podríamos decir que estos últimos eligen la dirección y la longitud de paso simultáneamente.

El tamaño de la región de confianza está sujeta a la efectividad del paso. Si la región es muy chiquita, el algoritmo puede perder la posibilidad de realizar un paso sustancial para moverse más cerca de un minimizador de la función objetivo. Si la región es muy grande, el minimizador del modelo puede estar lejos del minimizador de la función objetivo en la región, entonces deberíamos reducir la región e intentar otra vez.

Un excelente tratamiento de este método puede encontrarse en el libro [7].



Consideramos $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^2 y un iterado x_k . Un modelo cuadrático de f alrededor de x_k es:

$$q_k(x) = f(x_k) + \nabla f(x_k)^T(x - x_k) + 1/2(x - x_k)^T B_k(x - x_k)$$

donde $B_k \in \mathbb{R}^{n \times n}$ es una matriz simétrica que puede ser $\nabla^2 f(x_k)$ o se puede adaptar recursivamente.

Sabemos que $m_k(x)$ puede ser una buena aproximación de la función objetivo en un entorno de x_k , entonces, se considera la región:

$$\{x \in \mathbb{R}^n : \|x - x_k\| \leq \Delta_k\}$$

que es la *región de confianza* en el método de región de confianza.

Para simplificar la notación consideramos

$$d = x - x_k, m_k(d) = q_k(x_k + d).$$

Luego, el subproblema en el método de región de confianza es el siguiente problema cuadrático:

$$\begin{aligned} & \text{Minimizar } m_k(d) \\ & \text{s. a } \|d\| \leq \Delta_k. \end{aligned} \tag{36}$$

Al tratarse de un problema en una región compacta, el problema (36) tiene solución (sea B_k definida positiva o no) y la misma es un paso d_k . Una vez hallada la solución d_k del subproblema se evalúa si $x_k + d_k$ es aceptable o no para f . Se desea que el nuevo punto proporcione una reducción en la función objetivo y que la misma sea una fracción de la reducción que produce el modelo. Para formalizar esto se definen:

$$A_{red} = \underbrace{f(x_k) - f(x_k + d_k)}_{\text{reducción real en } f}$$

$$P_{red} = \underbrace{m_k(0) - m_k(d_k)}_{\text{reducción del modelo}}$$

Tanto la estrategia de adaptación del radio de la región de confianza como la aceptación del punto están basadas en la siguiente relación entre lo que baja el modelo y lo que baja la función objetivo:

$$\rho_k = \frac{f(x_k) - f(x_k + d_k)}{m_k(0) - m_k(d_k)} = \frac{A_{red}}{P_{red}}. \tag{37}$$

El numerador se denomina *reducción actual o real* y el denominador *reducción que predice el modelo*. Como se minimiza el modelo, la reducción que predice el modelo es siempre positiva.

Supongamos que se ha calculado una solución d_k del subproblema (36).

- Estrategia para aceptar o no $x_k + d_k$:

- Si $\rho_k > 0$ (o $\rho_k > 0.1$) entonces $x_{k+1} = x_k + d_k$.
- Sino, $x_{k+1} = x_k$.

- Estrategia para adaptar el radio de la región de confianza.

- Si $\rho_k \approx 1$ (por ejemplo $\rho_k > 0.75$) entonces $\Delta_{k+1} = 2\Delta_k$.
- Si $\rho_k \approx 0$ entonces $\Delta_{k+1} = \frac{\Delta_k}{2}$.
- Si no, entonces $\Delta_{k+1} = \Delta_k$.

Algoritmo de región de confianza. Dados $x_0 \in \mathbb{R}^n, \Delta_0 > 0, B_0$ simétrica y $\eta \in [0, 1/4]$. Mientras $\nabla f(x_k) \neq 0$:

1. Calcular d_k una solución del problema (36): $\min m_k(d)$ s. a $\|d\| \leq \Delta_k$.

Definir $x_{new} = x_k + d_k$.

2. Calcular ρ_k dado en (37): $\rho_k = \frac{f(x_k) - f(x_k + d_k)}{m_k(0) - m_k(d_k)} = \frac{A_{red}}{P_{red}}$.

(a) Aceptación del nuevo punto:

- Si $\rho_k > \eta$ entonces $x_{k+1} = x_{new}$.
 - Sino, $x_{k+1} = x_k$.
- (b) Adaptación del radio de confianza para la siguiente iteración:
- Si $\rho_k > 0.75$ entonces $\Delta_{k+1} = 2\Delta_k$.
 - Si $1/4 < \rho_k \leq 0.75$ entonces $\Delta_{k+1} = \Delta_k$.
 - Si no, entonces $\Delta_{k+1} = \frac{\Delta_k}{2}$.

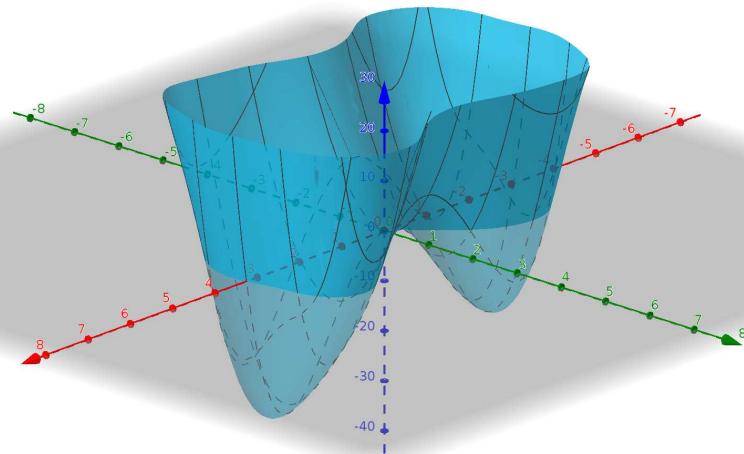
3. Hacer $k = k + 1$, definir una nueva B_{k+1} y repetir.

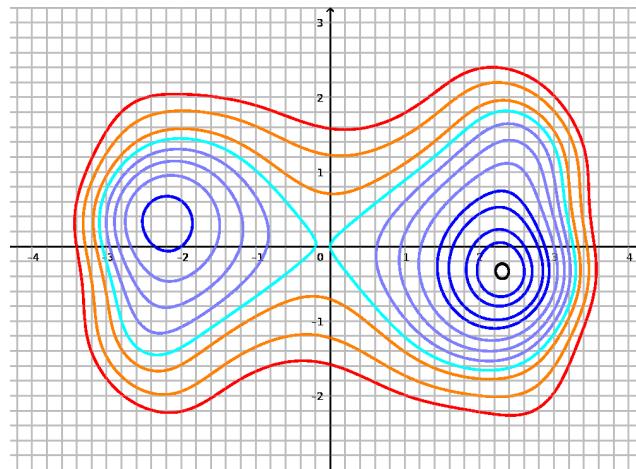
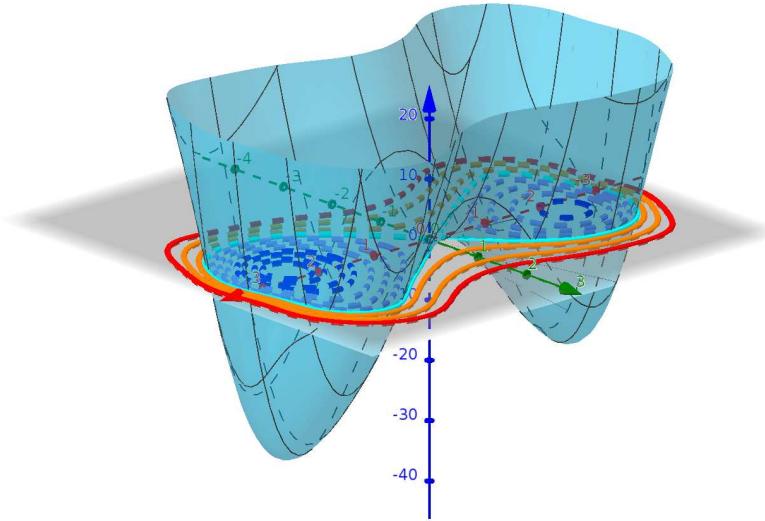
Ejemplo de aplicación del Método de Regiones de Confianza.

Consideremos la función

$$f(x, y) = -10x^2 + 10y^2 + 4\sin(xy) - 2x + x^4$$

Observar la gráfica de la función y sus curvas de nivel.



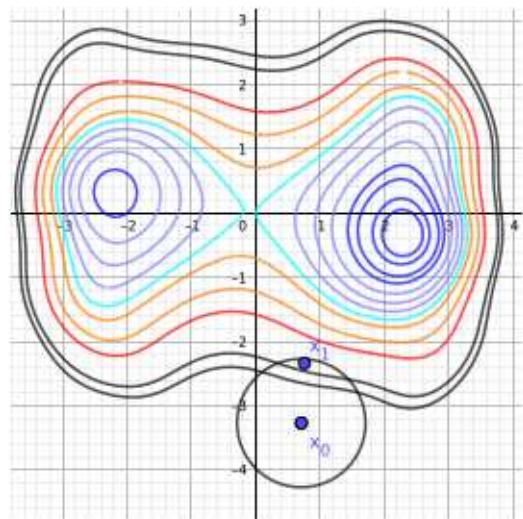


La función tiene dos minimizadores locales: $(-2.20, 0.32)$ y $(2.30, -0.34)$.

Se utilizará el Método de Regiones de Confianza para encontrar alguno de los minimizadores locales de f . Comenzamos con el punto inicial $x_0 = (0.71, -3.27)$.

$$x_0 = (0.71, -3.27) \\ \Delta_0 = 1.$$

$$d_0 = (0.05, 0.93). \\ \rho_0 = \frac{A_{red}}{P_{red}} = 0.998 > 0.1 \Rightarrow \\ x_1 = x_0 + d_0. \\ \rho_0 > 0.75 \Rightarrow \Delta_1 = 2\Delta_0 = 2$$



$$x_1 = x_0 + d_0.$$

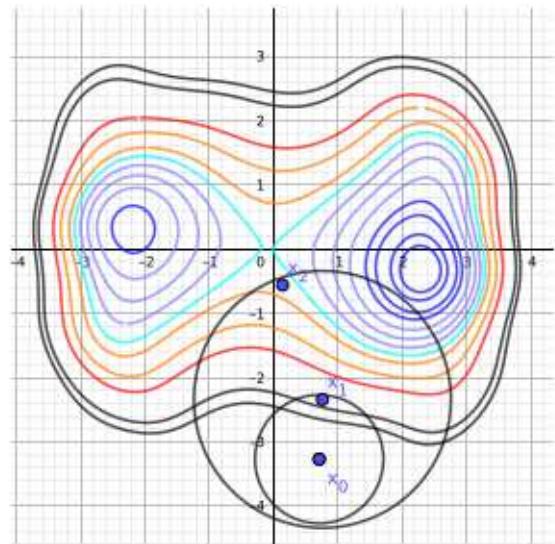
$$\Delta_1 = 2$$

$$d_1 = (-0.62, 1.78)$$

$$\rho_1 = \frac{A_{red}}{P_{red}} = 1.354 > 0.1 \Rightarrow$$

$$x_2 = x_1 + d_1$$

$$\rho_1 > 0.75 \Rightarrow \Delta_2 = 2\Delta_1 = 4$$



$$x_2 = x_1 + d_1$$

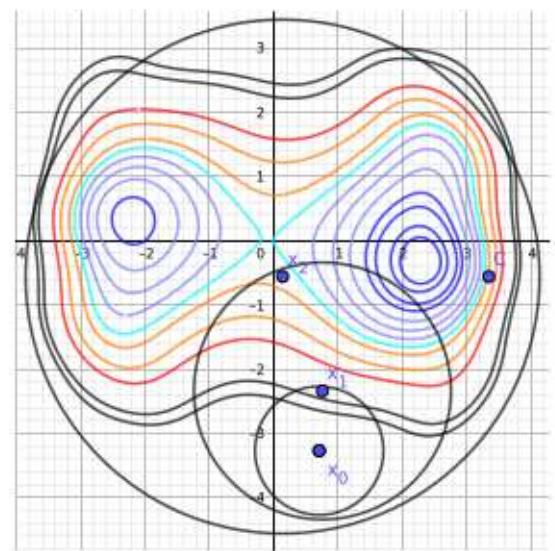
$$\Delta_2 = 4$$

$d_2 = (3.21, 0)$. Genera el punto C en el gráfico.

$$\rho_2 = \frac{A_{red}}{P_{red}} = -0.004 \not> 0.1 \Rightarrow$$

$$x_3 = x_2$$

$$\rho_2 \leq 1/4 \Rightarrow \Delta_3 = 1/2\Delta_2 = 2$$



En esta iteración, gráficamente se observa que el modelo ya no es una buena representación de la función cuando aumentamos el radio de la región de confianza. La solución del subproblema genera el punto C que se encuentra sobre una curva de nivel donde la función objetivo aumenta de valor y por eso se reduce el radio de la región de confianza.

$$x_3 = x_2 = x_1 + d_1$$

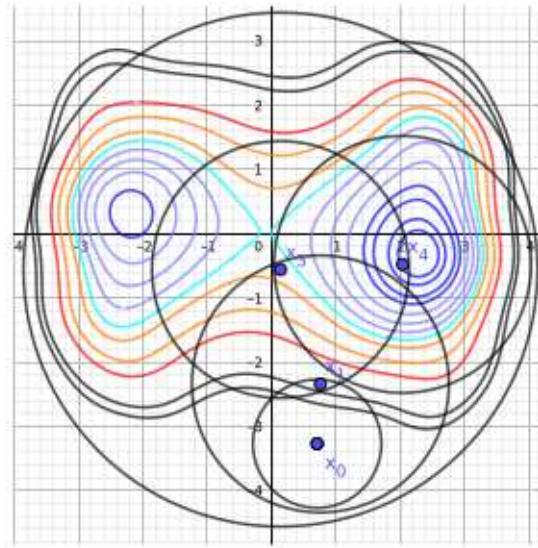
$$\Delta_3 = 2$$

$$d_3 = (1.90, 0.08).$$

$$\rho_3 = \frac{A_{red}}{P_{red}} = 0.649 > 0.1 \Rightarrow$$

$$x_4 = x_3 + d_3$$

$$\rho_3 \leq 0.75 \Rightarrow \Delta_4 = \Delta_3 = 2$$



$$x_4 = x_3 + d_3$$

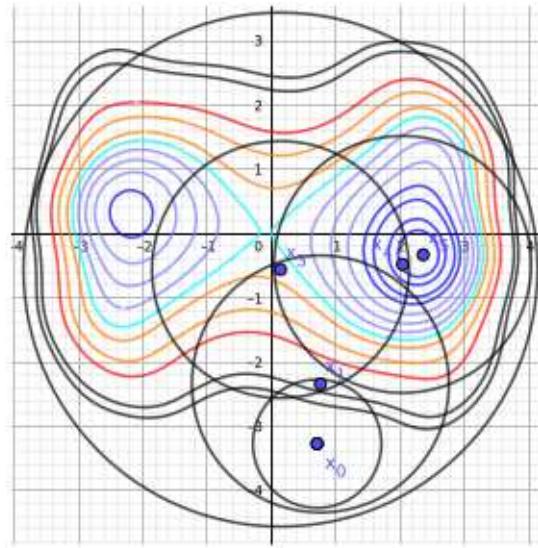
$$\Delta_4 = 2$$

$$d_4 = (0.32, 0.15)$$

$$\rho_4 = \frac{A_{red}}{P_{red}} = 0.857 > 0.1 \Rightarrow$$

$$x_5 = x_4 + d_4$$

$$\rho_4 > 0.75 \Rightarrow \Delta_5 = 2\Delta_4 = 4$$



Luego, se busca la solución con $x_5 = x_4 + d_4$ y $\Delta_5 = 4$.

Se obtiene $d_5 = (-0.03, -0.02)$ con $\rho_5 = \frac{A_{red}}{P_{red}} = 1.009$ entonces $x_6 = x_5 + d_5$.

Luego, en esta iteración la solución aproximada del modelo genera el punto $x_6 = (2.33, -0.35)$ y el algoritmo encuentra aproximadamente uno de los minimizadores de la función objetivo.

k	Δ_k	d_k	$f(x_k + d_k)$	$\frac{A_{red}}{P_{red}}$	x_{k+1}
0	1	(0.05, 0.93)	43.742	0.998	$x_0 + d_0$
1	2	(-0.62, 1.78)	2.306	1.354	$x_1 + d_1$
2	4	(3.21, 0.00)	6.295	-0.004	x_2
3	2	(1.90, 0.08)	-29.293	0.649	$x_2 + d_2$
4	2	(0.32, 0.15)	-31.131	0.857	$x_3 + d_3$
5	4	(-0.03, -0.02)	-31.176	1.009	$x_4 + d_4$

Observaciones:

1. La sucesión de valores funcionales $\{f(x_k)\}$ es no creciente:

- en cada iteración $x_{k+1} = x_k$ o $x_{k+1} = x_{new}$ para el cual $\rho_k > \eta > 0$, es decir $f(x_k) > f(x_{k+1})$.

2. En algunos casos se definen dos parámetros η_1, η_2 y se usan de la siguiente manera:

(a) Aceptación del nuevo punto:

- Si $\rho_k \geq \eta_1$ entonces $x_{k+1} = x_{new}$.
- Sino, $x_{k+1} = x_k$.

(b) Adaptación del radio de confianza para la siguiente iteración:

- Si $\rho_k \geq \eta_2$ entonces $\Delta_{k+1} \in [\Delta_k, \infty)$.
- Si $\rho_k \in [\eta_1, \eta_2)$ entonces $\Delta_{k+1} \in [\gamma_2 \Delta_k, \Delta_k]$.
- Si $\rho_k < \eta_1$, entonces $\Delta_{k+1} \in [\gamma_1 \Delta_k, \gamma_2 \Delta_k]$.

Valores que se consideran: $\eta_1 = 0.1, \eta_2 = 0.75$.

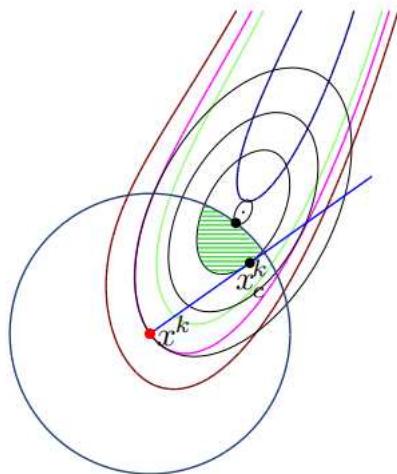
3. Tenemos que decir cómo calcular d_k .

Se busca el *punto de Cauchy*, que es el minimizador del modelo dentro de la región de confianza en la dirección de $-\nabla m_k(0) = -\nabla f(x_k)$ se denomina *punto de Cauchy*. Este punto es importante para garantizar convergencia de un método de región de confianza.

Denotamos $g_k = \nabla f(x_k)$. Si tenemos que $g_k^T B_k g_k \leq 0$ entonces la cuadrática no es definida positiva en esa dirección y el minimizador se encuentra en el borde. Si $g_k^T B_k g_k > 0$ entonces hay que calcular la longitud de paso máxima que se puede dar en esa dirección sin salir de la región. Si el minimizador esta fuera de la región nuevamente el punto de Cauchy estará en el borde. Se define $p_k^C = t_k(-g_k)$ y se busca la longitud de paso $t_k > 0$ que se la solución del problema de minimizar $m_k(t(-g_k))$ sujeto a $\|tg_k\| \leq \Delta_k$. Se puede demostrar que, si $t_k = \frac{\|g_k\|^2}{g_k^T B_k g_k}$ entonces

$$p_k^C = \begin{cases} t_k(-g_k) & \text{si } t_k \leq \frac{\Delta_k}{\|g_k\|} \text{ y } g_k^T B_k g_k > 0 \\ \frac{\Delta_k}{\|g_k\|}(-g_k) & \text{en otro caso.} \end{cases}$$

Este punto tiene importancia crucial para decidir si una solución aproximada del subproblema es aceptable. Concretamente, un método de región de confianza será globalmente convergente si el paso d_k produce una reducción en el modelo m_k que sea menor que un múltiplo fijo de la reducción que se obtiene con el punto de Cauchy.



Se puede demostrar la siguiente proposición.

Proposición 3.4. *La reducción que se obtiene en el modelo con p_k^C es la siguiente:*

$$m_k(0) - m_k(p_k^C) \geq \frac{1}{2} \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}.$$

Luego, se puede demostrar el siguiente teorema, una demostración puede encontrarse en [7].

Teorema 3.18. *Sea $x_0 \in \mathbb{R}^n$. Suponemos que f es dos veces continuamente diferenciable y acotada inferiormente en $\{x : f(x) \leq f(x_0)\}$. Suponemos además que $\|B_k\| \leq \gamma, \gamma > 0$ y que d_k cumple*

$$m_k(0) - m_k(d_k) \geq C_1 \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}$$

para $C_1 \in (0, 1)$. Si $0 \leq \eta < 1/4$ entonces $\liminf_{k \rightarrow \infty} \|g_k\| = 0$.

Teorema 3.19. *En las condiciones del teorema anterior, si $0 < \eta < 1/4$ entonces todo punto de acumulación es estacionario.*

3.10 Ejercicios.

1. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}, x, d \in \mathbb{R}^n, \lambda > 0$ tales que $x + \lambda d$ cumple la condición de Armijo. Sea $0 < \mu < \lambda$. Cumple μ la condición de Armijo? Pruébelo o de un contraejemplo que puede ser gráfico.
2. Considere la función $f(x, y) = x - y + 2x^2 + 2xy + y^2$.
 - (a) Muestre que $d = (-1, 0)$ es una dirección de descenso para f en $(0, 0)$. Analizar cual es el paso óptimo que se puede dar en esa dirección para hacer decrecer el valor de f utilizando búsqueda exacta.
 - (b) Para la dirección de máximo decrecimiento en $(0, 0)$ determinar el intervalo de paso máximo que se puede dar en esa dirección a partir de $(0, 0)$ para hacer decrecer el valor de f utilizando la regla de Armijo con parámetro $\sigma_1 = 1/4$.
3. Considere la función $f(x, y) = 2x^2 + y^2 - 2xy + 2x^3 + x^4$.
 - (a) Verificar que $d = (0, 1)$ es una dirección de descenso para f a partir de $(0, -2)$.
 - (b) Para la dirección a partir de $(0, -2)$ considerada en a), el valor $t = 1$ verifica la regla de Armijo con parámetro $\sigma_1 = 4/5$? Para qué valores de σ_1 el valor de longitud de paso $t = 1$ verifica la regla de Armijo?
4. Sea f una función diferenciable tal que $\nabla f(\bar{x}) \neq 0$. Mostrar que si $H : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ es una función continua que asigna a cada $x \in \mathbb{R}^n$ una matriz definida positiva $H(x)$ entonces la dirección:

$$d = -H(x)\nabla f(x)$$

es una dirección de descenso para f en \bar{x} .

5. Considere la función $f(x, y) = (x - 2y)^2 + x^4$. Calcular la dirección de Newton en el punto $(2, 1)$. Cumple el valor $t = 1$ la regla de Armijo con parámetro $\sigma_1 = 1/5$?
6. Considere el siguiente método:

- Dado x_k . Calcular d_k como se indica a continuación.

- Hacer $t = 1$.

Si $f(x_k + td_k) \leq f(x_k) + \frac{1}{2}td_k^T \nabla f(x_k)$ (*) hacer $x_{k+1} = x_k + td_k$,

Sino, reemplazar t por $t/2$ hasta que se verifique (*)

Sea $f(x, y) = x^2 + y^2 - xy, x_0 = (2, 0)$

- Dibuje algunas curvas de nivel de f .
- Hacer dos iteraciones del método utilizando la dirección de Cauchy. Dibuje los iterados obtenidos en el plano en el cual están las curvas de nivel de f .
- Resuelva el problema mediante el uso de la dirección de Newton.

7. Sean $u, v \in \mathbb{R}^n$ no nulos y $A \in \mathbb{R}^n$ una matriz no singular. Sea $B = A + uv^T$. Demuestre que B es no singular si y solo si $\sigma = 1 + v^T A^{-1} u \neq 0$. En este caso demuestre que $B^{-1} = A^{-1} - \frac{1}{\sigma} A^{-1} u v^T A^{-1}$.

Idea de la demostración:

- Mostrar que la matriz $B = I + xy^T$, para $x, y \in \mathbb{R}^n$ tiene dos autovalores: $\lambda_1 = 1$ con multiplicidad $n-1$ (hay que demostrar que $\dim(Nu(B-\lambda_1 I)) = n-1$, para eso considerar que la imagen de $B-\lambda_1 I = xy^T$ que es una matriz de rango 1 cuya imagen tiene dimensión 1) y $\lambda_2 = 1 + y^T x$.
- Lo anterior implica que

$$\det(B) = 1 + y^T x$$

luego,

$$\det(I + A^{-1}xy^T) = 1 + y^T A^{-1} x.$$

- (c) Finalmente mostrar que $A + uv^T$ es no singular si y solo si $I + A^{-1}uv^T = A^{-1}(A + uv^T)$ es invertible. Luego, hacer el producto de la matriz por su inversa para verificar la fórmula.

- Demostrar que la adaptada BFGS para la inversa cumple: Si H_k es simétrica definida positiva y se tiene que $s_k^T y_k > 0$ entonces H_{k+1} es simétrica definida positiva.
- Considere el método de Quasi-Newton con fórmula adaptada secante DFP de rango 2 con búsqueda lineal exacta y matriz inicial H_0 definida positiva. Demuestre que $y_k^T s_k > 0$ para todo k . Ídem si se utiliza la búsqueda de Wolfe.
- (Optativo) Considere la fórmula BFGS directa para la actualización de B_{k+1} :

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}.$$

Demuestre (usando dos veces la fórmula de Shermann Morrison) que

$$B_{k+1}^{-1} = B_k^{-1} + \left(1 + \frac{y_k^T B_k^{-1} y_k}{y_k^T s_k}\right) \frac{s_k s_k^T}{y_k^T s_k} - \frac{s_k y_k^T B_k^{-1} + B_k^{-1} y_k s_k^T}{y_k^T s_k}.$$

Demostrar que si B_k^{-1} es definida positiva y $y_k^T s_k > 0$ entonces B_{k+1}^{-1} es definida positiva.

Idea de la demostración:

(a) Observar que la matriz $Q = B_k + \frac{y_k y_k^T}{y_k^T s_k}$ es invertible.

(b) Aplicando la fórmula de Shermann-Morrison se puede ver que:

$$Q^{-1} = B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k} = H - \frac{H y_k y_k^T H}{y_k^T s_k + y_k^T H y_k}$$

llamando $H = B_k^{-1}$.

(c) Definir: $r = y_k^T s_k + y_k^T H y_k$. Luego,

$$Q^{-1} B_k = I - \frac{H y_k y_k^T}{r}, \quad B_k Q^{-1} = I - \frac{y_k y_k^T H}{r}, \quad B_k Q^{-1} B_k = B_k - \frac{y_k y_k^T}{r}. \quad (38)$$

Mostrar que

$$B_{k+1} = Q + uv^T$$

siendo $u = -\frac{B_k s_k}{s_k^T B_k s_k}$, $v = B_k s_k$. Usando (38) se puede ver que

$$1 + v^T Q^{-1} u = \frac{(s_k^T y_k)^2}{r(s_k^T B_k s_k)} \neq 0$$

lo que permite nuevamente usar la fórmula de Shermann-Morrison y obtener:

$$B_{k+1}^{-1} = Q^{-1} - \frac{Q^{-1} u v^T Q^{-1}}{1 + v^T Q^{-1} u} = Q^{-1} + r \frac{Q^{-1} B_k s_k s_k^T B_k Q^{-1}}{(s_k^T y_k)^2}.$$

Por (38) se tiene que

$$r(Q^{-1} B_k s_k s_k^T B_k Q^{-1}) = r s_k s_k^T - s_k^T y_k (s_k y_k^T H + H y_k s_k^T) + \frac{(s_k^T y_k)^2}{r} (H y_k y_k^T H)$$

y se puede obtener el resultado.

4 Teoría básica de Optimización con restricciones

En este capítulo vamos a analizar casos en los cuales el conjunto factible no es necesariamente todo \mathbb{R}^n . Este tipo de problemas se denominan *Problemas de optimización con restricciones*. Un problema general de optimización con restricciones tiene la siguiente forma

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0, g(x) \leq 0 \quad (39)$$

donde las funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ son todas diferenciables hasta el orden que sea necesario. La función f se denomina *función objetivo* y el conjunto $\Omega = \{x : h_i(x) = 0, g_j(x) \leq 0, i = 1, \dots, m; j = 1, \dots, p\}$ se denomina *conjunto factible*.

En principio y para fijas ideas analizaremos las condiciones de optimalidad para problemas que presentan solo restricciones de igualdad, luego para problemas con restricciones de desigualdad y finalmente extenderemos los resultados para problemas del tipo (39). Excelentes tratamientos de este tema pueden encontrarse en los libros [8, 5].

4.1 Condiciones necesarias de primer orden para problemas con restricciones de igualdad

Consideramos primero problemas con restricciones de igualdad, de la forma

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0 \quad (40)$$

donde las funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ son todas diferenciables hasta el orden que sea necesario. La función h es de la forma $h(x) = \begin{pmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{pmatrix}$ donde $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$.

Observemos que si consideramos solo una restricción de igualdad en \mathbb{R}^2 entonces el conjunto factible $\Omega = \{(x, y) \in \mathbb{R}^2 : h(x, y) = 0\}$ es una curva que puede estar dada en forma implícita.

Si consideramos solo una restricción de igualdad en \mathbb{R}^3 entonces el conjunto factible $\Omega = \{(x, y, z) \in \mathbb{R}^3 : h(x, y, z) = 0\}$ es una superficie. Si consideramos dos restricciones de igualdad en \mathbb{R}^3 entonces el conjunto factible $\Omega = \{(x, y, z) \in \mathbb{R}^3 : h_1(x, y, z) = 0, h_2(x, y, z) = 0\}$ es la intersección de dos superficies que generalmente es una curva en el espacio.

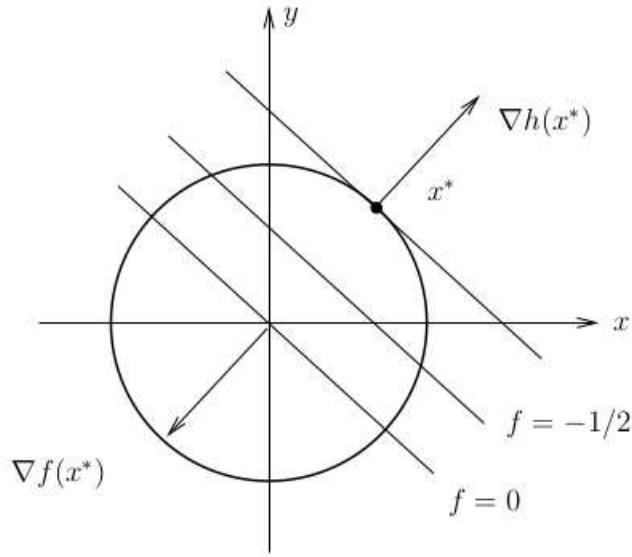
Consideremos por ejemplo el problema

$$\text{Minimizar } -x - y \text{ sujeto a } x^2 + y^2 = 1. \quad (41)$$

En la figura se observa el conjunto factible (en línea continua, la circunferencia unidad) y las curvas de nivel de la función objetivo f (rectas paralelas a $y = -x$). Sabemos que f crece en la dirección del gradiente $\nabla f(x, y) = (-1 - 1)^T$ y luego el minimizador de f se encuentra sobre la recta $y = x$ dentro de la región factible. Luego, la solución del problema es $x^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Observemos que, si llamamos $h(x, y) = x^2 + y^2 - 1$, en la solución se tiene que

$$\nabla f(x^*) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \nabla h(x^*) = \begin{pmatrix} \frac{2}{\sqrt{2}} \\ \frac{2}{\sqrt{2}} \end{pmatrix}.$$

Luego $-\nabla f(x^*) = \frac{\sqrt{2}}{2} \nabla h(x^*)$.



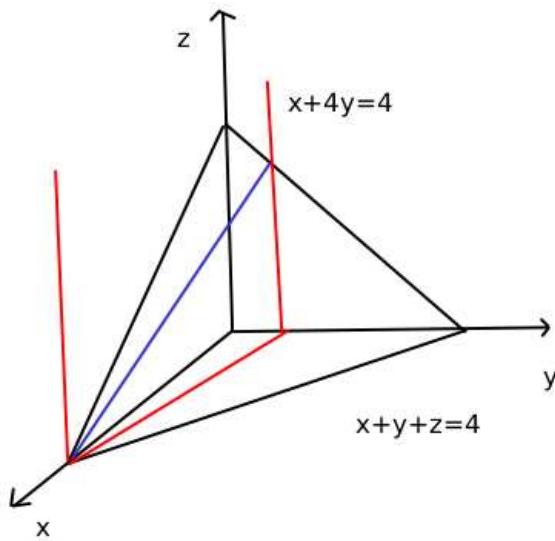
Ejercicio. Observar gráficamente que un comportamiento semejante ocurre para el problema

$$\text{Minimizar } \frac{1}{2}[(x-2)^2 + (y-2)^2] \text{ sujeto a } x^2 + y^2 = 1.$$

Consideremos ahora el siguiente problema en \mathbb{R}^3 :

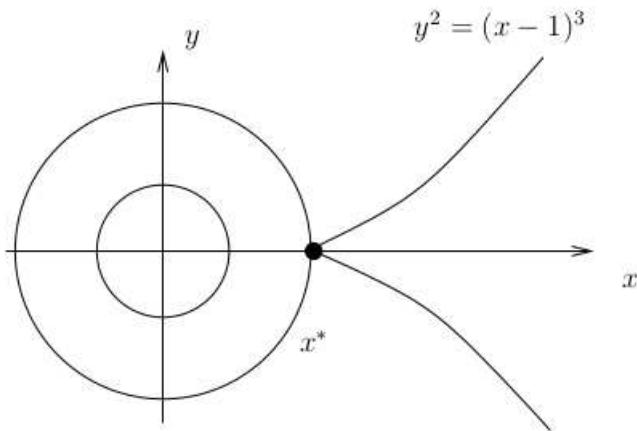
$$\text{Minimizar } x^2 + y^2 + z^2 \text{ sujeto a } x + y + z = 4, x + 4y = 4. \quad (42)$$

Se trata de la formulación del problema de encontrar un punto sobre la intersección de los dos planos que se esté a menor distancia del origen. Definimos $h_1(x, y, z) = x + y + z - 4$, $h_2(x, y, z) = x + 4y - 4$. La intersección de las dos superficies es una curva en el espacio que puede parametrizarse como el conjunto de los puntos de la forma $(4 - 4t, t, 3t)$. De esta manera, se obtiene que $f(4 - 4t, t, 3t) = 26t^2 - 32t + 16$ y tiene su minimizador global en $t = \frac{8}{13}$ obteniendo de esta manera la solución $x^* = (\frac{20}{13}, \frac{8}{13}, \frac{24}{13})$. Verificar en este caso que se tiene $-\nabla f(x^*) = \lambda_1 \nabla h_1(x^*) + \lambda_2 \nabla h_2(x^*)$ para $\lambda_1 = -48/13$, $\lambda_2 = 8/13$.



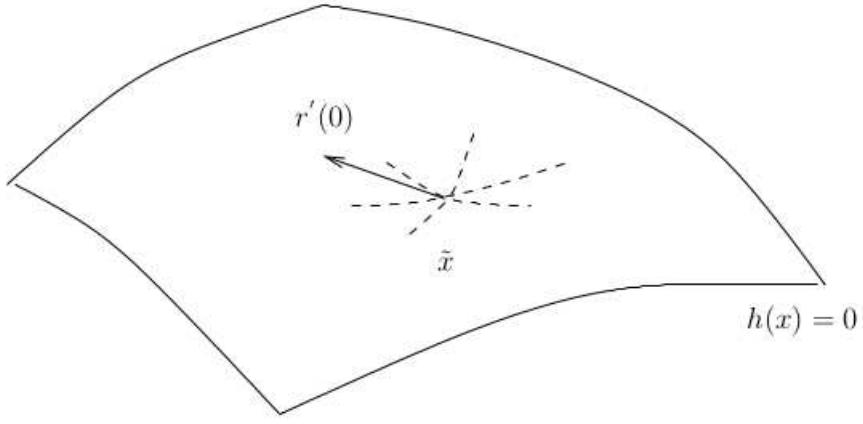
Lamentablemente esta condición ($-\nabla f(x^*)$ combinación lineal de los gradientes de las restricciones) no siempre sucede en un minimizador local.

Considerar el problema de encontrar un punto sobre la curva $y^2 = (x - 1)^3$ que se encuentra a menor distancia al origen. Gráficamente se observa que la solución de este problema es el punto $(1, 0)$, sin embargo, en este punto, como $\nabla h(x^*) = 0$ no existe $\lambda \in \mathbb{R}$ tal que $-\nabla f(x^*) = \lambda \nabla h(x^*)$.



Para que se verifique una combinación entre el gradiente de la función y de las restricciones en un minimizador local es necesario que se cumpla alguna condición extra. Para analizar cuál debe ser esta condición, veamos que sucede con curvas contenidas en la región.

Dado un punto factible \tilde{x} , los caminos que pasan por el punto contenidos en región factible son curvas. Sabemos que las curvas $r(t)$ diferenciables contenidas en la región factible que pasan por \tilde{x} tienen la propiedad de que $r'(t)$ es un vector tangente a $r(t)$ en el punto y es tangente a la superficie, es decir, se encuentra en el plano tangente a la gráfica de la superficie en \tilde{x} .



Si $\Omega = \{x \in \mathbb{R}^n : h(x) = 0\}$ y $r(t) : (-\varepsilon, \varepsilon) \rightarrow \Omega$ es una parametrización de un arco factible que pasa por \tilde{x} es decir $r(0) = \tilde{x}$ entonces $\forall t \in (-\varepsilon, \varepsilon)$

$$\Phi(t) = h(r(t)) = 0.$$

Derivando en función a t (mediante regla de la cadena) obtenemos que

$$\Phi'(t) = Jh(r(t)).r'(t) = 0 \quad \forall t \in (-\varepsilon, \varepsilon), \quad (43)$$

es decir

$$\nabla h_i(r(t))^T.r'(t) = 0 \quad \forall t \in (-\varepsilon, \varepsilon), \forall i = 1, \dots, m.$$

Si $r(0) = \tilde{x}$ entonces $\nabla h_i(\tilde{x})^T.r'(0) = 0, \forall i = 1, \dots, m.$

Es decir, dado un arco factible y diferenciable que pasa por el punto se tiene que el vector tangente al arco en \tilde{x} es ortogonal a los gradientes de las restricciones de igualdad evaluados en \tilde{x} .

Definición 4.1. *El conjunto*

$$T_{x^*} = \{d \in \mathbb{R}^n : \text{ existe } r(t) \text{ factible } C^1 \text{ tal que } r : [-\varepsilon, \varepsilon] \rightarrow \Omega, r(0) = x^*, r'(0) = d\}$$

se define como espacio tangente basado en curvas al conjunto factible en x^* .

Vimos que si $d \in T_{x^*}$ entonces $\nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, m.$

Es decir,

$$T_{x^*} \subset Nu(Jh(x^*)).$$

Observaciones:

- La otra contención no siempre se cumple.

Considerar $\Omega = \{(x, y) : x.y = 0\}$, es decir, se tiene que $h(x, y) = x.y$. Entonces $x.y = 0 \Leftrightarrow x = 0 \text{ o } y = 0$.

Considerar el punto $\tilde{x} = (0, 0)$.

$$Jh(x, y) = \left(\frac{\partial h}{\partial x} \quad \frac{\partial h}{\partial y} \right) = (y \quad x), Jh(\tilde{x}) = (0 \quad 0). \text{ Entonces, } Nu(Jh(\tilde{x})) = \mathbb{R}^2.$$

Ahora, una $d \in T_{\tilde{x}}$ si y solo si $d = r'(0)$ con $r(t) \subset \Omega : r(0) = \tilde{x}$. Es decir, $r(t) = (r_1(t), r_2(t))$ tal que $h(r_1(t), r_2(t)) = 0$ entonces $r_1(t).r_2(t) = 0$ lo que implica que $r_1(t) = 0$ o $r_2(t) = 0$. Entonces $r'_1(t) = 0$ o $r'_2(t) = 0$ y se tiene que $d = r'(0)$ cumple $d_1 = 0$ o $d_2 = 0$. Entonces $d \in \Omega$.

Por lo tanto, $T_{\tilde{x}} \subset \Omega$ y $Nu(Jh(\tilde{x})) \not\subseteq T_{\tilde{x}}$.

- que sucede con la función objetivo sobre este tipo de curvas si \tilde{x} es un minimizador local de $f(x)$ sujeto a $h(x) = 0$?

Sabemos que

$$f(r(t)) = f(\tilde{x}) + \nabla f(\tilde{x})^T r'(0)t + o(t)$$

luego, como \tilde{x} es un minimizador se tiene que

$$\nabla f(\tilde{x})^T r'(0)t + o(t) = f(r(t)) - f(\tilde{x}) \geq 0.$$

Dividiendo por $t > 0$ y haciendo t tender a cero obtenemos que $\nabla f(\tilde{x})^T r'(0) \geq 0$. Dividiendo por $t < 0$ y haciendo t tender a cero obtenemos que $\nabla f(\tilde{x})^T r'(0) \leq 0$, luego debe ser:

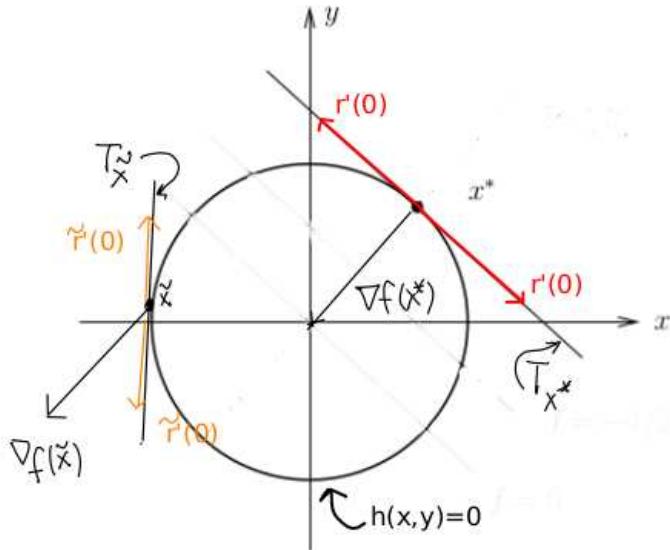
$$\nabla f(\tilde{x})^T r'(0) = 0.$$

Hemos demostrado el siguiente Teorema que nos proporciona una condición necesaria de primer orden.

Teorema 4.2. Si x^* es un minimizador local del problema (40) entonces $\nabla f(x^*)^T d = 0$, para cualquier $d \in T_{x^*}$.

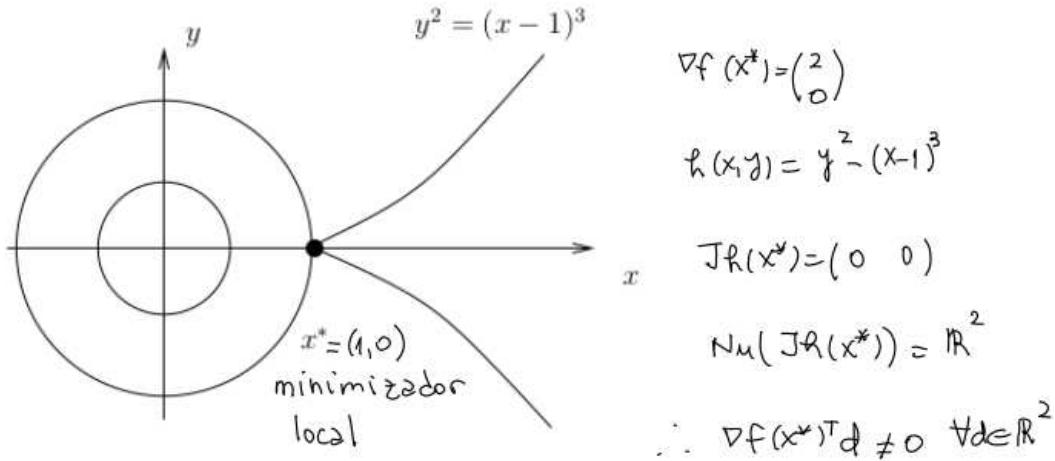
Observar en el problema

$$\text{Minimizar } -x - y \text{ sujeto a } x^2 + y^2 = 1.$$



Si bien el Teorema 4.2 da una condición necesaria de optimalidad de primer orden, la condición es poco práctica ya que no es fácil detectar todas las posibles curvas factibles diferenciables $r(t)$. Nos gustaría un resultado como el del teorema pero en un conjunto que sea más fácil de caracterizar, como podría ser $Nu(Jh(x^*))$. Pero lamentablemente no siempre sucede que:

- Si x^* es un minimizador local del problema (40) entonces $\nabla f(x^*)^T d = 0$, para cualquier $d \in Nu(Jh(x^*))$.



El principal problema en el ejemplo anterior es que

$$T_{x^*} \neq Nu(Jh(x^*)).$$

Nos preguntamos: bajo qué condiciones se puede asegurar que $T_{x^*} = Nu(Jh(x^*))$. Porque en este caso podemos asegurar que

- Si x^* es un minimizador local del problema (40) entonces $\nabla f(x^*)^T d = 0$, para cualquier $d \in Nu(Jh(x^*))$

y es una condición más práctica que la del Teorema 4.2.

Por ejemplo, si se tiene $h(x) = Ax - b$ con $A \in \mathbb{R}^{m \times n}$ se tiene que

$$T_{\tilde{x}} = Nu(Jh(\tilde{x})) = Nu(A)$$

para cualquier punto factible \tilde{x} .

Este caso es importante ya que si tenemos el problema de

$$\text{Minimizar } f(x) \text{ sujeto a } Ax = b$$

y suponemos que $A \in \mathbb{R}^{m \times n}$ una matriz con rango completo por filas (es decir, las filas de A son LI) sabemos que se puede escribir

$$A = [B \ N] \text{ donde } B \in \mathbb{R}^{m \times m} \text{ es no singular y } N \in \mathbb{R}^{m \times (n-m)}.$$

Es decir, que el problema original con restricciones se puede escribir de forma irrestricta, ya que, si

$$0 = Ax - b = [B \ N] \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - b = Bx_1 + Nx_2 - b$$

se tiene que $x_1 = B^{-1}(b - Nx_2)$. En este caso:

- El $Nu(A) = \{d : Ad = 0\} = \{(-B^{-1}Nd_2, d_2) : d_2 \in \mathbb{R}^{n-m}\}$.

- el problema original se puede reescribir como

$$\text{Minimizar } f(B^{-1}(b - Nx_2), x_2) \text{ sujeto a } x_2 \in \mathbb{R}^{n-m}. \quad (44)$$

Si llamamos $g(x_2) = f(B^{-1}(b - Nx_2), x_2)$ en (44) sabemos que un punto estacionario x_2^* cumple:

$$0 = \nabla g(x_2^*) = \nabla_{x_1} f(B^{-1}(b - Nx_2^*), x_2^*)^T (-B^{-1}N) + \nabla_{x_2} f(B^{-1}(b - Nx_2^*), x_2^*).$$

Es decir, para cualquier $d_2 \in \mathbb{R}^{n-m}$, como $\nabla g(x_2^*)^T d_2 = 0$:

$$\nabla_{x_1} f(B^{-1}(b - Nx_2^*), x_2^*)^T (-B^{-1}N) d_2 + \nabla_{x_2} f(B^{-1}(b - Nx_2^*), x_2^*)^T d_2 = 0$$

y se tiene que

$$0 = \nabla_{x_1} f(x_1^*, x_2^*)^T (-B^{-1}N) d_2 + \nabla_{x_2} f(x_1^*, x_2^*)^T d_2 = \nabla f(x^*)^T \begin{pmatrix} -B^{-1}N d_2 \\ d_2 \end{pmatrix}$$

es decir

$$\nabla f(x^*)^T d = 0 \text{ para cualquier } d \in Nu(A).$$

Afortunadamente, esta idea de “poder invertir” un sistema de ecuaciones lineales no cuadrado se generaliza, bajo hipótesis adecuadas, a restricciones no lineales.

Para eso, recordemos la siguiente una versión particular del Teorema de la Función Implícita y algunos resultados conocidos para matrices.

Teorema 4.3. *Teorema de la Función Implícita.*

Sea $h : \mathbb{R}^{n+1} \times \mathbb{R}^n$ dada por $h(z, t) = \begin{pmatrix} h_1(z, t) \\ \vdots \\ h_n(z, t) \end{pmatrix}$ una función diferenciable en un entorno de un punto $(z^*, 0)$ tal que

1. $h(z^*, 0) = 0$ para algún $z^* \in \mathbb{R}^n$,
2. $\frac{\partial(h_1, \dots, h_n)}{\partial(z_1, \dots, z_n)}(z^*, 0)$ es no singular.

Entonces existen entornos de z^* y de $t = 0$ y una función diferenciable $z(t)$ definida en forma implícita tal que $h(z(t), t) = 0$ y la derivada de $z(t)$ se determina por derivación implícita.

Lema 4.1. Sea $A \in \mathbb{R}^{m \times n}$. Si A tiene rango m entonces $AA^T \in \mathbb{R}^{m \times m}$ es definida positiva.

Lema 4.2. Sea $A \in \mathbb{R}^{m \times n}$. Entonces $Nu(A) = R(A^T)^\perp$.

Teorema 4.4. Si $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\}$ son linealmente independientes entonces

$$T_{x^*} = Nu(Jh(x^*)) = \{d \in \mathbb{R}^n : \nabla h_i(x^*)^T d = 0, i = 1, \dots, m\}.$$

Demostración. Vimos que la contención $T_{x^*} \subset \{d \in \mathbb{R}^n : \nabla h_i(x^*)^T d = 0, i = 1, \dots, m\}$ vale siempre. Veamos la otra contención.

Sea d tal que $\nabla h_i(x^*)^T d = 0, i = 1, \dots, m$. Consideramos el sistema de ecuaciones

$$F(u, t) = h(x^* + td + Jh(x^*)^T u) = 0$$

de m ecuaciones con $m + 1$ incógnitas. Tenemos que

1. $F \in \mathcal{C}^1$,
2. $F(0, 0) = h(x^*) = 0$,
3. $\frac{\partial(F_1, \dots, F_m)}{\partial(u_1, \dots, u_m)}(0, 0) = Jh(x^*)Jh(x^*)^T$.

Por hipótesis sabemos que $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\}$ son linealmente independientes, luego, por el Lema 4.1 $\frac{\partial(F_1, \dots, F_m)}{\partial(u_1, \dots, u_m)}(0, 0) = Jh(x^*)Jh(x^*)^T$ es no singular y podemos aplicar el Teorema de la Función Implícita. Luego, existen entornos de $t = 0$ y de $u = 0$ y una única función $u(t)$ diferenciable tal que $u(0) = 0$ y $F(u(t), t) = h(x^* + td + Jh(x^*)^T u(t)) = 0$ en los entornos. Definimos la curva factible diferenciable

$$r(t) = x^* + td + Jh(x^*)^T u(t).$$

Tenemos que $r(0) = x^*$, $r'(t) = d + Jh(x^*)^T u'(t)$ y se tiene que $r'(0) = d + Jh(x^*)^T u'(0)$. Luego, $r'(0) = d$ (como queremos probar) si y solo si $Jh(x^*)^T u'(0) = 0$. Veamos entonces que efectivamente se tiene que $u'(0) = 0$.

Tenemos que $h(x^* + td + Jh(x^*)^T u(t)) = 0$. Derivando respecto de t y evaluando en $t = 0$ obtenemos las igualdades

$$\begin{aligned} Jh(x^* + td + Jh(x^*)^T u(t))(d + Jh(x^*)^T u'(t)) &= 0 \\ Jh(x^*)(d + Jh(x^*)^T u'(0)) &= 0 \end{aligned}$$

que equivale a

$$Jh(x^*)d + Jh(x^*)Jh(x^*)^T u'(0) = 0.$$

Luego, como $\nabla h_i(x^*)^T d = 0, i = 1, \dots, m$ se tiene que

$$Jh(x^*)Jh(x^*)^T u'(0) = 0$$

pero, como $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\}$ son linealmente independientes la matriz $Jh(x^*)Jh(x^*)^T$ es no singular lo que implica que $u'(0) = 0$ como queríamos demostrar. \square

A partir de esta propiedad podemos demostrar el siguiente Teorema.

Teorema 4.5. *Si x^* es un minimizador local de (40) tal que $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\}$ son linealmente independientes entonces existen únicos escalares $\lambda_1, \dots, \lambda_m$ tales que*

$$-\nabla f(x^*) = \sum_{i=1}^m \lambda_i \nabla h_i(x^*). \quad (45)$$

Demostración. Queremos probar que la igualdad (45), es decir

$$-\nabla f(x^*) = \left(\begin{array}{ccc} \frac{\partial h_1(x^*)}{\partial x_1} & \cdots & \frac{\partial h_m(x^*)}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_1(x^*)}{\partial x_n} & \cdots & \frac{\partial h_m(x^*)}{\partial x_n} \end{array} \right) \left(\begin{array}{c} \lambda_1 \\ \vdots \\ \lambda_m \end{array} \right) = (Jh(x^*))^T \lambda.$$

Como $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\}$ son linealmente independientes tenemos que $Jh(x^*)$ tiene rango m entonces la matriz $Jh(x^*)(Jh(x^*))^T \in \mathbb{R}^{m \times m}$ es no singular. Luego, como $Jh(x^*) \in \mathbb{R}^{m \times n}$, $Nu(Jh(x^*)) \subset \mathbb{R}^n$ y se sabe que, por el lema 4.2, $\mathbb{R}^n = Nu(Jh(x^*)) \oplus R((Jh(x^*))^T)$. Como $-\nabla f(x^*) \in \mathbb{R}^n$ podemos

escribir $-\nabla f(x^*) = u + v$ con u, v únicos tales que $u \in Nu(Jh(x^*))$, $v \in R((Jh(x^*))^T)$. Luego, como $u \in Nu(Jh(x^*)) = T_{x^*}$ y x^* es minimizador local se tiene que $\nabla f(x^*)^T u = 0$. Entonces,

$$-\nabla f(x^*)^T u = u^T u + u^T v = \|u\|^2$$

ya que $u \in Nu(Jh(x^*))$, $v \in R((Jh(x^*))^T)$. Luego, debe ser $u = 0$ y $-\nabla f(x^*) = v \in R((Jh(x^*))^T)$ como queríamos demostrar. \square

Definición 4.6. Sea x^* un punto factible.

1. Decimos que x^* es un punto regular cuando los gradientes $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\}$ son linealmente independientes. En este caso decimos que x^* verifica la condición de regularidad.
2. Definimos $T^{lin}(x^*) = Nu(Jh(x^*)) = \{d \in \mathbb{R}^n : \nabla h_i(x^*)^T d = 0, i = 1, \dots, m\}$ como cono tangente linealizado en x^* .
3. La condición (45) se denomina condición de Lagrange y los escalares $\lambda_1, \dots, \lambda_m$ se denominan multiplicadores de Lagrange.
4. Una condición de calidad es una condición sobre los puntos factibles que, cuando es verificada por un minimizador local implica la condición de Lagrange.

Observación. Regularidad es una condición de calidad de primer orden.

Ahora nos podríamos preguntar: Qué sucede en problemas donde la solución no es regular?

Ejemplos.

(1) Consideremos el problema

$$\text{Minimizar } x^2 + y^2 \text{ sujeto a } y - 1 = 0, y - x^2 - 1 = 0.$$

Este problema tiene un único punto factible $x^* = (0, 1)$ que es la solución del problema. Si definimos $h_1(x, y) = y - 1$, $h_2(x, y) = y - x^2 - 1$, se tiene que

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \nabla h_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla h_2(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

luego, los gradientes de las restricciones son linealmente dependientes, es decir x^* no es regular, pero existen infinitos λ_1, λ_2 que verifican la condición de Lagrange ($\lambda_1 + \lambda_2 = -2$).

(2) Consideremos el problema

$$\text{Minimizar } x^2 + (y + 1)^2 \text{ sujeto a } x + y = 0, (x - 1)^2 + (y - 1)^2 = 2.$$

Nuevamente este problema tiene un único punto factible $x^* = (0, 0)$ que es la solución del problema. Si definimos $h_1(x, y) = x + y$, $h_2(x, y) = (x - 1)^2 + (y - 1)^2 - 2$, se tiene que

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \nabla h_1(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla h_2(x^*) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

luego, los gradientes de las restricciones son linealmente dependientes, es decir x^* no es regular, pero no existen escalares λ_1, λ_2 que verifican la condición de Lagrange.

Observación: Se ha demostrado que bajo la regularidad se tiene que

$$T_{x^*} = T^{lin}(x^*)$$

y gracias a esto se cumple la condición de Lagrange en la solución.

Luego, teniendo esto en cuenta se puede obtener el siguiente resultado.

Teorema 4.7. *Si x^* es un minimizador local del problema (40) y cumple $T_{x^*} = T^{lin}(x^*)$ entonces existen escalares $\lambda_1, \dots, \lambda_m$ tales que se verifica la condición (45).*

Demostración.

Tenemos que $T_{x^*} = T^{lin}(x^*) = Nu(Jh(x^*))$. Luego, por el Teorema 4.2 sabemos que $-\nabla f(x^*) \in (T_{x^*})^\perp = (Nu(Jh(x^*)))^\perp = R(Jh(x^*)^T)$ lo que demuestra el teorema. \square

Definición 4.8. *La condición $T_{x^*} = T^{lin}(x^*)$ se denomina condición de Kuhn-Tucker y es una condición de calidad, como muestra la propiedad anterior.*

Observaciones

- Si vale independencia lineal entonces vale la condición de Kuhn-Tucker $T_{x^*} = T^{lin}(x^*)$.
- La recíproca no vale. Alcanza con repetir una restricción que independencia lineal no se verifica.

Se define la función

$$l(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) \quad (46)$$

y se denomina función de Lagrange asociada al problema (40). Por lo que vimos anteriormente tenemos que, si x^* es un minimizador local de (40) y cumple alguna condición de calidad (regularidad por ejemplo) entonces existe $\lambda^* \in \mathbb{R}^m$ tal que $\nabla_x l(x^*, \lambda^*) = 0$. Luego,

$$\nabla_x l(x^*, \lambda^*) = 0$$

$$h(x^*) = 0,$$

es decir, el par (x^*, λ^*) es un punto estacionario de la función de Lagrange.

Tanto regularidad como Kuhn-Tucker son condiciones de calidad, es decir, condiciones que implican la existencia de multiplicadores de Lagrange en una solución (minimizador local). Existen en la literatura otras condiciones de calidad entre estas dos.

4.2 Condiciones necesarias de primer orden para problemas con restricciones de desigualdad

Consideramos problemas de la forma

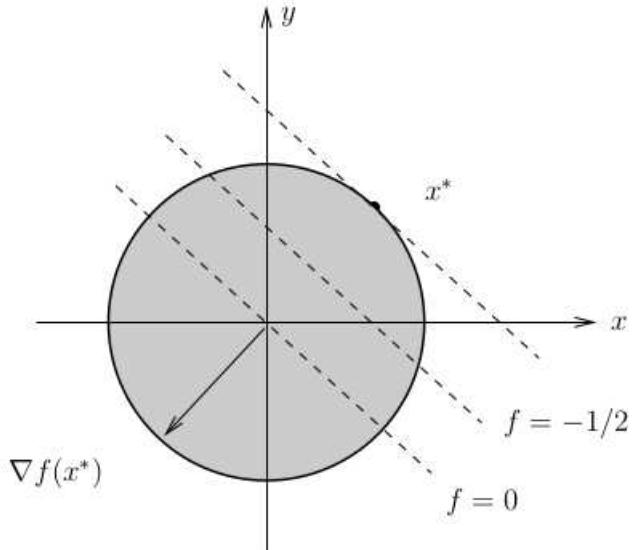
$$\text{Minimizar } f(x) \text{ sujeto a } g(x) \leq 0 \quad (47)$$

donde las funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ son todas diferenciables hasta el orden que sea necesario. La función g es de la forma $g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{pmatrix}$ donde $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$.

Consideremos por ejemplo el problema

$$\text{Minimizar } -x - y \text{ sujeto a } x^2 + y^2 \leq 1.$$

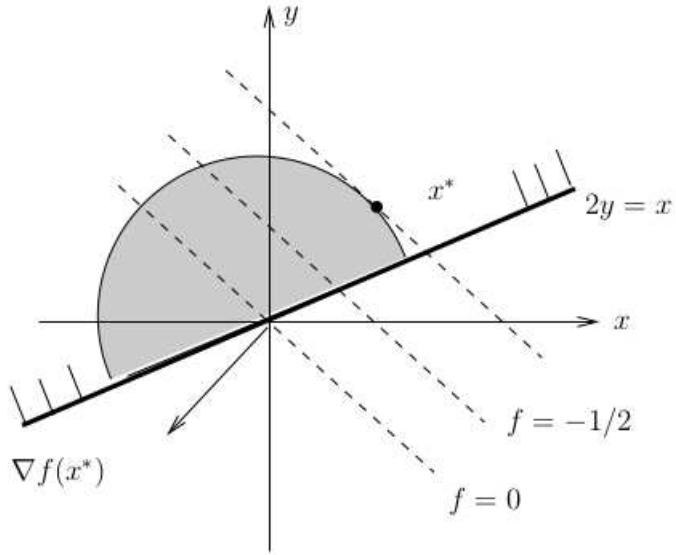
Gráficamente podemos observar que la solución del problema es, como encontramos para el problema con restricciones de igualdad, el punto $x^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.



Si consideramos el problema

$$\text{Minimizar } -x - y \text{ sujeto a } x^2 + y^2 \leq 1, 2y \geq x$$

vemos gráficamente que la solución sigue siendo la misma. Es decir, la restricción $2y \geq x$ no interfiere en el cálculo del minimizador local.



En este caso decimos que la restricción $g_2(x, y) = x - 2y$ es inactiva en x^* . Formalmente:

Definición 4.9. Dado un punto factible x , definimos el conjunto de restricciones de desigualdad que son activas en x :

$$A(x) = \{i \in \{1, \dots, p\} : g_i(x) = 0\}$$

Si $i \notin A(x)$ entonces decimos que esa restricción es inactiva en x . Es el caso de la restricción $2y \geq x$ en x^* en el ejemplo anterior.

Si $g_i(x) > 0$ decimos que esa restricción es violada en x .

Si consideramos el problema

$$\text{Minimizar } -x - y \text{ sujeto a } x^2 + y^2 \leq 1, y \geq 2x$$

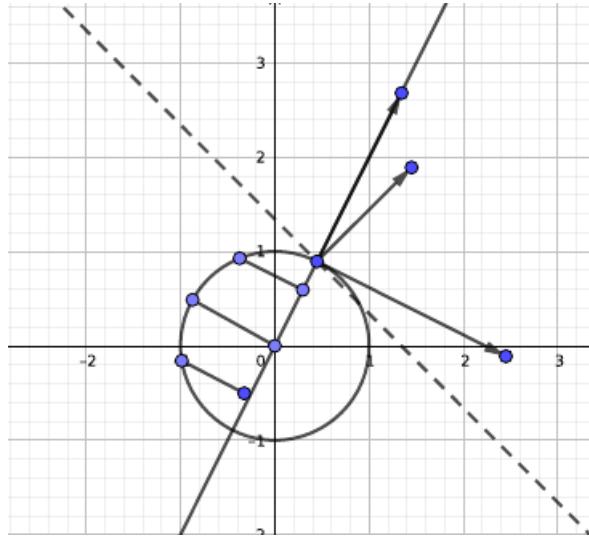
el punto $x^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ ya no es más la solución. Ahora la solución se alcanza en $x^* = (\frac{1}{\sqrt{5}}, 2\frac{1}{\sqrt{5}})$ y las dos restricciones son activas en ese punto. Si llamamos $g_1(x, y) = x^2 + y^2 - 1$, $g_2(x, y) = 2x - y$ tenemos que los gradientes en x^* son los vectores:

$$\nabla f(x^*) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \nabla g_1(x^*) = \begin{pmatrix} 2\frac{1}{\sqrt{5}} \\ 4\frac{1}{\sqrt{5}} \end{pmatrix}, \quad \nabla g_2(x^*) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

y se tiene que

$$-\nabla f(x^*) = \mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) \quad \text{para } \mu_1 = \frac{3\sqrt{5}}{10}, \mu_2 = \frac{1}{5}.$$

Observar que $\mu_1 > 0, \mu_2 > 0$.



Proposición 4.1. Si x^* es un minimizador del problema (47) entonces x^* es un minimizador del problema con restricciones de igualdad

$$\text{Minimizar } f(x) \text{ sujeto a } g_i(x) = 0, i \in A(x^*). \quad (48)$$

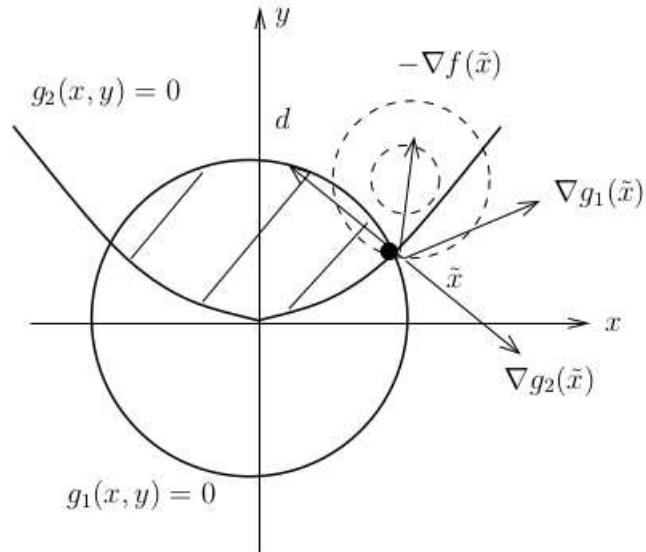
Luego, tenemos que, si los gradientes $\{\nabla g_i(x^*)\}_{i \in A(x^*)}$ son linealmente independientes entonces existen escalares λ_i tales que

$$-\nabla f(x^*) = \sum_{i \in A(x^*)} \lambda_i \nabla g_i(x^*). \quad (49)$$

Veremos que, como las restricciones son de desigualdad, nos importa el signo de los multiplicadores. Consideremos el problema

$$\text{Minimizar } (x - 1)^2 + (y - 1)^2 \text{ sujeto a } x^2 + y^2 \leq 1, 2y \geq x^2.$$

Observar el punto \tilde{x} en la figura.



Vemos que, en ese punto se tiene que

$$-\nabla f(\tilde{x}) = \mu_1 \nabla g_1(\tilde{x}) + \mu_2 \nabla g_2(\tilde{x}) \quad \text{con } \mu_1 > 0, \mu_2 < 0.$$

Sea d una dirección tal que $\nabla g_1(\tilde{x})^T d = 0, \nabla g_2(\tilde{x})^T d < 0$. Luego,

$$-\nabla f(\tilde{x})^T d = \mu_1 \nabla g_1(\tilde{x})^T d + \mu_2 \nabla g_2(\tilde{x})^T d > 0$$

es decir, si abandonamos g_2 moviéndonos en forma ortogonal al $\nabla g_1(\tilde{x})$ podemos hacer decrecer el valor de f . Esto muestra que no alcanza con ver la verificación de (49) sino que además es importante analizar el signo de los multiplicadores. Un multiplicador negativo dice que el punto no es solución.

Teorema 4.10. *Si x^* es un minimizador local de (47) tal que $\{\nabla g_i(x^*)\}_{i \in A(x^*)}$ son linealmente independientes entonces existen únicos escalares $\{\mu_i\}_{i \in A(x^*)}$ tales que $\mu_i \geq 0$*

$$-\nabla f(x^*) = \sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*). \quad (50)$$

Para demostrar este teorema necesitamos un resultado previo.

Analicemos qué sucede con curvas factibles. Sea $r(t)$ una curva factible diferenciable tal que $r(0) = x^*$. Si $i \in A(x^*)$ entonces utilizando el desarrollo de Taylor

$$g_i(r(t)) = g_i(x^*) + \nabla g_i(x^*)^T r'(0)t + o(t)$$

entonces, por la factibilidad $g_i(r(t)) \leq 0$ y

$$\nabla g_i(x^*)^T r'(0)t + o(t) \leq 0$$

dividiendo por $t > 0$ y haciendo t tender a cero obtenemos que

$$\nabla g_i(x^*)^T r'(0) \leq 0.$$

Luego, en el caso de restricciones de desigualdad tenemos que

$$T_{x^*} \subset \{d \in \mathbb{R}^n : \nabla g_i(x^*)^T d \leq 0, i \in A(x^*)\}.$$

Definimos

$$T^{lin}(x^*) = \{d \in \mathbb{R}^n : \nabla g_i(x^*)^T d \leq 0, i \in A(x^*)\}$$

como tangente linealizado en x^* para problemas con restricciones de desigualdad.

Como sucede en el caso de restricciones de igualdad, se puede demostrar que si $\{\nabla g_i(x^*)\}_{i \in A(x^*)}$ son linealmente independientes entonces $T^{lin}(x^*) \subset T_{x^*}$.

Demostración del teorema 4.10.

Sabemos que x^* es un minimizador local del problema (48), luego, por hipótesis existen μ_i tales que

$$-\nabla f(x^*) = \sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*).$$

Supongamos que algún $\mu_j < 0$. Por la independencia lineal existe $d \neq 0$ tal que $\nabla g_i(x^*)^T d = 0, \forall i \neq j, \nabla g_j(x^*)^T d < 0$. Entonces

$$-\nabla f(x^*)^T d = \mu_j \nabla g_j(x^*)^T d > 0.$$

Como $d \in T^{lin}(x^*) = T_{x^*}$ tenemos que existe $r(t)$ factible diferenciable tal que $r(0) = x^*, r'(0) = d$. Luego, usando Taylor,

$$\lim_{t \rightarrow 0} \frac{f(r(t)) - f(x^*)}{t} = \nabla f(x^*)^T d < 0$$

lo que implica que existe $t > 0$ tal que $f(r(t)) < f(x^*)$ absurdo. Luego, debe ser $\mu_i \geq 0, \forall i \in A(x^*)$. \square

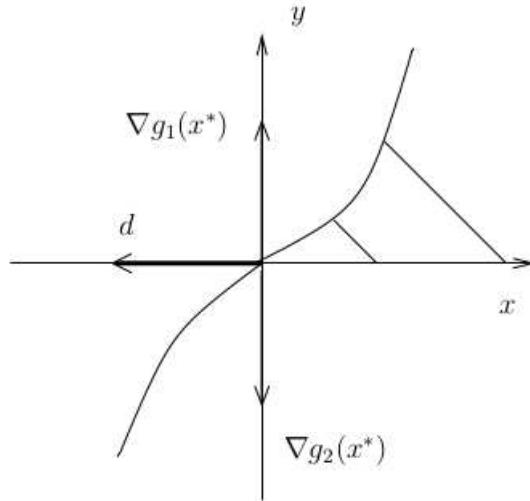
Definición 4.11. 1. La condición de que los gradientes $\{\nabla g_i(x^*)\}_{i \in A(x^*)}$ sean linealmente independientes se denomina regularidad. Es una condición de calidad para el problema con restricciones de desigualdad.

2. Las condiciones

$$\begin{cases} -\nabla f(x^*) = \sum_{i=1}^p \mu_i \nabla g_i(x^*) \\ \mu_i \geq 0, g_i(x^*) \leq 0, \mu_i g_i(x^*) = 0, i = 1, \dots, p \end{cases}$$

se denominan condiciones de Karush-Kuhn-Tucker (KKT). La segunda condición se llama condición de complementariedad. Los escalares μ_i se denominan multiplicadores de Lagrange.

La inclusión $T^{lin}(x^*) \subset T_{x^*}$ no vale en general: consideremos las restricciones de desigualdad $g_1(x, y) = y - x^3, g_2(x, y) = -y$ en el punto factible $x^* = (0, 0)$.



Los gradientes son linealmente dependientes en x^* y se tiene que $d = (-1, 0) \in T^{lin}(x^*)$ pero $d \notin T_{x^*}$.

Problemas de optimización con restricciones mixtas

Si consideramos el problema general de la forma

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0, g(x) \leq 0 \quad (51)$$

donde las funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}, h : \mathbb{R}^n \rightarrow \mathbb{R}^m, g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ son todas diferenciables hasta el orden que sea necesario, podemos generalizar los resultados obtenidos teniendo en cuenta las siguientes definiciones.

Definición 4.12. Dado un punto factible x^* ,

1. Decimos que x^* es regular si los gradientes $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\} \cup \{\nabla g_i(x^*)_{i \in A(x^*)}\}$ son linealmente independientes.
2. Definimos el tangente linealizado en x^* al conjunto $T^{lin}(x^*) = \{d \in \mathbb{R}^n : \nabla h_i(x^*)^T d = 0, i = 1, \dots, m; \nabla g_i(x^*)^T d \leq 0, i \in A(x^*)\}$.
3. Decimos que x^* cumple la condición de Kuhn-Tucker si $T_{x^*} = T^{lin}(x^*)$.
4. La función $l(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{i=1}^p \mu_i g_i(x)$ se denomina función de Lagrange.

Proposición 4.2. Si x^* es regular entonces x^* cumple la condición de Kuhn-Tucker.

Teorema 4.13. Condiciones KKT. Si x^* es un minimizador local regular de (51) entonces existen únicos escalares $\lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_p$ tales que

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{i=1}^p \mu_i \nabla g_i(x^*) = 0 \\ \mu_i \geq 0, g_i(x^*) \leq 0, \mu_i g_i(x^*) = 0. \end{cases} \quad (52)$$

El sistema (52) se denomina sistema KKT.

Decimos que una solución de este sistema es un punto KKT o punto estacionario.

4.3 El abordaje de conos

En las secciones anteriores aparecen conjuntos que permiten definir una teoría mucho más rica relacionada con las condiciones de calidad.

Consideramos el problema general con restricciones de igualdad y de desigualdad (51). Denotemos por $\Omega = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$ al conjunto factible.

La condición (52) es equivalente a:

$$-\nabla f(x^*) \in \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{i=1}^p \mu_i \nabla g_i(x^*), \mu_i \geq 0, \mu_i g_i(x^*) = 0 \right\}. \quad (53)$$

El conjunto

$$G(x) = \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{i=1}^p \mu_i \nabla g_i(x), \mu_i \geq 0, \mu_i g_i(x) = 0 \right\}$$

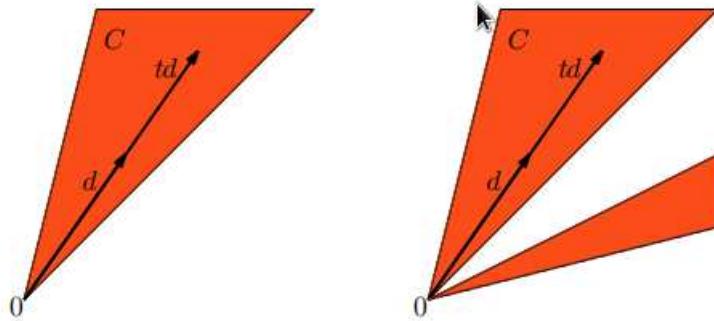
es llamado cono generado por los gradientes de las restricciones en x .

El conjunto $T^{lin}(x) = \{d \in \mathbb{R}^n : \nabla h_i(x)^T d = 0, i = 1, \dots, m; \nabla g_i(x)^T d \leq 0, i \in A(x)\}$ es el cono tangente linealizado en x .

Hablemos un poco sobre conos.

Definición 4.14. Un subconjunto no vacío $C \subset \mathbb{R}^n$ es un cono si, para todo $t \geq 0$ y $d \in C$, se cumple que $td \in C$.

Informalmente, un cono es un conjunto de direcciones. Note que el vector nulo pertenece a cualquier cono. Además, un cono es un conjunto no acotado. En la figura vemos dos ejemplos de conos, uno convexo y otro no convexo.



Por ejemplo, dada una matriz $B \in \mathbb{R}^{n \times m}$, el conjunto

$$C = \{By \mid y \in \mathbb{R}^m, y \geq 0\}$$

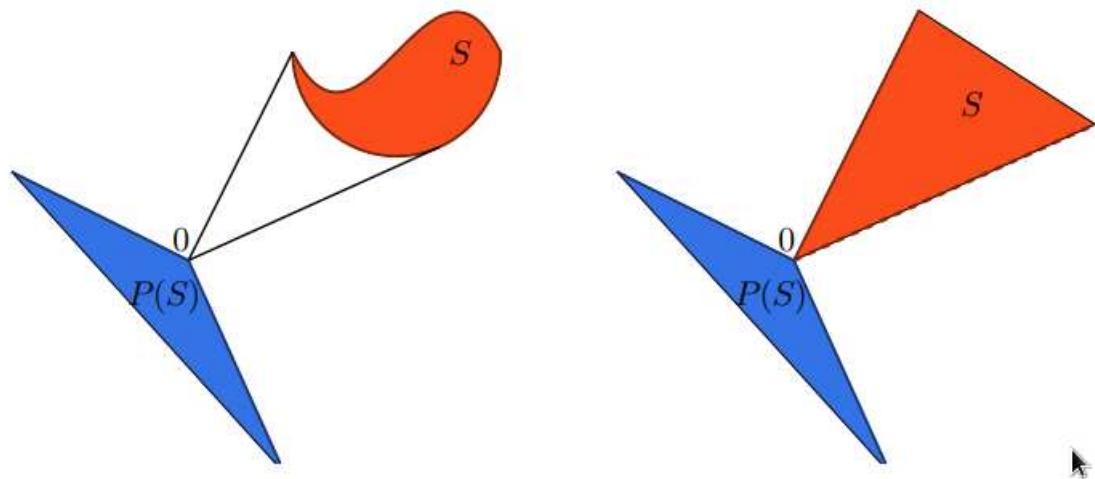
es un cono convexo.

Un ejemplo de cono que es muy útil en optimización es el cono polar, que en \mathbb{R}^2 o \mathbb{R}^3 puede ser caracterizado por los vectores que forman un ángulo mayor o igual a $\pi/2$ con los elementos de un conjunto dado.

Definición 4.15. Dado un conjunto $S \subset \mathbb{R}^n$, definimos el polar de S por

$$S^\circ = \{p \in \mathbb{R}^n \mid p^T x \leq 0, \forall x \in S\}.$$

La siguiente figura ilustra el polar de algunos conjuntos (el polar es el conjunto en azul).



Se tienen las siguientes propiedades:

Proposición 4.3. 1. Dado $S \subset \mathbb{R}^n$, S° es un cono, convexo y cerrado.

2. Sea $C \subset \mathbb{R}^n$ un cono convexo cerrado. Para $z \in \mathbb{R}^n$ sea $\bar{z} = P_C(z)$. Entonces,

$$z - \bar{z} \in \{\bar{z}\}^\perp \cap C^\circ.$$

No siempre el polar del polar da el polar. Se tiene el siguiente resultado:

Lema 4.3. Dado $S \subset \mathbb{R}^n$, se tiene que

$$S \subset (S^\circ)^\circ.$$

Demostración.

Considere $x \in S$ y $C = S^\circ$. Dado $d \in C$, tenemos

$$x^T d \leq 0.$$

Luego,

$$x \in C^\circ = (S^\circ)^\circ,$$

como queríamos demostrar. □

Básicamente, hay tres motivos que impiden la igualdad entre el conjunto y el polar del polar: el hecho de no ser un cono, no ser convexo o no ser cerrado. El clásico Lema de Farkas, presentado a continuación, garantiza la igualdad.

Lema 4.4. Farkas geométrico Si $C \subset \mathbb{R}^n$ es un cono convexo cerrado no vacío, entonces

$$(C^\circ)^\circ = C.$$

Demostración.

Por el Lema 4.3, basta mostrar que

$$(C^\circ)^\circ \subset C.$$

Sea $z \in (C^\circ)^\circ$ y $\bar{z} = P_C(z)$. Vamos a probar que $z = \bar{z}$. Por 2. de la Proposición 4.3, tenemos que $(z - \bar{z}) \in C^\circ$. Como $z \in (C^\circ)^\circ$, vale que

$$(z - \bar{z})^T z \leq 0.$$

Usando 2. de la Proposición 2.1($(z - \bar{z})^T (y - \bar{z}) \leq 0 \forall y \in C$), obtenemos que

$$\|z - \bar{z}\|^2 = (z - \bar{z})^T z - (z - \bar{z})^T \bar{z} = (z - \bar{z})^T z \leq 0,$$

lo que implica que $z = \bar{z} \in C$ como queríamos demostrar. \square .

El Lema de Farkas tiene una versión muy conocida en optimización, conocida como versión algebraica:

Lema 4.5. Farkas algebraico. Sean $A \in \mathbb{R}^{m \times n}$ y $c \in \mathbb{R}^n$. Entonces, exactamente uno de los dos sistemas siguientes tiene solución:

$$\begin{cases} Ax \leq 0 & y \quad c^T x > 0, \\ A^T y = c & y \quad y \geq 0. \end{cases}$$

Demostración. Si el sistema

$$A^T y = c, \quad y \geq 0$$

tiene solución, entonces $c = A^T y$ con $y \geq 0$. Así, dado $x \in \mathbb{R}^n$ tal que $Ax \leq 0$, tenemos

$$c^T x = y^T A x \leq 0,$$

lo que implica que el sistema

$$Ax \leq 0, \quad c^T x > 0$$

no tiene solución.

Suponga ahora que el sistema

$$A^T y = c, \quad y \geq 0$$

no tiene solución. Por lo tanto,

$$c \notin C = \{A^T y \mid y \geq 0\}.$$

Como C es un cono convexo cerrado, se tiene $C = (C^\circ)^\circ$. Luego,

$$c \notin (C^\circ)^\circ,$$

lo que significa que existe $x \in C^\circ$ tal que

$$c^T x > 0.$$

Además, como $x \in C^\circ$ y $A^T y \in C$ si $y \geq 0$ se tiene que

$$(Ax)^T y = x^T A^T y \leq 0,$$

para todo $y \geq 0$. En particular, tomando $y = e_j, j = 1, \dots, m$, obtenemos $Ax \leq 0$. Así, el sistema

$$Ax \leq 0, \quad c^T x > 0$$

tiene solución. \square

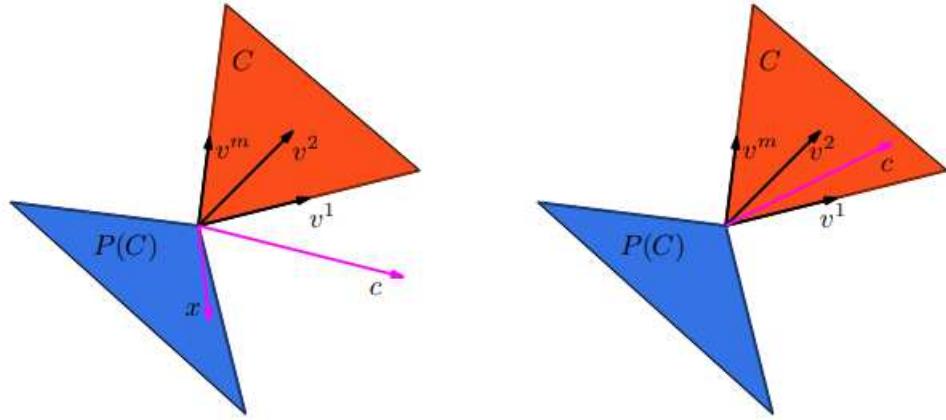
Aunque es una versión algebraica, el Lema puede ser interpretado geométricamente. Los vectores $v_1, v_2, \dots, v_m \in \mathbb{R}^n$ son las filas de A . En la ilustración del lado izquierdo tenemos el caso en que el sistema

$$Ax \leq 0, \quad c^T x > 0$$

tiene solución. En el lado derecho,

$$A^T y = c, \quad y \geq 0$$

tiene solución.

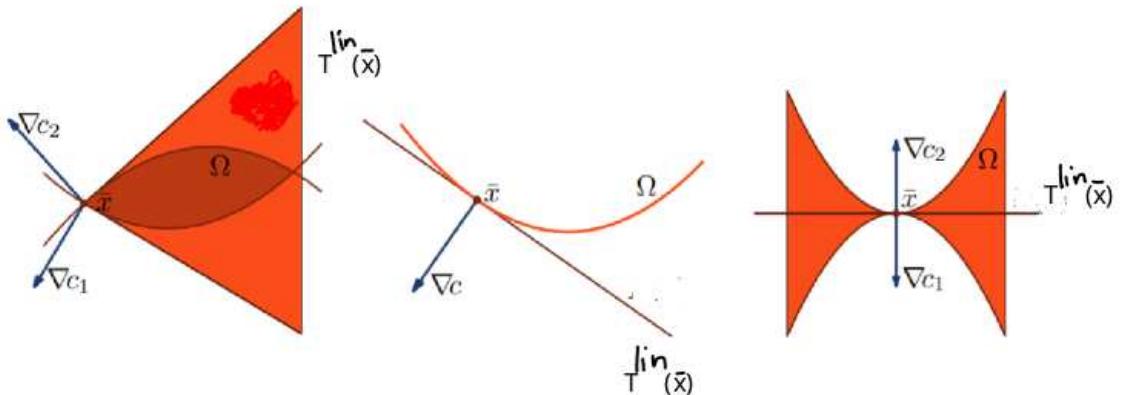


Retomemos ahora con los conos que aparecen en las condiciones de optimalidad de problemas generales de optimización no lineal.

El cono tangente linealizado

$$T^{lin}(x) = \{d \in \mathbb{R}^n : \nabla h_i(x)^T d = 0, i = 1, \dots, m; \nabla g_i(x)^T d \leq 0, i \in A(x)\} \quad (54)$$

En la siguiente figura vemos algunos ejemplos que surgen cuando consideramos el cono tangente linealizado $T^{lin}(x)$. En la primera, tenemos desigualdades activas y los gradientes son linealmente independientes. Esto le da cierta regularidad al conjunto Ω , en el sentido de que está "bien" aproximado por $T^{lin}(\bar{x})$ en un entorno de \bar{x} . En la segunda, tenemos una igualdad y también podemos decir que $T^{lin}(\bar{x})$ es una buena aproximación para Ω . Sin embargo, la última situación muestra un caso donde el cono $T^{lin}(\bar{x})$ es una recta, pero el conjunto factible es una región del plano. Note que, en este caso, los gradientes activos son linealmente dependientes.

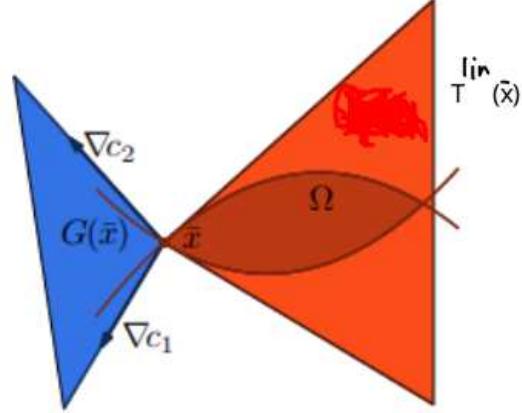


Se puede demostrar que el cono tangente linealizado es un cono convexo cerrado.

El cono generado por los gradientes de las restricciones

$$G(x) = \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{i=1}^p \mu_i \nabla g_i(x), \mu_i \geq 0, \mu_i g_i(x) = 0 \right\} \quad (55)$$

Observar la relación entre G y T^{lin} en el siguiente ejemplo:



Lema 4.6. Dado $\bar{x} \in \Omega$, se tiene que

$$T^{lin}(\bar{x}) = (G(\bar{x}))^\circ.$$

Demostación.

Si $m = 0$ y $A(\bar{x}) = \emptyset$ en \bar{x} se tiene $T^{lin}(\bar{x}) = \mathbb{R}^n$ y $G(\bar{x}) = \{0\}$ y vale la igualdad. Si no, sea $d \in T^{lin}(\bar{x})$ y $s \in G(\bar{x})$. Entonces,

$$d^T s = \sum_{i=1}^m \lambda_i d^T \nabla h_i(\bar{x}) + \sum_{i \in A(\bar{x})} \mu_i d^T \nabla g_i(\bar{x}).$$

Como $d \in T^{lin}(\bar{x})$, se cumple que $d^T \nabla h_i(\bar{x}) = 0$ para todo $i = 1, \dots, m$ y $d^T \nabla g_i(\bar{x}) \leq 0$ para todo $i \in A(\bar{x})$. Luego, como $\mu_i \geq 0$ se tiene que

$$d^T s \leq 0,$$

y, $d \in (G(\bar{x}))^\circ$.

Para probar la inclusión contraria, sea $d \in (G(\bar{x}))^\circ$. Entonces,

$$d^T s \leq 0, \quad \forall s \in G(\bar{x}).$$

En particular, para $i = 1, \dots, m$, tenemos que $\nabla h_i(\bar{x})$ y $-\nabla h_i(\bar{x})$ son elementos de $G(\bar{x})$. Por lo tanto,

$$d^T \nabla h_i(\bar{x}) \leq 0 \quad \text{y} \quad d^T (-\nabla h_i(\bar{x})) \leq 0,$$

de donde se sigue que

$$d^T \nabla h_i(\bar{x}) = 0.$$

Además, para $i \in A(\bar{x})$, tenemos $\nabla g_i(\bar{x}) \in G(\bar{x})$ y,

$$d^T \nabla g_i(\bar{x}) \leq 0.$$

De esta forma, $d \in T^{lin}(\bar{x})$. □

Como $G(x)$ es un cono convexo cerrado, podemos asegurar que

$$(T^{lin}(x))^\circ = G(x). \quad (56)$$

El cono tangente

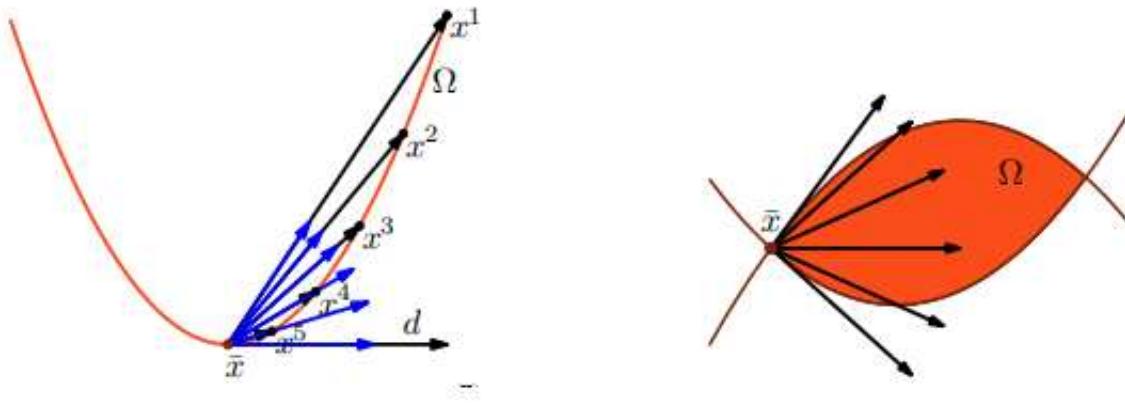
Existe además otro cono importante para el desarrollo de la teoría que también aproxima el conjunto viable Ω , pero a diferencia del cono $T^{lin}(\bar{x})$, que se basa en los gradientes de las restricciones y del espacio $T_{\bar{x}}$ que se basa en curvas, este nuevo cono considera los vectores que son tangentes a la región Ω que se pueden obtener a partir de sucesiones factibles.

Definición 4.16. Una dirección $d \in \mathbb{R}^n$ se dice tangente a $\Omega \subset \mathbb{R}^n$ en el punto $\bar{x} \in \Omega$ cuando $d = 0$ o existe una sucesión de puntos factibles $\{x_k\} \subset \Omega$ tal que $x_k \rightarrow \bar{x}$ y

$$\frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \rightarrow \frac{d}{\|d\|}.$$

Denotaremos a este cono como $T(\bar{x})$.

En la siguiente figura vemos algunos ejemplos. El conjunto factible de la izquierda es una curva definida por una restricción de igualdad, en la cual representamos una dirección tangente d . En la otra figura, el conjunto factible es una región determinada por dos restricciones de desigualdad.



Se tiene la siguiente relación:

Lema 4.7. Dado $\bar{x} \in \Omega$, se tiene que

$$T(\bar{x}) \subset T^{lin}(\bar{x}).$$

Demostración.

Sea $d \in T(\bar{x})$, con $d \neq 0$. Entonces existe una sucesión $\{x_k\} \subset \Omega$ tal que $x_k \rightarrow \bar{x}$ y

$$\frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \rightarrow \frac{d}{\|d\|}.$$

Por la diferenciabilidad de las restricciones, se tiene que para todo $i = 1, \dots, m; j \in A(\bar{x})$,

$$h_i(x_k) = h_i(\bar{x}) + \nabla h_i(\bar{x})^T(x_k - \bar{x}) + o(\|x_k - \bar{x}\|),$$

$$g_j(x_k) = g_j(\bar{x}) + \nabla g_j(\bar{x})^T(x_k - \bar{x}) + o(\|x_k - \bar{x}\|).$$

Como $x_k, \bar{x} \in \Omega$, se tiene que

$$\frac{\nabla h_i(\bar{x})^T(x_k - \bar{x})}{\|x_k - \bar{x}\|} + \frac{o(\|x_k - \bar{x}\|)}{\|x_k - \bar{x}\|} = 0,$$

y

$$\frac{\nabla g_j(\bar{x})^T(x_k - \bar{x})}{\|x_k - \bar{x}\|} + \frac{o(\|x_k - \bar{x}\|)}{\|x_k - \bar{x}\|} \leq 0.$$

Tomando el límite cuando $k \rightarrow \infty$, obtenemos

$$\nabla h_i(\bar{x})^T \frac{d}{\|d\|} = 0,$$

y

$$\nabla g_j(\bar{x})^T \frac{d}{\|d\|} \leq 0.$$

Es decir, $d \in T^{lin}(\bar{x})$. □

No siempre vale la igualdad entre los conos.

Ejemplo Consideremos las restricciones $h_1, g_1, g_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$h_1(x_1, x_2) = x_1 x_2, \quad g_1(x_1, x_2) = -x_1, \quad g_2(x_1, x_2) = -x_2.$$

en el punto $\bar{x} = 0$.

Tenemos

$$\nabla h_1(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla g_1(\bar{x}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \nabla g_2(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

$$T^{lin}(\bar{x}) = \{(d_1, d_2) \mid d_1 \geq 0, d_2 \geq 0\}, \quad G(\bar{x}) = \{(d_1, d_2) \mid d_1 \leq 0, d_2 \leq 0\},$$

y

$$T(\bar{x}) = \{(d_1, d_2) \mid d_1 \geq 0, d_2 \geq 0, d_1 \cdot d_2 = 0\}.$$

Veamos ahora como se relacionan estos conceptos con las condiciones KKT. Comencemos con un resultado que da una condición necesaria de optimalidad.

Lema 4.8. Si $x^* \in \Omega$ es un minimizador local del problema (51), entonces

$$\nabla f(x^*)^T d \geq 0, \quad \forall d \in T(x^*).$$

Demostración.

Sea $d \in T(x^*)$, $d \neq 0$. Entonces existe una sucesión $\{x_k\} \subset \Omega$ tal que $x_k \rightarrow x^*$ y

$$\frac{x_k - x^*}{\|x_k - x^*\|} \rightarrow \frac{d}{\|d\|}.$$

Por otro lado, tenemos

$$0 \leq f(x_k) - f(x^*) = \nabla f(x^*)^T (x_k - x^*) + o(\|x_k - x^*\|),$$

para todo k suficientemente grande. Dividiendo por $\|x_k - x^*\|$ y pasando al límite obtenemos

$$\nabla f(x^*)^T d \geq 0,$$

como queríamos demostrar. \square

Este es un resultado sumamente importante pero, de la misma forma que sucede cuando se utiliza el cono tangente basado en curvas, es poco práctico en el sentido de que no podemos usarlo para calcular posibles minimizadores. Sin embargo es muy útil para obtener el siguiente resultado.

Teorema 4.17. Sea $x^* \in \Omega$ un minimizador local del problema (51) y supongamos que

$$(T(x^*))^\circ = (T^{lin}(x^*))^\circ.$$

Entonces existen vectores λ^* y μ^* tales que

$$-\nabla f(x^*) = \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla g_i(x^*),$$

con

$$\mu_i^* \geq 0,$$

y

$$\mu_i^* g_i(x^*) = 0, \quad i = 1, \dots, p.$$

Demostración.

Por el Lema anterior, se tiene que

$$-\nabla f(x^*)^T d \leq 0, \quad \forall d \in T(x^*).$$

Así, usando la hipótesis y la relación (56), obtenemos

$$-\nabla f(x^*) \in (T(x^*))^\circ = (T^{lin}(x^*))^\circ = G(x^*).$$

Esto significa que existen vectores λ y μ , con $\mu \geq 0$, tales que

$$-\nabla f(x^*) = \sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*)$$

y se obtiene el resultado deseado. \square

La condición $(T(x^*))^\circ = (T^{lin}(x^*))^\circ$ fue introducida por Guignard [9] para dimensión infinita y reformulada para el caso finito por Gould y Tolle [10], donde se establece que esta condición, además de ser suficiente, también es necesaria para que se cumplan las condiciones KKT. Es la condición de calidad más débil que permite obtener las condiciones KKT.

Si bien teóricamente es un resultado muy importante, del punto de vista práctico puede ser muy difícil obtener los conos y verificar la igualdad. Es por este motivo que se desarrollan otras condiciones que no son tan débiles pero son mas simples de verificar analíticamente, ver [4].

4.4 Condiciones de calidad

Recordemos que una condición de calidad para un problema de optimización es una condición sobre los puntos factibles que, cuando es verificada por un minimizador local implica las existencia de los multiplicadores de Lagrange y garantiza que se verifican las condiciones Karush-Kuhn-Tucker.

Ya hemos presentado algunas que existen en la literatura:

- regularidad o independencia lineal de los gradientes de las restricciones activas;
- condición de Kuhn-Tucker (Definición 4.8);
- Guignard $(T(x^*))^\circ = (T^{lin}(x^*))^\circ$.

Si bien la condición de Guignard es la más débil que permite demostrar la existencia de los multiplicadores de Lagrange en un minimizador local, existen muchas otras condiciones de calidad en la literatura. Veremos las condiciones de Mangasarian-Fromovitz y de Rango constante.

Definición 4.18. *Mangasarian-Fromovitz, 1967 [11]. Decimos que un punto factible x^* del problema (51) verifica la condición de Mangasarian-Fromovitz si los gradientes*

$$\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\} \cup \{\nabla g_j(x^*)\}_{j \in A(x^*)}$$

son positivo linealmente independientes. Es decir, si

$$\sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*) = 0, \text{ con } \mu_j \geq 0 \text{ entonces } \lambda_i = 0 \text{ y } \mu_j = 0.$$

La definición anterior es en realidad una definición alternativa de la definición original dada por Mangasarian-Fromovitz que es conocida como definición dual.

En optimización es usual que una restricción de igualdad $h(x) = 0$ se reemplace por dos restricciones de desigualdad opuestas: $h(x) \leq 0, -h(x) \leq 0$. En este caso la condición de Mangasarian-Fromovitz ya no se verifica. Lo mismo sucede si en la descripción del conjunto factible aparecen dos restricciones repetidas.

Definición 4.19. *Condición de Rango Constante, 1984. [12] Decimos que un punto factible x^* del problema (51) verifica la condición de rango constante si existe un entorno $B(x^*, \epsilon)$ de x^* tal que, para todo*

$$I_0 \subseteq \{1, \dots, m\}, \quad J_0 \subseteq A(x^*),$$

el rango de los gradientes

$$\{\nabla h_i(x)\}_{i \in I_0} \cup \{\nabla g_j(x)\}_{j \in J_0}$$

es constante para todo $x \in B(x^, \epsilon)$.*

Basicamente lo que establece la condición de rango constante es que cualquier subconjunto de los gradientes de las restricciones de igualdad y de desigualdad activas que son LD en el punto continúan siendo LD en el entorno del punto.

Se tiene que:

- Regularidad implica Mangasarian-Fromovitz;
- Regularidad implica Rango constante.

Sin embargo las implicaciones son estrictas, existen ejemplos que muestran que Mangasarian-Fromovitz y rango constante son condiciones estrictamente más débiles que regularidad.

Se puede demostrar además que:

- Mangasarian-Fromovitz implica la condición de Kuhn-Tucker;
- rango constante implica la condición de Kuhn-Tucker;
- la condición de Kuhn-Tucker implica Guignard;

y que las reciprocas no valen. Ver el tratamiento en [8, 4, 1].

Se tiene el siguiente conjunto de implicaciones:

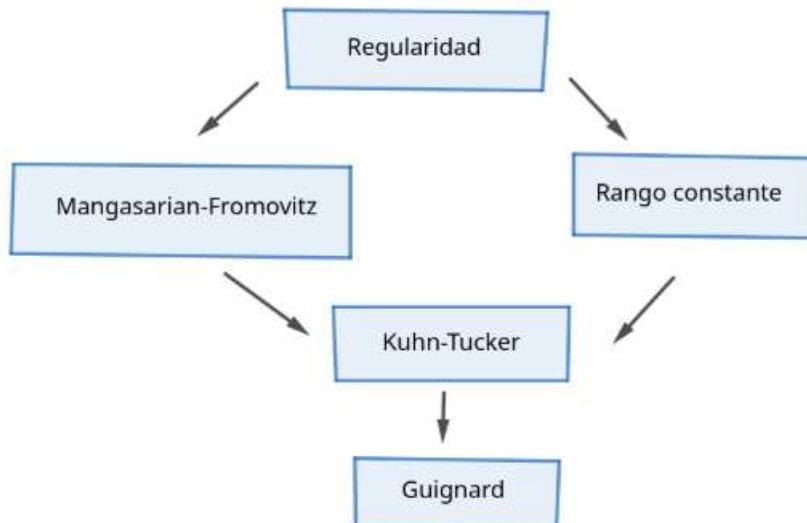


Figure 1: Implicaciones entre las condiciones de calidad

4.5 Condiciones de optimalidad de segundo orden

Como en el caso de minimización irrestricta, existen condiciones necesarias y suficientes de segundo orden. Suponemos que todas las funciones involucradas poseen derivadas hasta orden dos continuas.

Problemas con restricciones de igualdad

Teorema 4.20. Condición necesaria de segundo orden para problemas con restricciones de igualdad. *Sea x^* un minimizador local regular del problema (40) entonces existe $\lambda^* \in \mathbb{R}^m$ tal que*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0 \quad (57)$$

y

$$d^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*) \right) d \geq 0, \quad \forall d \in T^{lin}(x^*). \quad (58)$$

Demostración.

Por hipótesis y por el Teorema 4.5 sabemos que existe $\lambda^* \in \mathbb{R}^m$ verificando (57). Además, como $T^{lin}(x^*) = T_{x^*}$, si $d \in T^{lin}(x^*)$ entonces existe $r(t)$ una curva factible y diferenciable tal que $r(0) = x^*, r'(0) = d$. Luego, veamos cómo es el desarrollo de Taylor para $l(r(t), \lambda^*)$ alrededor de $t = 0$ donde l es la función de Lagrange (46).

Sea $g(t) = l(r(t), \lambda^*)$. Entonces

$$\begin{aligned} g'(t) &= \nabla_x l(r(t), \lambda^*)^T r'(t), \\ g''(t) &= r'(t)^T \nabla_x^2 l(r(t), \lambda^*) r'(t) + \nabla_x l(r(t), \lambda^*)^T r''(t). \end{aligned}$$

Entonces

$$l(r(t), \lambda^*) = l(x^*, \lambda^*) + \nabla_x l(x^*, \lambda^*)^T r'(0)t + \frac{1}{2} t^2 \left(r'(0)^T \nabla_x^2 l(x^*, \lambda^*) r'(0) + \nabla_x l(x^*, \lambda^*)^T r''(0) \right) + o(t^2).$$

Así, como $\nabla_x l(x^*, \lambda^*) = 0$ (por (57)) y por definición de l se tiene

$$f(r(t)) = f(x^*) + \frac{1}{2} t^2 r'(0)^T \nabla_x^2 l(x^*, \lambda^*) r'(0) + o(t^2).$$

Como x^* es minimizador local y $r(t)$ es factible se tiene que

$$f(r(t)) \geq f(x^*) \quad \forall t \in (-\varepsilon, \varepsilon).$$

Entonces

$$\frac{1}{2} t^2 r'(0)^T \nabla_x^2 l(x^*, \lambda^*) r'(0) + o(t^2) \geq 0 \quad \forall t \in (-\varepsilon, \varepsilon).$$

Dividiendo por $t^2 \neq 0$ y tomando límite cuando t tiende a cero obtenemos (58). \square

Nos gustaría establecer ahora condiciones suficientes de segundo orden.

Teorema 4.21. Condición suficiente de segundo orden para problemas con restricciones de igualdad Sea x^* tal que $h(x^*) = 0$. Sea $\lambda^* \in \mathbb{R}^m$ tal que

$$\nabla_x l(x^*, \lambda^*) = 0$$

y suponemos que

$$v^T \nabla_x^2 l(x^*, \lambda^*) v > 0, \quad \forall v \in T^{lin}(x^*) - \{0\}$$

entonces x^* es un minimizador local del problema (40).

Demostración.

Suponemos que x^* no es minimizador local. Entonces, existe una sucesión $\{y_k\}$ tal que $h(y_k) = 0, f(y_k) \leq f(x^*), \|y_k - x^*\| \leq \frac{1}{k}, y_k \neq x^*$.

Sea $\delta_k = \|y_k - x^*\|$. Definimos $s_k = \frac{y_k - x^*}{\delta_k}$ para todo k tal que $\delta_k \neq 0$. Luego, $\{s_k\}$ es una sucesión acotada y tiene una subsucesión convergente. Sean $K_1 \subset \mathbb{N}, s^* \in \mathbb{R}^n$ tales que $\lim_{k \in K_1} s_k = s^*$.

Usando el desarrollo de Taylor tenemos que

$$h_i(y_k) - h_i(x^*) = \delta_k \nabla h_i(x^*)^T s_k + o(\delta_k \|s_k\|).$$

Dividiendo por δ_k para todo k tal que $\delta_k \neq 0$ y tomando límite para $k \in K_1$ tendiendo a infinito obtenemos que $\nabla h_i(x^*)^T s^* = 0$. Luego, $s^* \in T^{lin}(x^*)$. Además, $s^* \neq 0$ ya que $\|s_k\| = 1$.

Usando el desarrollo de Taylor de segundo orden para $l(x, \lambda^*)$ alrededor de x^* y evaluando en y_k obtenemos que

$$l(y_k, \lambda^*) = l(x^*, \lambda^*) + \delta_k \nabla_x l(x^*, \lambda^*)^T s_k + \frac{1}{2} \delta_k^2 s_k^T \nabla_x^2 l(\tilde{x}, \lambda^*) s_k$$

siendo $\tilde{x} = x^* + \alpha(y_k - x^*), 0 < \alpha < 1$ un punto entre y_k y x^* .

Luego, como $h_i(y_k) = 0, h_i(x^*) = 0$ y $\nabla_x l(x^*, \lambda^*) = 0$ obtenemos que

$$l(y_k, \lambda^*) - l(x^*, \lambda^*) = \frac{1}{2} \delta_k^2 s_k^T \nabla_x^2 l(\tilde{x}, \lambda^*) s_k$$

Como $f(y_k) \leq f(x^*)$ tenemos que

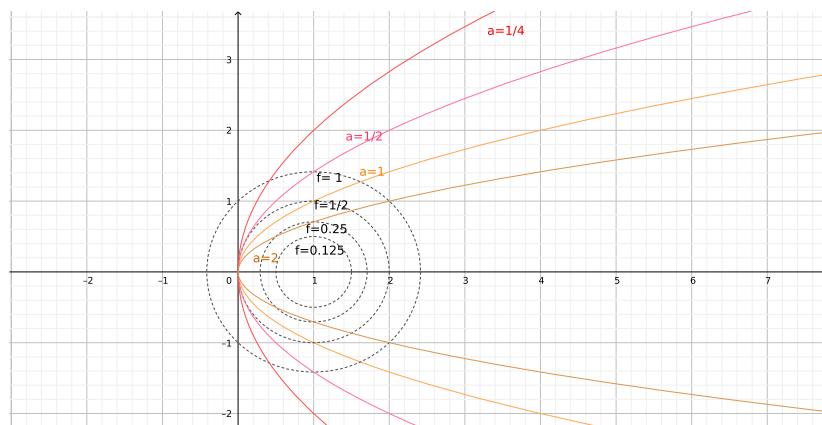
$$\frac{1}{2} \delta_k^2 s_k^T \nabla_x^2 l(\tilde{x}, \lambda^*) s_k \leq 0.$$

Dividiendo por δ_k^2 y haciendo $k \in K_1$ tender a infinito obtenemos que

$$(s^*)^T \nabla^2 l(x^*, \lambda^*) s^* \leq 0$$

que es un absurdo. \square

Ejemplo: Consideremos el problema de minimizar $f(x, y) = 1/2((x - 1)^2 + y^2)$ sujeto a $h(x, y) = x - ay^2$ para $a > 0$. La restricción $h(x, y) = 0$ es una parábola de eje x :



El punto $x^* = (0, 0)$ cumple las condiciones de primer orden con $\lambda = 1$. Observar que el punto es regular:

$$\nabla f(x, y) = \begin{pmatrix} x - 1 \\ y \end{pmatrix}, \nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

$$\nabla h(x, y) = \begin{pmatrix} 1 \\ -2ay \end{pmatrix}, \nabla h(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Veamos las condiciones de segundo orden en el origen:

$$T^{lin}(x^*) = \{d : \nabla h(x^*)^T d = 0\} = \{d = (d_1, d_2) : d_1 = 0\} = \overline{\{(0, 1)\}}$$

$$l((x, y), 1) = f(x, y) + h(x, y) = 1/2((x - 1)^2 + y^2) + x - ay^2$$

$$\nabla_x l((x, y), 1) = \begin{pmatrix} x - 1 \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -2ay \end{pmatrix}, \nabla_x^2 l((x, y), 1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 - 2a \end{pmatrix}$$

Si $d = (0, 1)$

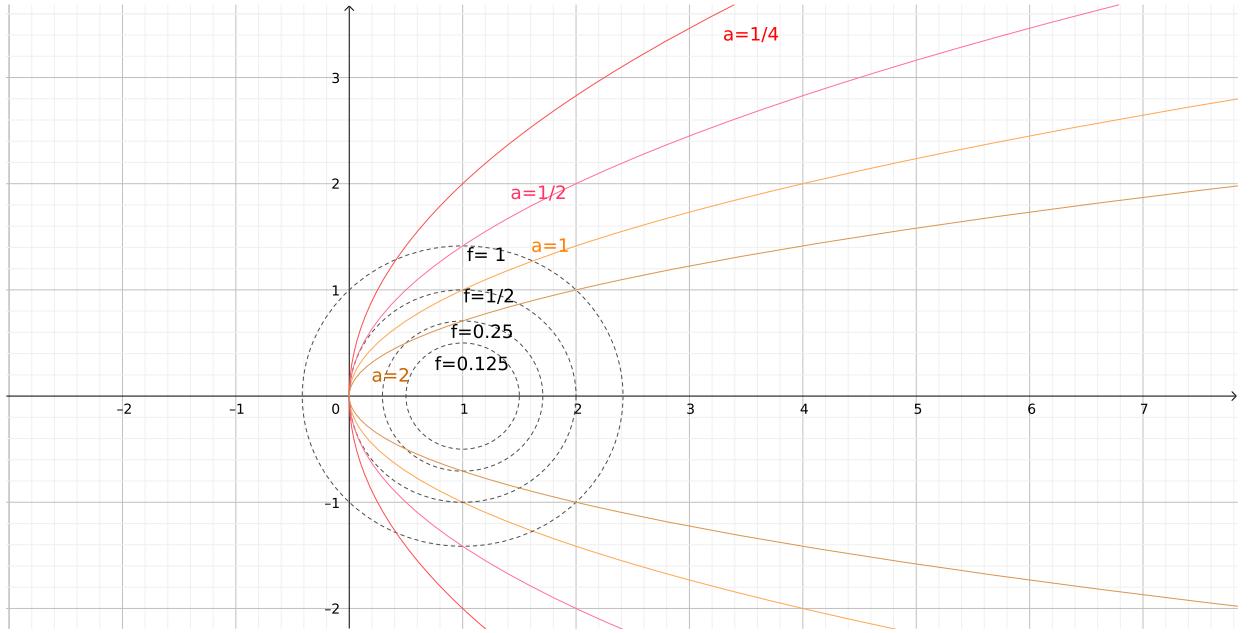
$$d^T \nabla_x^2 l(x^*, 1) d = (0 \ 1) \begin{pmatrix} 1 & 0 \\ 0 & 1 - 2a \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 - 2a.$$

Por lo tanto

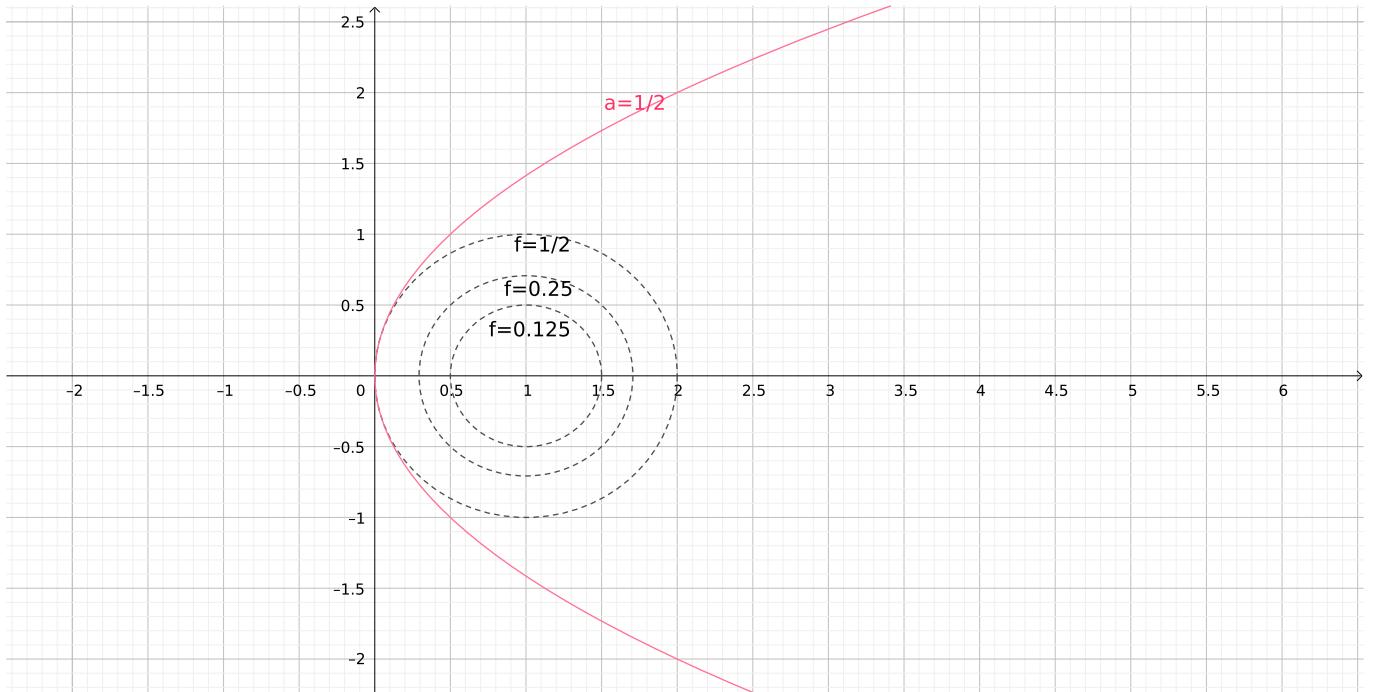
$$d^T \nabla_x^2 l(x^*, 1) d = 1 - 2a.$$

Luego.

- Si $a < 1/2$ valen las condiciones suficientes de segundo orden y x^* es un minimizador local.
- Si $a > 1/2$ el hessiano no es semidefinito positivo y podemos asegurar que x^* NO es un minimizador local. (Verificar que el problema tiene dos minimizadores locales en este caso).



- Si $a = 1/2$ valen las condiciones necesarias de segundo orden pero no las suficientes y no podemos concluir con estas herramientas si x^* es o no minimizador en este caso. Observar en el dibujo que el punto es minimizador local en este caso:



Problemas con restricciones de desigualdad

Analicemos ahora las condiciones necesarias de segundo orden para problemas con restricciones de desigualdad.

Consideremos el problema

$$\text{Minimizar } x - x^2 \text{ sujeto a } x \geq 0.$$

La función objetivo es una parábola con ramas hacia abajo que tiene su máximo en $x = 1/2$. Tenemos que $x = 0$ es un minimizador local del problema. Además, $\mu = 1$ es un multiplicador de Lagrange asociado. Consideremos la función de Lagrange $l(x, \mu) = x - x^2 + \mu(-x)$, luego, para todo $d \in T^{lin}(0) = \{d : -d \leq 0\}$ tenemos que $d^T \nabla^2 l(x, \mu)d < 0$. Luego, $x = 0$ es un minimizador local pero $d^T \nabla^2 l(x, \mu)d \not\geq 0, \forall d \in T^{lin}(0)$.

Vemos entonces que $T^{lin}(x^*)$ no es el espacio adecuado para analizar condiciones de segundo orden para problemas con restricciones de desigualdad.

Consideremos el problema reducido (48):

$$\text{Minimizar } f(x) \text{ sujeto a } g_i(x) = 0, i \in A(x^*).$$

Si x^* es un minimizador local de (47) entonces, usando el Teorema 4.20 y la Proposición 4.1 se puede demostrar el siguiente teorema sobre condiciones necesarias de segundo orden para el problema (47).

Teorema 4.22. Condición necesaria de segundo orden para problemas con restricciones de desigualdad *Sea x^* un minimizador local de (47) tal que el conjunto de gradientes $\{\nabla g_i(x^*)\}_{i \in A(x^*)}$ es linealmente independiente. Entonces existe $\mu \in \mathbb{R}^p$ tal que se verifican las condiciones KKT y*

$$d^T \nabla^2 l(x^*, \mu)d \geq 0 \quad \forall d \in \tilde{T}^{lin}(x^*) = \{d \in \mathbb{R}^n : \nabla g_i(x^*)^T d = 0, \forall i \in A(x^*)\}$$

siendo $l(x, \mu)$ la función de Lagrange asociada: $l(x, \mu) = f(x) + \sum_{i=1}^p \mu_i g_i(x)$.

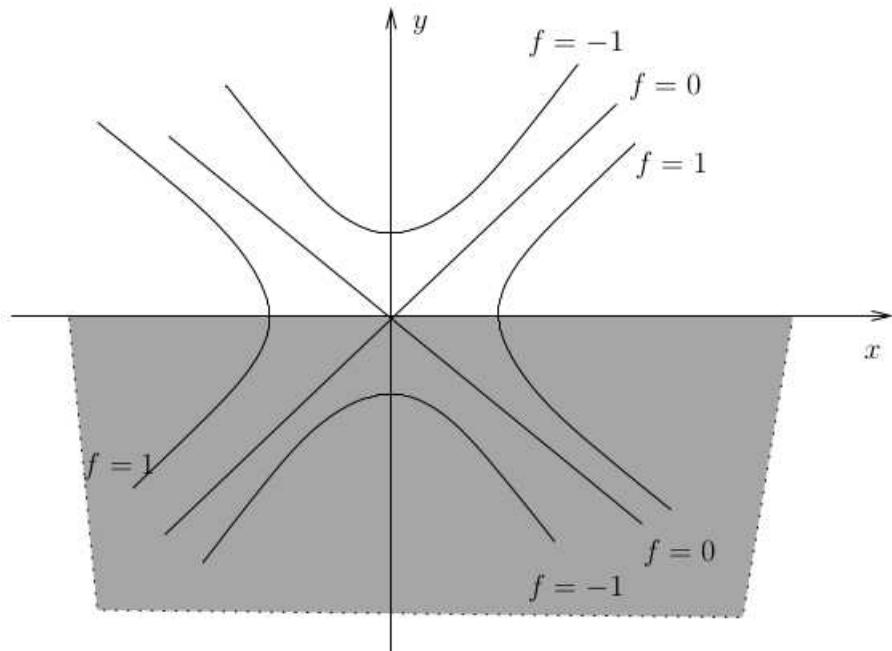
Nos gustaría establecer ahora condiciones suficientes de segundo orden.

Lamentablemente, para problemas con restricciones de desigualdad el espacio $\tilde{T}^{lin}(x^*)$ no sirve para analizar las condiciones suficientes de segundo orden:

Consideremos lo problema

$$\text{Minimizar } 1/2(x^2 - y^2) \text{ sujeto a } y \leq 0.$$

Las curvas de nivel de la función objetivo son hipérbolas como puede verse en la figura.



Luego, gráficamente vemos que $x = (0, 0)$ no es solución, sin embargo, con el multiplicador $\mu = 0$ se tienen las condiciones KKT. Se tiene además que, siendo $l(x, \mu) = 1/2(x^2 - y^2) + \mu y$,

$$\nabla^2 l(x, \mu) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Luego, como $\tilde{T}^{lin}(x) = \{d = (d_1, d_2) : d_2 = 0\}$ se tiene que $d^T \nabla^2 l(x, \mu) d > 0, \forall d \in \tilde{T}^{lin}(x)$ siendo que x no es minimizador local.

Dado x^* un punto factible para el problema (47) y un multiplicador de Lagrange asociado μ , definimos el espacio tangente

$$T_2(x^*) = \left\{ d \in \mathbb{R}^n : \begin{array}{ll} \nabla g_i(x^*)^T d = 0, & i \in A(x^*) : \mu_i > 0 \\ \nabla g_i(x^*)^T d \leq 0, & i \in A(x^*) : \mu_i = 0. \end{array} \right\} \quad (59)$$

Teorema 4.23. Condición suficiente de segundo orden para problemas con restricciones de desigualdad. *Sea x^* un punto factible para (47) tal que existe $\mu \in \mathbb{R}^p$ tal que se verifican las condiciones KKT y*

$$d^T \nabla^2 l(x^*, \mu) d > 0 \quad \forall d \in T_2(x^*) - \{0\}$$

entonces x^ es un minimizador local del problema (47).*

Demostración. Como se hizo para el problema con restricciones de igualdad, suponemos que x^* no es minimizador local. Entonces, existe una sucesión $\{y_k\}$ tal que $g(y_k) \leq 0, f(y_k) \leq f(x^*)$, $\|y_k - x^*\| \leq \frac{1}{k}, y_k \neq x^*$.

Sea $\delta_k = \|y_k - x^*\|$. Definimos $s_k = \frac{y_k - x^*}{\delta_k}$. Luego, $\{s_k\}$ es una sucesión acotada y tiene una subsucesión convergente. Sean $K_1 \subset \mathbb{N}$, $s^* \in \mathbb{R}^n$ tales que $\lim_{k \in K_1} s_k = s^*$.

Si $i \in A(x^*)$ entonces $g_i(x^*) = 0$ y tenemos que

$$g_i(y_k) - g_i(x^*) = \delta_k \nabla g_i(x^*)^T s_k + o(\delta_k \|s_k\|) \leq 0.$$

Entonces, dividiendo por δ_k y tomando límite para $k \in K_1$ tendiendo a infinito obtenemos que $\nabla g_i(x^*)^T s^* \leq 0$. Además, $s^* \neq 0$ ya que $\|s_k\| = 1$.

Usando el desarrollo de Taylor de segundo orden para $f(x)$ y para cada $g_i(x), i \in A(x^*)$ alrededor de x^* y evaluando en y_k obtenemos que

$$\begin{aligned} 0 &\geq f(y_k) - f(x^*) = \delta_k \nabla f(x^*)^T s_k + \frac{1}{2} \delta_k^2 s_k^T \nabla^2 f(x^*) s_k + o(\delta_k^2 \|s_k\|^2) \\ 0 &\geq g_i(y_k) - g_i(x^*) = \delta_k \nabla g_i(x^*)^T s_k + \frac{1}{2} \delta_k^2 s_k^T \nabla^2 g_i(x^*) s_k + o(\delta_k^2 \|s_k\|^2). \end{aligned} \quad (60)$$

Luego,

$$\begin{aligned} 0 &\geq l(y_k, \mu) - l(x^*, \mu) = (f(y_k) + \sum_{i \in A(x^*)} \mu_i g_i(y_k)) - (f(x^*) + \sum_{i \in A(x^*)} \mu_i g_i(x^*)) = \\ &\quad \delta_k \nabla_x l(x^*, \mu)^T s_k + \frac{1}{2} \delta_k^2 s_k^T \nabla_x^2 l(x^*, \mu) s_k + o(\delta_k^2 \|s_k\|^2). \end{aligned}$$

Como $\nabla_x l(x^*, \mu) = 0$ obtenemos que

$$\frac{1}{2} \delta_k^2 s_k^T \nabla_x^2 l(x^*, \mu) s_k \leq 0.$$

Dividiendo por δ_k^2 y haciendo $k \in K_1$ tender a infinito obtenemos que

$$(s^*)^T \nabla_x^2 l(x^*, \mu) s^* \leq 0.$$

Para contradecir la hipótesis tenemos que demostrar que $s^* \in T_2(x^*)$. Tenemos que $\nabla g_i(x^*)^T s^* \leq 0, \forall i \in A(x^*)$.

Suponemos que existe $\mu_j > 0$ tal que $\nabla g_j(x^*)^T s^* < 0$. Entonces,

$$\nabla f(x^*)^T s^* + \sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*)^T s^* = 0.$$

Luego

$$\nabla f(x^*)^T s^* = - \sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*)^T s^* = -\mu_j \nabla g_j(x^*)^T s^* - \sum_{i \in A(x^*) - \{j\}} \mu_i \nabla g_i(x^*)^T s^* > 0.$$

Pero, usando el desarrollo de Taylor de orden 1 para f alrededor de x^* , obtenemos que

$$\nabla f(x^*)^T s^* \leq 0$$

que genera una contradicción. Por lo tanto, si $\mu_j > 0$ debe ser $\nabla g_j(x^*)^T s^* = 0$ y esto demuestra que $s^* \in T_2(x^*)$. \square

Utilizando este tangente se puede obtener un nuevo teorema sobre condición necesaria de segundo orden para el problema (47).

Teorema 4.24. Condición necesaria de segundo orden. Sea x^* un minimizador local de (47) regular. Sea $\mu \in \mathbb{R}^p$ tal que se verifican las condiciones KKT, entonces

$$d^T \nabla^2 l(x^*, \mu) d \geq 0 \quad \forall d \in T_2(x^*).$$

Para una demostración ver [8].

Problemas de optimización con restricciones mixtas

Teorema 4.25. Condiciones necesarias de segundo orden Sea x^* un minimizador local regular de (51). Entonces existen escalares $\lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_p$ tal que se verifican las condiciones KKT y

$$d^T \nabla^2 l(x^*, \lambda, \mu) d \geq 0 \quad \forall d : \begin{cases} \nabla h_i(x^*)^T d = 0, & i = 1, \dots, m; \\ \nabla g_i(x^*)^T d = 0, & \forall i \in A(x^*). \end{cases} \quad (61)$$

Teorema 4.26. Condición suficiente de segundo orden. Sea x^* un punto factible tal que existen $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ tal que se verifican las condiciones KKT y, $\forall d, d \neq 0$ tal que

$$\begin{aligned} \nabla h_i(x^*)^T d &= 0, & i &= 1, \dots, m \\ \nabla g_i(x^*)^T d &= 0, & i \in A(x^*) : \mu_i &> 0 \\ \nabla g_i(x^*)^T d &\leq 0, & i \in A(x^*) : \mu_i &= 0. \end{aligned}$$

se cumple que

$$d^T \nabla^2 l(x^*, \lambda, \mu) d > 0$$

entonces x^* es un minimizador local del problema (51).

En el teorema 4.25 estamos estableciendo que la condición de regularidad es una condición de calidad de segundo orden, en el sentido de que garantiza no solo la existencia de los multiplicadores de Lagrange sino que además vale la condición (61) que establece que el hessiano de la función de Lagrange es semidefinito positivo en el espacio $\nabla h_i(x^*)^T d = 0, i = 1, \dots, m; \nabla g_i(x^*)^T d = 0, \forall i \in A(x^*)$.

De las condiciones de calidad de primer orden dadas en la Figura 4.4 podemos decir lo siguiente:

- Mangasarian Fromovitz NO es una condición de calidad de segundo orden: existe un ejemplo de un problema de programación no lineal en el cual no vale Mangasarian-Fromovitz pero vale la condición necesaria de segundo orden en el minimizador local, [13].
- La condición de rango constante es una condición de calidad de segundo orden, [14].

4.6 Ejercicios.

1. Demostrar que los escalares $\lambda_1, \dots, \lambda_m$ del Teorema 4.4 son únicos.
2. Demostrar que si x^* es un minimizador local de f sujeto a $g_i(x) \leq 0, i = 1, \dots, p$ y $A(x^*)$ es el conjunto de restricciones activas entonces x^* es un minimizador local del siguiente problema con restricciones de igualdad:

$$\begin{array}{ll} \text{Min} & f(x) \\ \text{s.a} & g_i(x) = 0, i \in A(x^*). \end{array}$$

3. Como en el caso de restricciones de igualdad, demostrar que si los gradientes $\{\nabla g_i(x^*), i \in A(x^*)\}$ son linealmente independientes entonces $T_{x^*} = T^{lin}(x^*)$, siendo $T^{lin}(x^*) = \{d \in \mathbb{R}^n : \nabla g_i(x^*)^T d \leq 0, i \in A(x^*)\}$.
4. Dados los vectores $v_1, v_2, \dots, v_m \in \mathbb{R}^n \setminus \{0\}$, demostrar que el conjunto

$$C = \left\{ \sum_{i=1}^m y_i v_i \mid y_i \geq 0, i = 1, \dots, m \right\}$$

es un cono convexo y cerrado.

5. Demostrar la Proposición 4.3.
6. Mostrar con un ejemplo numérico que la condición de calidad de Mangasarian-Fromovitz es estrictamente más débil que regularidad. Esto quiere decir, encontrar un ejemplo en el cual se verifique Mangasarian-Fromovitz pero no regularidad.
7. Mostrar con un ejemplo numérico que la condición de calidad de rango constante es estrictamente más débil que regularidad. Esto quiere decir, encontrar un ejemplo en el cual se verifique rango constante pero no regularidad.
8. Mostrar con un ejemplo numérico que las condiciones de calidad de rango constante y Mangasarian-Fromovitz no se implican mutuamente.
9. Considere el problema

$$\begin{aligned} \text{Min } & x_1^2 + 4x_1x_2 + x_2^2 \\ \text{s.a } & x_1^2 + x_2^2 = 1 \end{aligned}$$

- (a) Usando las condiciones KKT, encontrar una solución x^* del problema. Interpretar las condiciones KKT geométricamente en x^* . Como son las curvas de nivel de f ?
- (b) Calcular $T^{lin}(x^*)$ y analizar si se verifican las condiciones necesarias y suficientes de segundo orden.
- (c) Tiene el problema una única solución?
10. Considere el problema

$$\begin{aligned} \text{Min } & \left(x_1 - \frac{9}{4}\right)^2 + (x_2 - 2)^2 \\ \text{s.a } & x_1^2 - x_2 \leq 0 \\ & x_1 + x_2 \leq 6 \end{aligned}$$

- (a) Escribir las condiciones KKT y verificar que estas condiciones se cumplen en el punto $x^* = (\frac{3}{2}, \frac{9}{4})$.
- (b) Analizar las condiciones suficientes de segundo orden (Teorema 4.23).
- (c) Mostrar geométricamente que x^* es el único minimizador global.
11. Considere el problema

$$\begin{aligned} \text{Min } & x_1 x_2 \\ \text{s.a } & x_1 \leq x_2 \end{aligned}$$

Mostrar que se cumplen las condiciones KKT en el punto $x^* = (0, 0)$.

- (a) Analizar si se verifican las condiciones necesarias de segundo orden en el tangente $T_{x^*} = \{d : \nabla g_i(x^*)^T d = 0, \forall i \in A(x^*)\} - \{0\}$.
- (b) Analizar si se verifican las condiciones suficientes de segundo orden.
- (c) Es x^* es un minimizador local? Puede justificarlo gráficamente.
12. Mostrar que cualquier punto factible del conjunto definido por $\{x \in \mathbb{R}^n : x \geq 0\}$ es regular.
13. Demostrar el Teorema 4.22.
14. (Optativo) En el siguiente ejemplo se muestra que Mangasarian-Fromovitz NO es una condición de calidad que pueda reemplazar a la condición de regularidad en los teoremas de condiciones necesarias de segundo orden.

$$\begin{array}{ll}\text{Min} & x_3 \\ \text{s.a} & 2\sqrt{3}x_1x_2 - 2x_2^2 - x_3 \leq 0 \\ & x_2^2 - 3x_1^2 - x_3 \leq 0 \\ & -2\sqrt{3}x_1x_2 - 2x_2^2 - x_3 \leq 0\end{array}$$

- (a) Mostrar que $(0, 0, 0)$ es minimizador local del problema.
- (b) Mostrar que en $(0, 0, 0)$ vale la condición de Mangasarian-Fromovitz.
- (c) Mostrar que el conjunto de escalares que cumplen las condiciones KKT es el siguiente:

$$\Delta(x^*) = \{\mu = (\mu_1, \mu_2, \mu_3) : \mu_i \geq 0, \mu_1 + \mu_2 + \mu_3 = 1\}.$$

- (d) Mostrar que para ningún multiplicador de Lagrange se verifica el Teorema 4.22.

5 Problemas con restricciones especiales

En esta sección nos interesa analizar problemas de optimización con restricciones de la forma

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } x \in \Omega, \end{aligned} \tag{62}$$

para $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^1 y Ω un subconjunto de \mathbb{R}^n con restricciones fáciles. Por restricciones fáciles nos referimos a restricciones que poseen una estructura específica que debe ser aprovechada para el desarrollo de métodos para resolver esos problemas.

Algunos ejemplos de este tipo de problemas son:

- Problemas con restricciones de cotas en las variables:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } l \leq x \leq u, \end{aligned} \tag{63}$$

para $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^1 y $l, u \in \mathbb{R}^n, l \leq u$ pueden tener alguna coordenada ∞ .

- Problemas con restricciones lineales de igualdad:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } Ax = b, \end{aligned}$$

donde asumimos que $A \in \mathbb{R}^{m \times n}, m < n$.

- Problemas con restricciones lineales de desigualdad:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } Ax \leq b. \end{aligned} \tag{64}$$

Estos ejemplos tienen en común que el conjunto factible es un conjunto convexo y cerrado y por ende está bien definida la proyección ortogonal de cualquier punto sobre el conjunto.

Para el conjunto $\Omega = \{x \in \mathbb{R}^n : x \geq 0\}$, la proyección de x sobre Ω es el vector $y = P_\Omega(x)$ cuyas componentes son

$$y_i = \begin{cases} 0 & \text{si } x_i < 0 \\ x_i & \text{si } x_i \geq 0, \end{cases}$$

es decir $y_i = \max\{x_i, 0\}$.

Para el conjunto $\Omega = \{x \in \mathbb{R}^n : l \leq x \leq u\}$, la proyección sobre Ω es el vector $y = P_\Omega(x)$ cuyas componentes son $y_i = \min\{\max\{x_i, l_i\}, u_i\}$.

Para el problema

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } x \geq 0 \end{aligned} \tag{65}$$

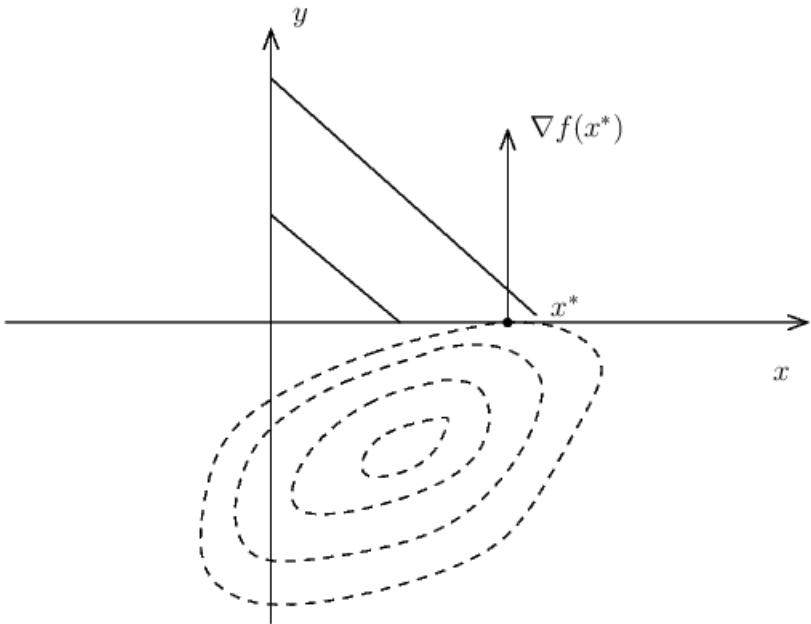
la restricción $x \geq 0$ se puede escribir como $g(x) = -x \leq 0$, es decir, hay n restricciones de desigualdad dadas por $g_i(x) = -x_i$. Luego, cualquier punto factible es regular (los gradientes de las restricciones activas son linealmente independientes) y tenemos que, si x^* es solución del problema entonces existe $\mu \in \mathbb{R}^n, \mu \geq 0$ tal que

$$\begin{aligned} -\nabla f(x^*) &= \sum_{i=1}^n \mu_i (-e_i) \\ \mu_i x_i^* &= 0, \forall i = 1, \dots, n. \end{aligned}$$

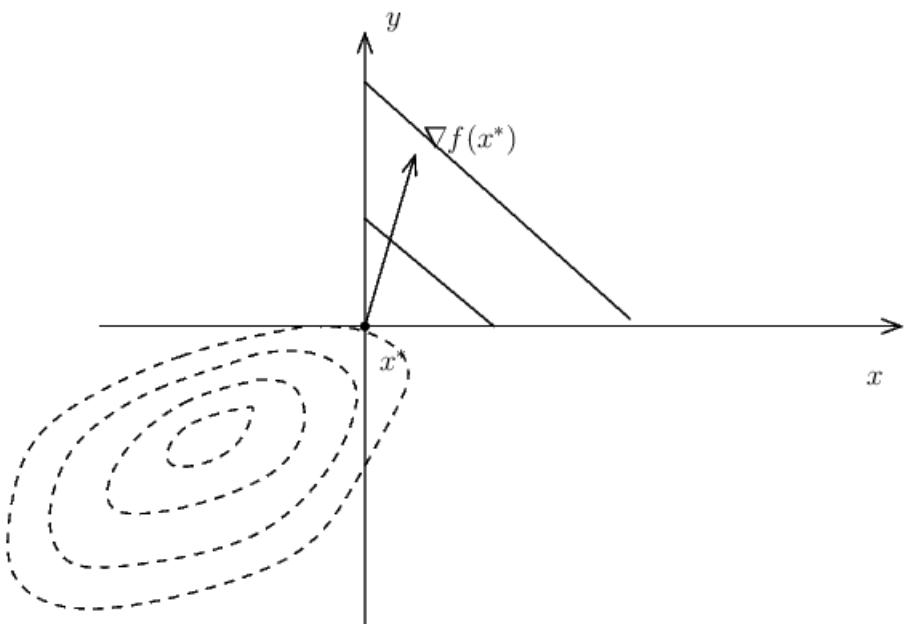
Observar que, si la i -ésima restricción es activa ($x_i^* = 0$) entonces, la i -ésima ecuación del sistema KKT es la igualdad $-\frac{\partial f}{\partial x_i}(x^*) = -\mu_i$. Si la i -ésima restricción es inactiva ($x_i^* > 0$) entonces $\mu_i = 0$ y se tiene que $\frac{\partial f}{\partial x_i}(x^*) = 0$. Luego, si x^* es un minimizador local del problema (65) entonces las condiciones KKT se escriben como

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x^*) &= 0, && \text{si } x_i^* > 0 \\ \frac{\partial f}{\partial x_i}(x^*) &\geq 0, && \text{si } x_i^* = 0. \end{aligned} \quad (66)$$

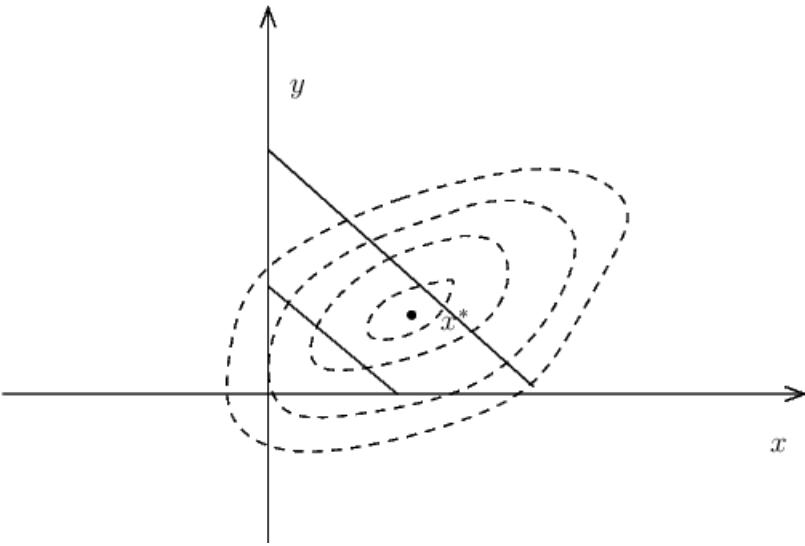
Un punto x^* que cumple (66) es un punto estacionario del problema (65). Observar los siguientes ejemplos.



En x^* se tiene:
 $x_1^* > 0$ y $\frac{\partial f}{\partial x_1}(x^*) = 0$,
 $x_2^* = 0$ y $\frac{\partial f}{\partial x_2}(x^*) > 0$



En x^* se tiene:
 $x_1^* = 0$ y $\frac{\partial f}{\partial x_1}(x^*) > 0$,
 $x_2^* = 0$ y $\frac{\partial f}{\partial x_2}(x^*) > 0$



En x^* se tiene:
 $x_1^* > 0$ y $\frac{\partial f}{\partial x_1}(x^*) = 0$,
 $x_2^* > 0$ y $\frac{\partial f}{\partial x_2}(x^*) = 0$

Para el conjunto $\Omega = \{x \in \mathbb{R}^n : x \geq 0\}$ se tiene el siguiente resultado:

Proposición 5.1. x^* es un punto KKT para el problema (65) si y solo si $P_\Omega(x^* - \nabla f(x^*)) - x^* = 0$.

Demostración. Supongamos que x^* es un punto KKT. Luego, por (66) se tiene que:

- Si $x_i^* = 0$ entonces $\frac{\partial f}{\partial x_i}(x^*) = \mu_i \geq 0$.

Entonces $x_i^* - \frac{\partial f}{\partial x_i}(x^*) \leq 0 \Rightarrow [P_\Omega(x^* - \nabla f(x^*)) - x^*]_i = 0$

- Si $x_i^* > 0$ entonces $\frac{\partial f}{\partial x_i}(x^*) = 0$.

Luego $x_i^* - \frac{\partial f}{\partial x_i}(x^*) = x_i^* > 0 \Rightarrow [P_\Omega(x^* - \nabla f(x^*)) - x^*]_i = 0$

Por lo tanto $P_\Omega(x^* - \nabla f(x^*)) - x^* = 0$ como queríamos demostrar.

Supongamos ahora que se tiene $P_\Omega(x^* - \nabla f(x^*)) - x^* = 0$. Veamos que x^* es un punto KKT.

- Si $x_i^* = 0$, como $[P_\Omega(x^* - \nabla f(x^*)) - x^*]_i = x_i^* = 0$ se tiene que

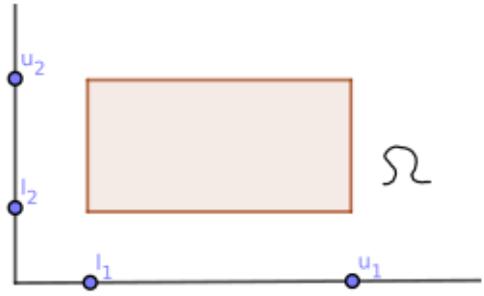
$$x_i^* - \frac{\partial f}{\partial x_i}(x^*) \leq 0 \text{ entonces } \frac{\partial f}{\partial x_i}(x^*) \geq 0.$$

- Si $x_i^* > 0$, como $[P_\Omega(x^* - \nabla f(x^*)) - x^*]_i = x_i^* > 0$ se tiene que

$$x_i^* - \frac{\partial f}{\partial x_i}(x^*) = x_i^* \text{ entonces } \frac{\partial f}{\partial x_i}(x^*) = 0.$$

Si tenemos en cuenta el caso más general (63), el conjunto factible $\Omega = \{x \in \mathbb{R}^n : l \leq x \leq u\}$ está definido por las siguientes restricciones de desigualdad:

$$\begin{aligned}
g_1(x) &= x_1 - u_1 \leq 0 & g_{n+1}(x) &= l_1 - x_1 \leq 0 \\
g_2(x) &= x_2 - u_2 \leq 0 & g_{n+2}(x) &= l_2 - x_2 \leq 0 \\
&\vdots &&\vdots \\
g_n(x) &= x_n - u_n \leq 0 & g_{2n}(x) &= l_n - x_n \leq 0.
\end{aligned}$$

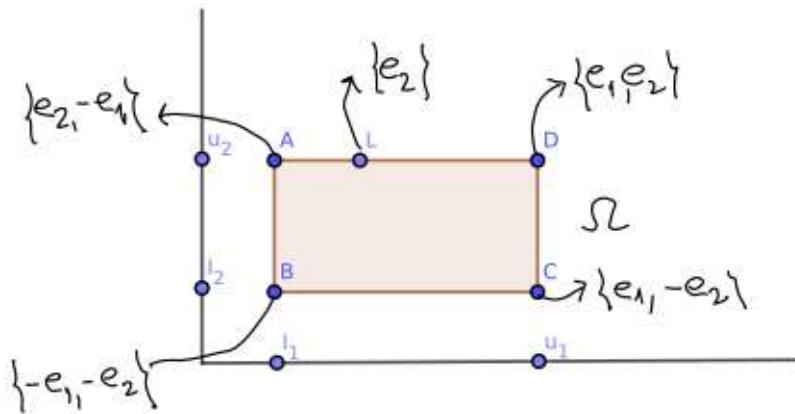


Todo punto $x \in \Omega$ es regular:

$$\nabla g_i(x) = e_i, \forall i = 1, \dots, n; \nabla g_{n+i}(x) = -e_i, \forall i = 1, \dots, n.$$

- Si $g_i(x)$ es activa en x^* entonces $g_{n+i}(x)$ es inactiva en x^* .
- Si $g_{n+i}(x)$ es activa en x^* entonces $g_i(x)$ es inactiva en x^* .

Por lo tanto, $\forall x^* \in \Omega, \{\nabla g_i(x^*)\}_{i \in A(x^*)} \subset \{n \text{ vectores LI entre } \{e_i\}_{i=1}^n \cup \{-e_i\}_{i=1}^n\}$ lo que implica que x^* es regular.



Luego, si x^* es un minimizador local de (63) existen $\mu, \nu \geq 0$ tales que

$$\begin{aligned}
\nabla f(x^*) + \sum_{i=1}^n \mu_i e_i + \sum_{i=1}^n \nu_i (-e_i) &= 0 \\
\mu_i (x_i^* - u_i) &= 0, i = 1, \dots, n \\
\nu_i (l_i - x_i^*) &= 0, i = 1, \dots, n.
\end{aligned}$$

Observación:

- Si g_i es activa en x^* entonces $x_i^* = u_i \Rightarrow \nu_i = 0$ y $\frac{\partial f}{\partial x_i}(x^*) = -\mu_i \leq 0$.
- Si g_{n+i} es activa en x^* entonces $x_i^* = l_i \Rightarrow \mu_i = 0$ y $\frac{\partial f}{\partial x_i}(x^*) = \nu_i \geq 0$.

Por lo tanto, si x^* es un minimizador local entonces

$$\begin{aligned}\frac{\partial f}{\partial x_i}(x^*) &\geq 0 && \text{si } x_i^* = l_i \\ \frac{\partial f}{\partial x_i}(x^*) &= 0 && \text{si } l_i < x_i^* < u_i \\ \frac{\partial f}{\partial x_i}(x^*) &\leq 0 && \text{si } x_i^* = u_i\end{aligned}\tag{67}$$

Un punto x^* que cumple (67) es un punto estacionario del problema (63).

En el caso de una caja general, como sucede con el problema (65) se puede demostrar que vale la Proposición 5.1. El siguiente teorema es el resultado más general que vale para un conjunto factible convexo y cerrado cualquiera y es una generalización de la Proposición 5.1:

Teorema 5.1. *Si x^* es un minimizador local para el problema (62) con $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función C^1 y Ω un conjunto convexo cerrado entonces $P_\Omega(x^* - \nabla f(x^*)) - x^* = 0$.*

Demostración. Como x^* es un minimizador local se tiene que

$$f(x^*) \leq f((1-t)x^* + ty), \quad \forall y \in \Omega, t \geq 0 \text{ suficientemente pequeño.}$$

Entonces

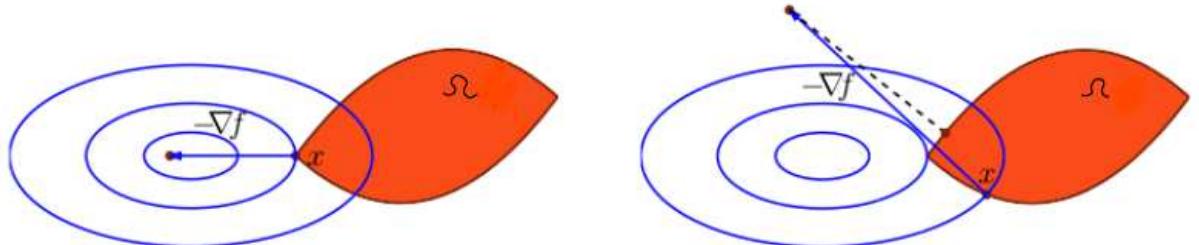
$$0 \leq f(x^* + t(y - x^*)) - f(x^*) = t\nabla f(x^*)^T(y - x^*) + o(t)$$

donde $\lim_{t \rightarrow 0^+} \frac{o(t)}{t} = 0$.

Dividiendo por $t \geq 0$ y tomando límite cuando $t \rightarrow 0^+$ se obtiene que $\nabla f(x^*)^T(y - x^*) \geq 0, \forall y \in \Omega$. Entonces, $\forall y \in \Omega$ se tiene que

$$([x^* - \nabla f(x^*)] - x^*)^T(y - x^*) \leq 0.$$

Luego, usando la Proposición 2.1 $x^* = P_\Omega(x^* - \nabla f(x^*))$. □



Un punto x^* que cumple $x^* = P_D(x^* - \nabla f(x^*))$ es un *punto estacionario* del problema (62).

5.1 Método del gradiente proyectado

La condición de optimalidad dada en el Teorema 5.1 sugiere definir el método del gradiente proyectado para resolver el problema (62).

Método del gradiente proyectado. Dado x_0 factible inicial. En un $x_k \in \Omega$ hacer los siguientes pasos:

1. Calcular

$$\begin{aligned}y_k &= x_k - \nabla f(x_k) \\z_k &= P_\Omega(y_k) \\d_k &= z_k - x_k\end{aligned}$$

Si $d_k = 0$ parar. (x_k es un punto estacionario)

Si no, hacer búsqueda de Armijo en la dirección d_k :

2. Definir $t = 1$. Si

$$f(x_k + td_k) \leq f(x_k) + \sigma_1 t \nabla f(x_k)^T d_k \Rightarrow \quad (68)$$

(a) definir $t_k = t$ y $x_k + t_k d_k$, e ir a 1.

(b) si no, hacer $t = t/2$ e ir a (68).

Observaciones.

1. Si $\Omega = \mathbb{R}^n$ es el método del gradiente con búsqueda de Armijo.

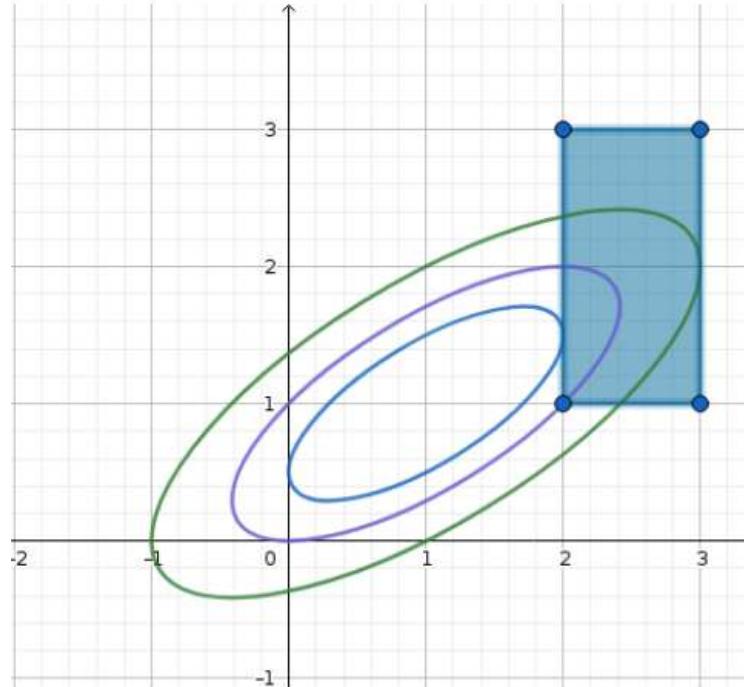
2. El algoritmo es aplicable a cualquier conjunto Ω siempre que sea convexo y cerrado.

Ejemplo. Hacer algunas iteraciones del método del gradiente proyectado para resolver el problema

Minimizar $-2xy - 2y + x^2 + 2y^2$ sujeto a $2 \leq x \leq 3, 1 \leq y \leq 3$

comenzando desde $x_0 = (3, 1)$ y utilizando $\sigma_1 = 1/2$.

Gráficamente se observa que la solución del problema es el punto $x^* = (2, 3/2)$.



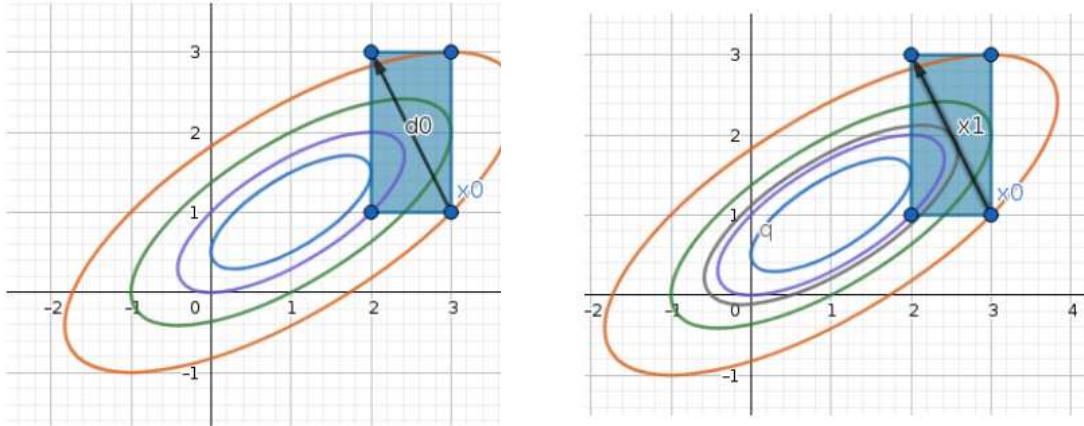
- $k = 0, y_0 = x_0 - \nabla f(x_0) = \begin{pmatrix} 3 \\ 1 \end{pmatrix} - \begin{pmatrix} 4 \\ -4 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$

$$z_0 = P_\Omega(y_0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$d_0 = z_0 - x_0 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Vale Armijo con $t = 1$? NO

Vale Armijo con $t = 1/2$? Si $\Rightarrow x_1 = x_0 + 1/2d_0 = (5/2, 2)$.



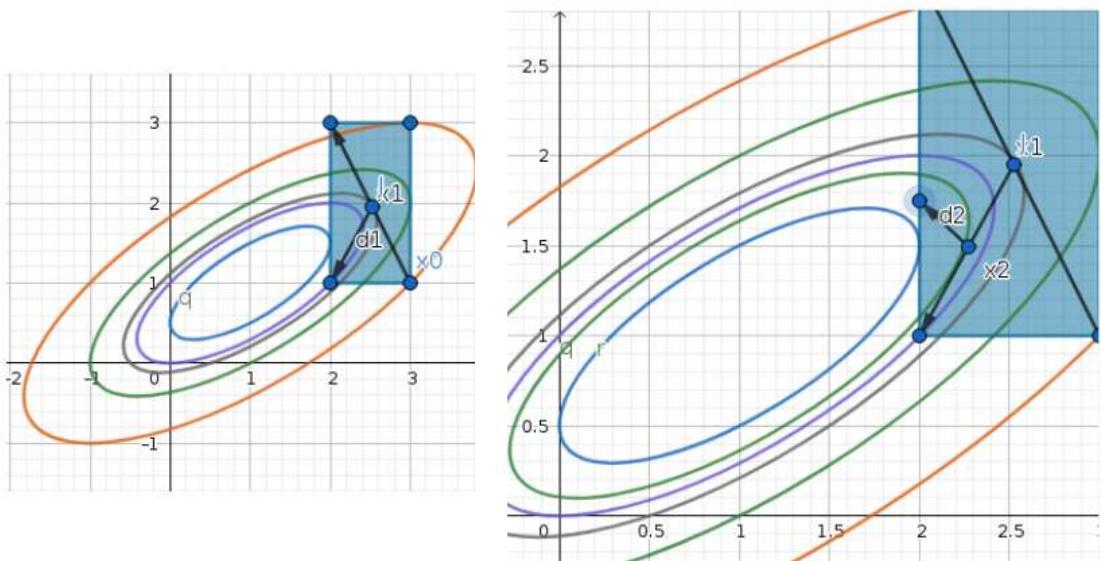
- $y_1 = x_1 - \nabla f(x_1) = \begin{pmatrix} 5/2 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}$

$$z_1 = P_\Omega(y_1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$d_1 = z_1 - x_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 5/2 \\ 2 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1 \end{pmatrix}$$

Vale Armijo con $t = 1$? NO

Vale Armijo con $t = 1/2$? Si $\Rightarrow x_2 = x_1 + 1/2d_1 = (9/4, 3/2)$.



En el siguiente teorema se demuestra que si el algoritmo no se detiene entonces genera una sucesión infinita para la cual cualquier punto de acumulación es estacionario.

Teorema 5.2. Sea $f \in C^1$

1. Para todo k , si $z_k \neq x_k$, d_k es una dirección de descenso para f en x_k .
2. Si $z_k = x_k$ para algún k entonces x_k es un punto estacionario.
3. Si se genera una sucesión infinita $\{x_k\}$ entonces todo punto de acumulación es estacionario.

Demostración. 1. Por definición $z_k = P_\Omega(y_k)$. Entonces, por la Proposición 2.1 tenemos que

$$(y_k - P_\Omega(y_k))^T(x_k - P_\Omega(y_k)) \leq 0.$$

Usando que $y_k = x_k - \nabla f(x_k)$ tenemos que

$$(x_k - z_k)^T(x_k - z_k) - \nabla f(x_k)^T(x_k - z_k) \leq 0.$$

Luego, como $d_k = z_k - x_k$ y $z_k \neq x_k$ se obtiene que

$$\nabla f(x_k)^T d_k \leq -\|d_k\|^2 < 0 \quad (69)$$

y d_k es una dirección de descenso.

2. Se obtiene directamente de la definición de punto estacionario.
3. Sean $K \subset \mathbb{N}$, $x^* \in \mathbb{R}^n$ tales que $\lim_{k \in K} x_k = x^*$.

Por la continuidad de f y de la proyección tenemos que

$$\lim_{k \in K} d_k = \lim_{k \in K} P_\Omega(x_k - \nabla f(x_k)) - x_k = P_\Omega(x^* - \nabla f(x^*)) - x^*.$$

Definimos $d^* = P_\Omega(x^* - \nabla f(x^*)) - x^*$.

Por un lado, como d_k es una dirección de descenso tenemos que

$$\nabla f(x_k)^T d_k < 0, \forall k \in \mathbb{N}$$

luego, tomando límite para $k \in K$ tenemos que

$$\nabla f(x^*)^T d^* \leq 0. \quad (70)$$

Por la búsqueda de Armijo tenemos que

$$f(x_{k+1}) - f(x_k) \leq \sigma_1 t_k \nabla f(x_k)^T d_k < 0.$$

Luego, la sucesión $\{f(x_k)\}$ es monótona decreciente y sabemos que

$$\lim_{k \rightarrow \infty} f(x_{k+1}) - f(x_k) = 0.$$

Entonces

$$\lim_{k \in K} t_k \nabla f(x_k)^T d_k = 0.$$

Consideramos dos posibilidades para la sucesión $\{t_k\}$:

- (a) $\lim_{k \in K} t_k = 0$.
- (b) $\exists K_1 \subset K : t_k \geq c > 0$.

En el caso (a) existe $t'_k > 0 : \lim_{k \in K_1} t'_k = 0$ para el cual Armijo no vale:

$$f(x_k + t'_k d_k) > f(x_k) + \sigma_1 t'_k \nabla f(x_k)^T d_k \Rightarrow$$

$$\frac{f(x_k + t'_k d_k) - f(x_k)}{t'_k} > \sigma_1 \nabla f(x_k)^T d_k.$$

Luego, por el TVM tenemos que

$$\nabla f(x_k + \xi_k t'_k d_k)^T d_k > \sigma_1 \nabla f(x_k)^T d_k, \xi_k \in (0, 1).$$

Tomando límite para $k \in K_1$ obtenemos

$$\nabla f(x^*)^T d^* \geq \sigma_1 \nabla f(x^*)^T d^*.$$

Luego, como $\sigma_1 \in (0, 1)$ debe ser $\nabla f(x^*)^T d^* \geq 0$. Entonces, por (70) obtenemos que $\nabla f(x^*)^T d^* = 0$.

Si se cumple (b) entonces

$$f(x_{k+1}) \leq f(x_k) + \sigma_1 c \nabla f(x_k)^T d_k < 0$$

y tomando límite para $k \in K_1$ obtenemos que $\nabla f(x^*)^T d^* \geq 0$ y nuevamente por (70) obtenemos que $\nabla f(x^*)^T d^* = 0$.

Finalmente, usando (69) tenemos que debe ser $d^* = 0$ y x^* es un punto estacionario. \square

5.2 Problemas con restricciones lineales

El método de gradiente proyectado de la sección anterior se aplica para resolver el problema con restricciones lineales de igualdad:

$$\begin{aligned} &\text{Minimizar } f(x) \\ &\text{sujeto a } Ax = b, \end{aligned} \tag{71}$$

donde asumimos que $A \in \mathbb{R}^{m \times n}$, $m < n$, $\text{rango}(A) = m$.

La restricción $Ax = b$ equivale a m restricciones lineales de igualdad:

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m. \end{aligned}$$

La hipótesis $\text{rango}(A) = m$ no es necesaria para demostrar que se cumplen las condiciones KKT en una solución pero si para definir algoritmos basados en proyecciones. Cuando el problema tiene restricciones lineales se puede demostrar que $T_{x^*} = T^{lin}(x^*)$ y está garantizada la existencia de multiplicadores de Lagrange en una solución.

Cuando $\text{rango}(A) \neq m$ entonces, el sistema o no tiene solución o puede tener ecuaciones redundantes que pueden ser eliminadas sin afectar la solución del problema.

Elegir m columnas LI se puede hacer mediante eliminación gaussiana.

Dado un punto $x_k \in \Omega = \{x : Ax = b\}$, nos preguntamos cómo se calcula la proyección del vector $x_k - \nabla f(x_k)$ sobre Ω . Sabemos que $P_\Omega(x_k - \nabla f(x_k))$ es la solución del problema cuadrático de

$$\begin{aligned} \min & \quad 1/2 \|y - (x_k - \nabla f(x_k))\|^2 \\ \text{s. a} & \quad y \in \Omega \end{aligned}$$

y existe $\lambda \in \mathbb{R}^m$ tal que

$$P_\Omega(x_k - \nabla f(x_k)) - (x_k - \nabla f(x_k)) = \sum_{i=1}^m \lambda_i a_i = A^T \lambda \quad (72)$$

donde $a_i = (a_{i1} \dots a_{in})^T$. Entonces

$$A(P_\Omega(x_k - \nabla f(x_k)) - (x_k - \nabla f(x_k))) = AA^T \lambda.$$

Pero $P_\Omega(x_k - \nabla f(x_k)) \in D$, luego, $A(P_\Omega(x_k - \nabla f(x_k))) - Ax_k = 0$ y se obtiene que

$$A\nabla f(x_k) = AA^T \lambda.$$

Esto implica que $\lambda = (AA^T)^{-1}A\nabla f(x_k)$. Reemplazando en (72) obtenemos que

$$P_\Omega(x_k - \nabla f(x_k)) - (x_k - \nabla f(x_k)) = A^T(AA^T)^{-1}A\nabla f(x_k)$$

es decir

$$P_D(x_k - \nabla f(x_k)) = x_k - (I - A^T(AA^T)^{-1}A)\nabla f(x_k)$$

Observar que la condición de optimalidad del Teorema 5.1 se puede establecer, para el problema (71), de la siguiente manera:

$$P_\Omega(x^* - \nabla f(x^*)) - x^* = 0 \text{ si y solo si } (I - A^T(AA^T)^{-1}A)(-\nabla f(x^*)) = 0.$$

Para cada vector $u \in \mathbb{R}^n$ las proyecciones:

$$P_{Nu(A)}(u) = (I - A^T(AA^T)^{-1}A)u, \quad (73)$$

$$P_{R(A^T)}(u) = A^T(AA^T)^{-1}Au \quad (74)$$

son las proyecciones ortogonales sobre $Nu(A)$ y $R(A^T)$ respectivamente. La matriz $A^+ = A^T(AA^T)^{-1}$ se llama pseudoinversa de A . Cumple $AA^+ = I$, $A^+A = [P_{R(A^T)}]$.

Método del gradiente proyectado para problemas con restricciones lineales de igualdad. Dado x_0 factible inicial. En cada iteración k hacer los siguientes pasos:

1. Calcular $d_k = (I - A^T(AA^T)^{-1}A)(-\nabla f(x_k)) = P_\Omega(x_k - \nabla f(x_k)) - x_k$.

Si $d_k = 0$, parar.

Si no, ir a 2.

2. Hacer búsqueda de Armijo a partir de x_k en la dirección d_k :

Definir $t = 1$. Si

$$f(x_k + td_k) \leq f(x_k) + \sigma_1 t \nabla f(x_k)^T d_k \Rightarrow \quad (75)$$

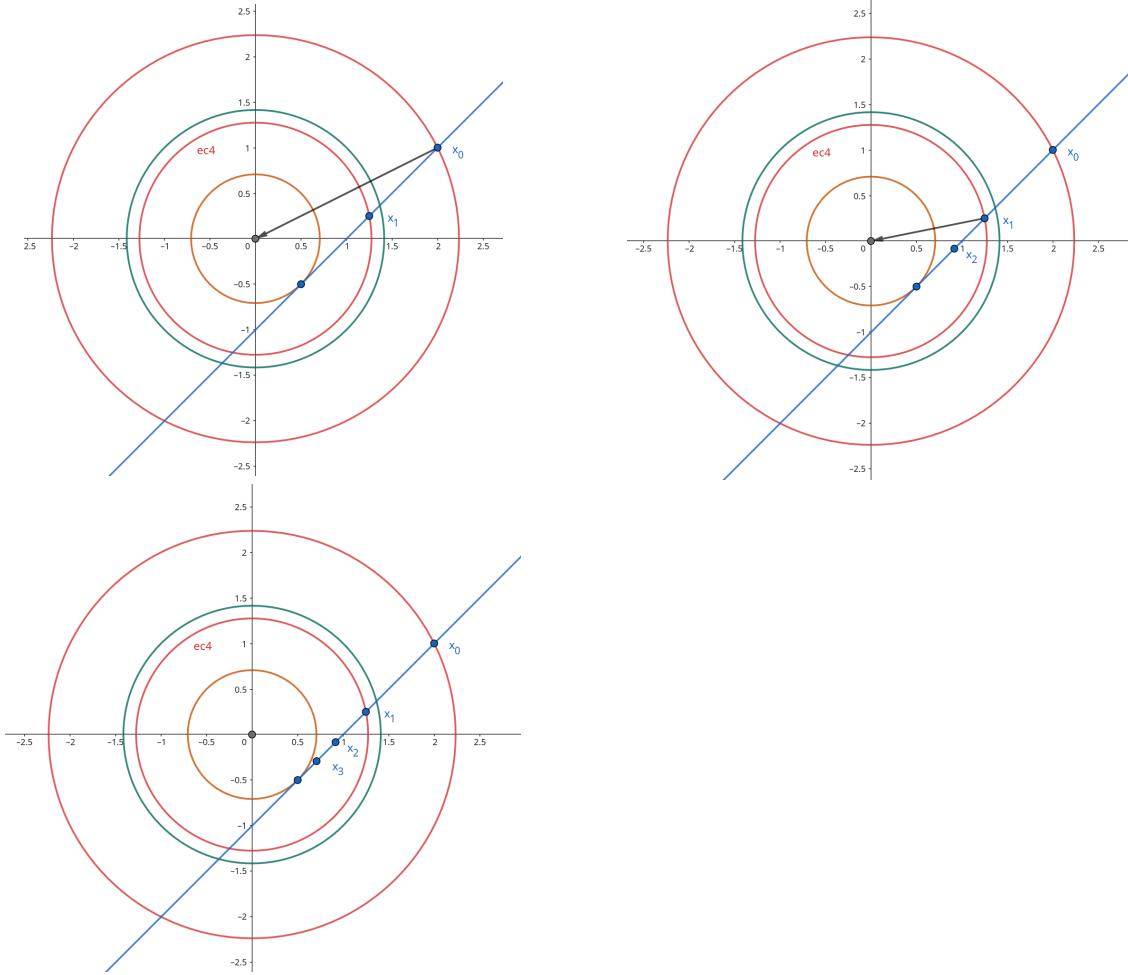
- (a) Definir $x_{k+1} = x_k + td_k \in \Omega$ e ir a 1.
- (b) si no, hacer $t = t/2$ e ir a (75).

Aplicar el método de gradiente proyectado al problema con restricciones lineales: minimizar $1/2(x^2 + y^2)$ sujeto a $y = -1 + x$ comenzando desde $(2, 1)$ con $\sigma_1 = 2/3$.

Se tiene que $A = (-1 \ 1)$ entonces

$$(I - A^T(AA^T)^{-1}A) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

Se tienen las siguientes iteraciones:



Observaciones:

1. La dirección d_k es una dirección factible y de descenso para f en x_k . (Dado un punto factible x_0 , decimos que d es una dirección es factible a partir de x_0 para el problema (71) si cumple:

$$A(x_0 + td) = b \text{ para } t > 0,$$

es decir, debe ser $Ad = 0$.)

2. El método es la versión del método del gradiente proyectado de la sección anterior adaptado para el problema (71).
3. Se puede demostrar que si x_k no es un punto estacionario y el algoritmo genera una sucesión infinita de iterados entonces todo punto de acumulación es un punto KKT de (71).

Un caso particular del problema (71) es el problema en el cual la función objetivo es una función cuadrática:

$$\begin{aligned} & \text{Minimizar } 1/2x^T Qx + c^T x \\ & \text{sujeto a } Ax = b, \end{aligned} \quad (76)$$

donde $A \in \mathbb{R}^{m \times n}$, $m < n$, $\text{rango}(A) = m$, $Q \in \mathbb{R}^{n \times n}$ simétrica y $c \in \mathbb{R}^n$. Se trata de un problema llamado de programación cuadrática.

En este caso $\nabla f(x) = Qx + c$ y las condiciones KKT para este problema son equivalentes al sistema:

$$\begin{array}{lcl} Qx + A^T \lambda & = -c \\ Ax & = b \end{array} \Leftrightarrow \begin{pmatrix} Q & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix} \quad (77)$$

El siguiente teorema establece condiciones sobre A y Q bajo las cuales la matriz del sistema anterior es no singular lo que asegura que existe solución del sistema KKT.

Teorema 5.3. Si $x^T Qx > 0 \forall x \neq 0 : Ax = 0$ entonces la matriz $\begin{pmatrix} Q & A^T \\ A & 0 \end{pmatrix}$ es no singular.

Ejemplo. Considerar el problema de minimizar $q(x, y, z) = 3x^2 + 2xy + xz - 2.5y^2 + 3z^2 - 8x - 3z$ sujeto a $x + z = 3$, $y + z = 0$.

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

$$\nabla q(x, y, z) = \begin{pmatrix} 6x + 2y + z - 8 \\ 2x - 5y \\ x + 6z - 3 \end{pmatrix}, Q = \begin{pmatrix} 6 & 2 & 1 \\ 2 & -5 & 0 \\ 1 & 0 & 6 \end{pmatrix}$$

$$\{x : Ax = 0\} = \{(x, y, z) : x + z = 0, y + z = 0\} = \{(-z, -z, z)\} = \overline{(-1, -1, 1)}.$$

Luego, consideramos el vector $(-1, -1, 1)$:

$$\begin{pmatrix} -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 6 & 2 & 1 \\ 2 & -5 & 0 \\ 1 & 0 & 6 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} -6 - 2 + 1 \\ -2 + 5 \\ -1 + 6 \end{pmatrix} =$$

$$\begin{pmatrix} -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} -7 \\ 3 \\ 5 \end{pmatrix} = 9 > 0$$

por lo tanto el problema tiene única solución.

Luego, una posibilidad podría ser tratar de resolver el sistema KKT (77) mediante algún método para resolver sistemas lineales.

En el caso del ejemplo:

$$\begin{pmatrix} 6 & 2 & 1 & 1 & 0 \\ 2 & -5 & 0 & 0 & 1 \\ 1 & 0 & 6 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \\ 3 \\ 3 \\ 0 \end{pmatrix}$$

5.3 Método de Conjuntos activos para programación cuadrática con restricciones lineales de desigualdad

En esta sección consideramos el problema cuadrático con restricciones lineales de desigualdad de la forma

$$\begin{aligned} & \text{Minimizar } \frac{1}{2}x^T Qx + c^T x \\ & \text{sujeto a } a_i^T x \leq b_i, i = 1, \dots, p \end{aligned} \quad (78)$$

donde $Q \in \mathbb{R}^{n \times n}$ es una matriz simétrica. El conjunto factible $\{x \in \mathbb{R}^n : a_i^T x \leq b_i, i = 1, \dots, p\}$ también se representa en forma matricial como $Ax \leq b$ donde A es la matriz en $\mathbb{R}^{p \times n}$ cuyas filas son los vectores a_i^T y asumimos que $p \leq n$.

Sabemos que un punto x^* es KKT del problema (78) si y solo si existen $\mu \in \mathbb{R}^p, \mu \geq 0$ tal que

$$\begin{aligned} Qx^* + c + A^T \mu &= 0 \\ a_i^T x^* \leq b_i, \mu_i (a_i^T x^* - b_i) &= 0, i = 1, \dots, p. \end{aligned}$$

Observar que, si A tiene rango completo ($r(A) = p$), el sistema

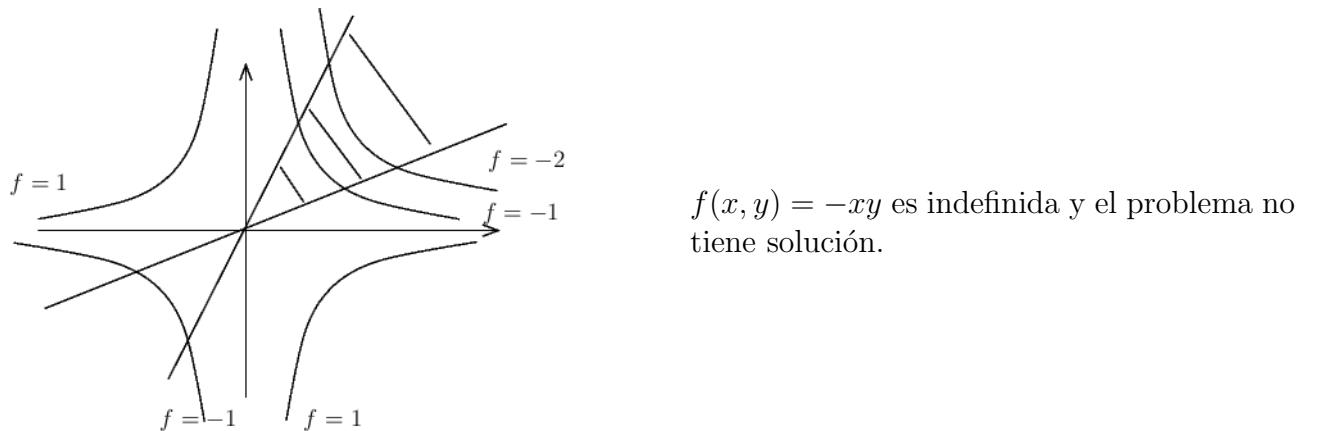
$$\nabla q(x^*) + A^T \mu = 0$$

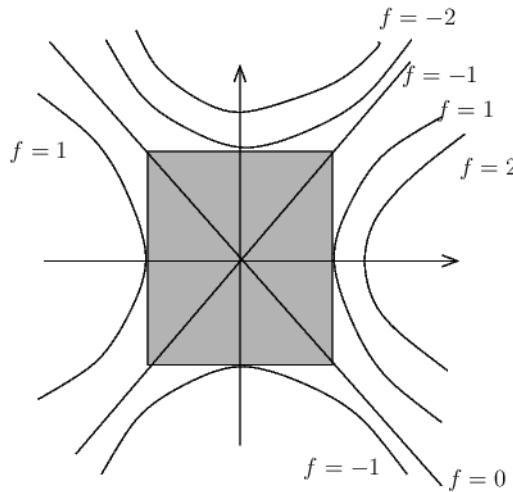
implica que $\mu = -(AA^T)^{-1}A\nabla q(x^*)$, donde $q(x) = \frac{1}{2}x^T Qx + c^T x$.

La región factible $\Omega = \{x \in \mathbb{R}^n : a_i^T x \leq b_i, i = 1, \dots, p\}$ es un poliedro, definido por la intersección de los semiespacios $a_i^T x \leq b_i$. Puede ser vacío, acotado o no acotado.

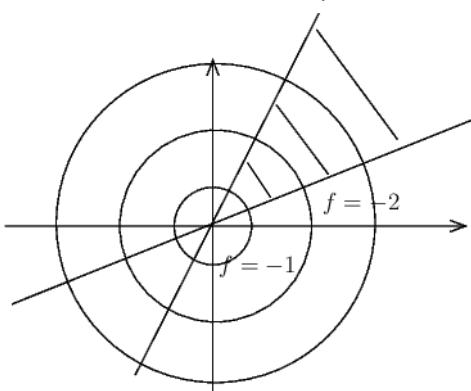
Si la matriz Q es definida positiva entonces la cuadrática es convexa y existe un único minimizador global.

Si Q no es definida positiva entonces, puede no haber minimizador local o puede haber más de un minimizador local que cumple las condiciones necesarias de segundo orden.





$f(x, y) = x^2 - y^2$ es indefinida y el problema tiene 2 minimizadores locales.



$f(x, y) = -(x^2 + y^2)$ es definida negativa y el problema no tiene solución.

Problemas con restricciones de desigualdad son más difíciles de resolver que problemas con restricciones de igualdad. La razón principal es que no se conocen de antemano las restricciones que son activas en la solución.

Una posibilidad podría ser resolver todos los posibles subproblemas formados por restricciones de igualdad y elegir la solución con menor valor funcional para la función objetivo. Pero aún para un problema pequeño la cantidad de subproblemas puede ser muy grande. Los métodos denominados de restricciones activas tratan de superar esta dificultad moviéndose sucesivamente de una elección de restricciones activas a otra de modo de reducir el valor objetivo de la función. Es decir, se mueven de un conjunto de restricciones activas a otro que produce al menos un punto más óptimo que el anterior. La esperanza es que solo algunos de los subproblemas con restricciones de igualdad sean considerados. En cada iteración de un método de conjuntos activos, se elige un conjunto de trabajo formado por restricciones que se asumen activas en el óptimo y se minimiza la función objetivo en ese conjunto de trabajo formado por restricciones de igualdad.

Veamos cómo funciona un método de restricciones activas. Asumimos que Q es definida positiva. Recordemos que estamos buscando un punto factible x^* para el cual

$$-\nabla q(x^*) = A^T \mu$$

con $\mu \geq 0$, $\mu_i(a_i^T x^* - b_i) = 0$, $i = 1, \dots, p$.

Dado un punto x_0 factible, en un iterado x_k se identifican las restricciones que son activas en ese punto:

$$\mathcal{A}(x_k) = \{i : a_i^T x_k = b_i\}$$

y se plantea el subproblema

$$\begin{aligned} & \text{Minimizar } q(x) \\ & \text{sujeto a } x \in W_k = \{x \in \mathbb{R}^n : a_i^T x = b_i, i \in \mathcal{A}(x_k)\}. \end{aligned} \tag{79}$$

W_k es el conjunto de trabajo de esta iteración.

Se define la matriz $A_k = [a_i]_{i \in \mathcal{A}(x_k)}$ formada a partir de A por las filas que se corresponden con índices en $\mathcal{A}(x_k)$. Si \tilde{x} es una solución del problema (79) entonces $d_k = \tilde{x} - x_k$ es solución del problema

$$\begin{aligned} & \text{Minimizar } \frac{1}{2} d^T Q d + \nabla q(x_k)^T d \\ & \text{sujeto a } A_k d = 0. \end{aligned} \quad (80)$$

Consideramos dos posibilidades:

1. $d_k = 0$.

En este caso, $\tilde{x} = x_k$ y como x_k es solución de (79), por las condiciones KKT sabemos que existe μ tal que $\nabla q(x_k) + A_k^T \mu = 0$. Si $\{a_i\}_{i \in \mathcal{A}(x_k)}$ son linealmente independientes entonces $\mu = -(A_k A_k^T)^{-1} A_k \nabla q(x_k)$.

Si $\mu \geq 0$ entonces x_k cumple las condiciones KKT para (78).

Si algún $\mu_i < 0, i \in \mathcal{A}(x_k)$ entonces x_k no cumple las condiciones KKT y tenemos que abandonar alguna restricción del conjunto $\mathcal{A}(x_k)$ para la cual $\mu_i < 0$.

2. $d_k \neq 0$.

En este caso se busca el paso máximo que se puede dar en esa dirección sin salir de la región factible. Aquí intervienen las restricciones que no se encuentran en el conjunto de trabajo.

Si $j \notin \mathcal{A}(x_k)$ entonces $a_j^T(x_k + t d_k) = a_j^T x_k + t a_j^T d_k$.

Si $a_j^T d_k \leq 0$ entonces $a_j^T(x_k + t d_k) = a_j^T x_k + t a_j^T d_k \leq b_j$ y no se alcanza el borde de la región.

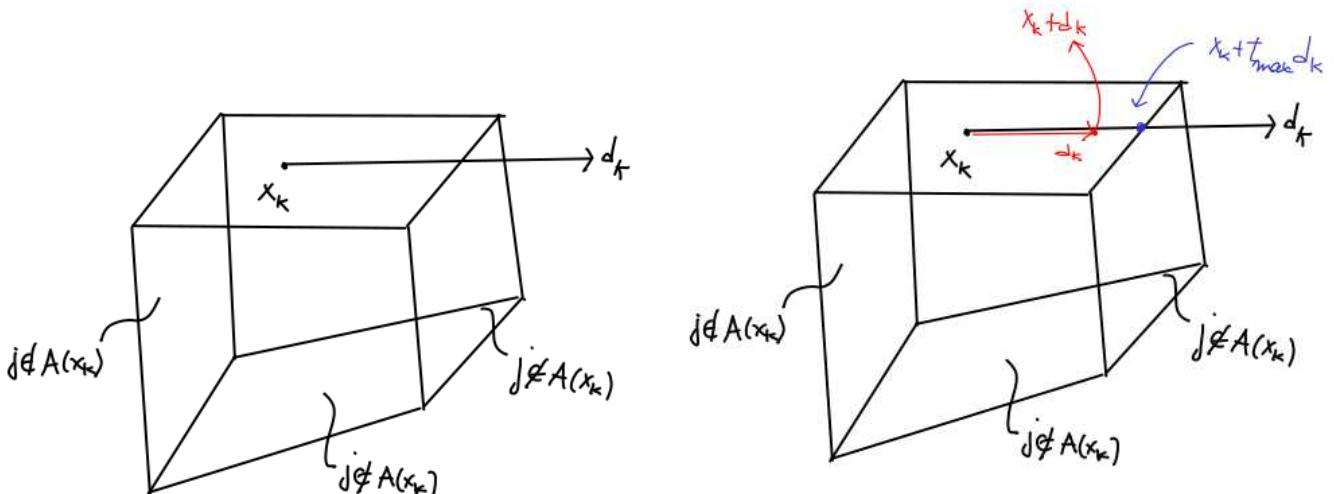
Si $a_j^T d_k > 0$ entonces $a_j^T(x_k + t d_k) = a_j^T x_k + t a_j^T d_k = b_j$ si y solo si $t = \frac{b_j - a_j^T x_k}{a_j^T d_k}$.

Luego, el paso máximo que se puede dar en la dirección sin salir del conjunto factible es

$$t_{max} = \min_{j \notin \mathcal{A}(x_k), a_j^T d_k > 0} \left\{ \frac{b_j - a_j^T x_k}{a_j^T d_k} \right\}.$$

Si $t_{max} \leq 1$ entonces se ha alcanzado alguna restricción. Supongamos que fue la restricción j , luego, esa restricción se debe incorporar al nuevo conjunto de trabajo. Se define $t_k = t_{max}, x_{k+1} = x_k + t_k d_k, W_{k+1} = W_k \cup \{j\}$.

Si no, entonces no se ha alcanzado ninguna restricción y se define $t_k = 1, x_{k+1} = x_k + d_k, W_{k+1} = W_k$ (no cambia el conjunto de trabajo).



Algoritmo de conjuntos activos. Sean x_0 inicial, factible, $\mathcal{A}(x_0), [a_i]_{i \in \mathcal{A}(x_0)}$ linealmente independientes. En cada iteración k :

Paso 1. Resolver el subproblema (80) correspondiente.

Si $d_k = 0$ ir al Paso 3. Sino

Paso 2. Calcular t_{max} .

Si $t_{max} \leq 1$ entonces definir $t_k = t_{max}, x_{k+1} = x_k + t_k d_k, W_{k+1} = W_k \cup \{j\}$ siendo j el índice para el cual $t_k = \frac{b_j - a_j^T x_k}{a_j^T d_k}$ e ir al Paso 1.

Si no, entonces definir $t_k = 1, x_{k+1} = x_k + d_k, W_{k+1} = W_k$, hacer $k = k + 1$ e ir al Paso 1.

Paso 3. Si $d_k = 0$, calcular $\mu = -(A_k A_k^T)^{-1} A_k \nabla q(x_k)$.

Si $\mu \geq 0$ terminar. Sino, definir $x_{k+1} = x_k, \mathcal{A}(x_{k+1}) = \mathcal{A}(x_k) - \{j\}$, siendo j tal que la coordenada j de μ es negativa e ir al Paso 1.

Ejemplo. Considerar el siguiente problema:

$$\begin{aligned} & \text{Min } 2x^2 + xy + y^2 - 12x - 10y \\ & \text{s. a } x + y \leq 4, x \geq 0, y \geq 0. \end{aligned}$$

$$\begin{aligned} \text{Para } x_0 = (0, 0), \text{ se tiene que } \mathcal{A}(x_0) = \{2, 3\}. \quad A_0 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\ & \text{Min } \frac{1}{2} d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d + (-12 - 10)^T d \\ & \text{s. a } A_0 d = 0 \end{aligned}$$

Luego, $d_0 = 0$. Calculamos μ :

$$\mu = -(A_0 A_0^T)^{-1} A_0 \nabla q(x_0) = \begin{pmatrix} -12 \\ -10 \end{pmatrix} \leq 0$$

entonces $x_1 = x_0, \mathcal{A}(x_1) = \mathcal{A}(x_0) - \{2\} = \{3\}$.

$$\begin{aligned} 1. \quad A_1 &= \begin{pmatrix} 0 & -1 \end{pmatrix} \\ & \text{Min } \frac{1}{2} d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d + (-12 - 10)^T d \\ & \text{s. a } A_1 d = 0. \end{aligned}$$

Luego, $d = (d_1, 0)$, entonces $q(d) = 1/2(d_1 0) \begin{pmatrix} 4d_1 \\ d_1 \end{pmatrix} - 12d_1 = 2d_1^2 - 12d_1$ y se tiene que $d_1 = (3 \ 0)$.

Luego

$$a_1^T d_1 = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = 3 > 0, \text{ y}$$

$$\frac{b_1 - a_1^T x_1}{a_1^T d_1} = \frac{4}{3}.$$

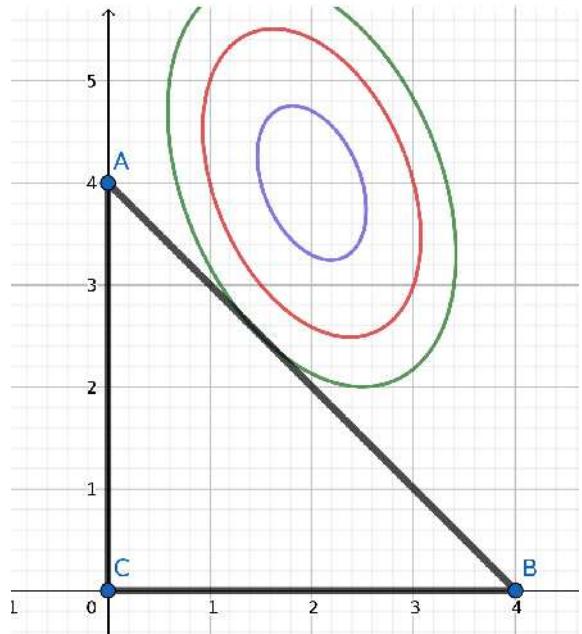
$$a_2^T d_1 = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = -3 < 0$$

entonces

$$t_{max} = 4/3 > 1$$

y

$$t_1 = 1$$



Luego, $x_2 = x_1 + d_1 = \begin{pmatrix} 3 & 0 \end{pmatrix}$, $\mathcal{A}(x_2) = \mathcal{A}(x_1) = \{3\}$, $A_2 = A_1 = \begin{pmatrix} 0 & -1 \end{pmatrix}$.

2. $A_2 = \begin{pmatrix} 0 & -1 \end{pmatrix}$

$$\begin{aligned} \text{Min } & \frac{1}{2} d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d + (0 - 7)^T d \\ \text{s. a } & A_2 d = 0 \end{aligned}$$

Luego, $d = (d_1, 0)$, entonces $q(d) = 1/2(d_1 0) \begin{pmatrix} 4d_1 \\ d_1 \end{pmatrix} = 4d_1^2$ y se tiene que $d_2 = (0, 0)$. Calculamos μ :

$$\mu = -(A_2 A_2^T)^{-1} A_2 \nabla q(x_2) = -7 \leq 0$$

entonces $x_3 = x_2$, $\mathcal{A}(x_3) = \emptyset$, $A_3 = 0$.

3. $A_3 = 0$

$$\text{Min } \frac{1}{2} d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d + (0 - 7)^T d$$

Luego, $q(d) = 1/2d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d + (0 - 7)^T d$ y se tiene que $d_3 = \begin{pmatrix} -1 & 4 \end{pmatrix}$.

$$a_1^T d_3 = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = 3 > 0, \text{ y}$$

$$\frac{b_1 - a_1^T x_3}{a_1^T d_3} = \frac{4 - \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix}}{3} = \frac{1}{3}.$$

$$a_2^T d_3 = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = 1 > 0 \text{ y}$$

$$\frac{b_2 - a_2^T x_3}{a_2^T d_3} = \frac{0 - \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix}}{1} = \frac{3}{1}.$$

$$a_3^T d_3 = \begin{pmatrix} 0 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = -4 < 0$$

entonces

$$t_{max} = \min\{1/3, 3\} = 1/3 < 1$$

y

$$t_3 = 1/3$$

Luego, $x_4 = x_3 + t_3 d_3 = \begin{pmatrix} 3 & 0 \end{pmatrix} + 1/3 \begin{pmatrix} -1 & 4 \end{pmatrix} = \begin{pmatrix} 8/3 & 4/3 \end{pmatrix}$, $\mathcal{A}(x_4) = \{1\}$, $A_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}$.
4. $A_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}$.

$$\begin{aligned} \text{Min } & \frac{1}{2} d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d + (0 - 14/3)^T d \\ \text{s. a } & A_4 d = 0 \end{aligned}$$

Luego, $d_2 = -d_1$ y $q(d) = 1/2 d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d - \begin{pmatrix} 0 & -14/3 \end{pmatrix}^T d = 2d_1^2 + 14/3d_1$ y se tiene que $d_4 = \begin{pmatrix} -7/6 & 7/6 \end{pmatrix}$.

$$a_2^T d_4 = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} -7/6 \\ 7/6 \end{pmatrix} = 7/6 > 0 \text{ y}$$

$$\frac{b_2 - a_2^T x_4}{a_2^T d_4} = \frac{0 - \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 8/3 \\ 4/3 \end{pmatrix}}{7/6} = \frac{8/3}{7/6} =$$

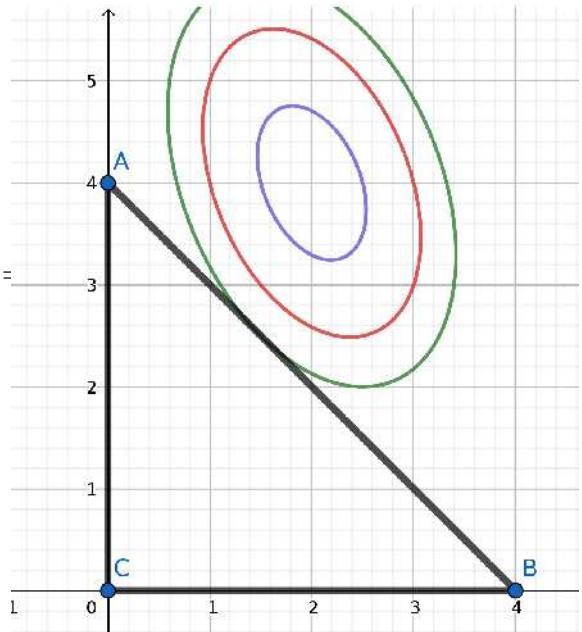
$$a_3^T d_4 = \begin{pmatrix} 0 & -1 \end{pmatrix} \begin{pmatrix} -7/6 \\ 7/6 \end{pmatrix} = -7/6 < 0$$

entonces

$$t_{max} = 16/7 > 1$$

y

$$t_4 = 1$$



Luego, $x_5 = x_4 + d_4 = \begin{pmatrix} 8/3 & 4/3 \end{pmatrix} + \begin{pmatrix} -7/6 & 7/6 \end{pmatrix} = \begin{pmatrix} 3/2 & 5/2 \end{pmatrix}$, $\mathcal{A}(x_5) = \{1\}$, $A_5 = \begin{pmatrix} 1 & 1 \end{pmatrix}$.
5. $A_5 = \begin{pmatrix} 1 & 1 \end{pmatrix}$.

$$\begin{aligned} \text{Min } & \frac{1}{2} d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d + (-7/2 - 7/2)^T d \\ \text{s. a } & A_5 d = 0 \end{aligned}$$

Luego, $d_2 = -d_1$ y $q(d) = 1/2 d^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} d - \begin{pmatrix} -7/2 & -7/2 \end{pmatrix}^T d = 2d_1^2$.

Luego, $d_5 = \begin{pmatrix} 0 & 0 \end{pmatrix}$ y calculamos μ :

$$\mu = -(A_5 A_5^T)^{-1} A_5 \nabla q(x_5) = -1/2 \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} -7/2 \\ -7/2 \end{pmatrix} > 0$$

entonces x_5 es la solución.

Propiedades:

1. Sea d solución del subproblema

$$\begin{aligned} \text{Min } & \frac{1}{2} d^T Q d + g_k^T d \\ \text{s. a } & A_k d = 0 \end{aligned}$$

donde $g_k = \nabla q(x_k)$, $A_k = [a_i]_{i \in \mathcal{A}(x_k)}$. Entonces, si $d \neq 0$, d es una dirección factible de descenso para q en x_k .

2. Si x_0 es factible entonces x_k es factible para todo k .

3. Si $\{a_i\}_{i \in \mathcal{A}(x_k)}$ es LI entonces $\{a_i\}_{i \in \mathcal{A}(x_{k+1})}$ es LI: tenemos que analizar el caso en que una restricción se agrega al conjunto de trabajo. Supongamos que $\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{j\}$. Es decir $a_j^T d > 0$. Si $\{a_i\}_{i \in \mathcal{A}(x_k)}$ es LI y suponemos que $\{a_i\}_{i \in \mathcal{A}(x_{k+1})}$ no lo es, tenemos que $a_j = \sum_{j \in \mathcal{A}(x_k)} \alpha_i a_i$. Entonces $a_j^T d = \sum_{j \in \mathcal{A}(x_k)} \alpha_i a_i^T d = 0$ que es absurdo.

Teorema. Si Q es simétrica definida positiva y $\Omega = \{x : Ax \leq b\}$ es acotado, en un número finito de iteraciones el método encuentra un punto estacionario.

Demostración.

1. Si $d_k = 0$ entonces, como Q es definida positiva, x_k es la única solución global de $q(x)$ en el conjunto de trabajo W_k .

Si $\mu_k \not\geq 0$ entonces se elimina una restricción y se calcula una nueva d_{k+1} a partir de x_{k+1} (en este caso se tiene que $x_{k+1} = x_k$). Luego, en el próximo iterado x_{k+2} se tiene que $q(x_{k+2}) < q(x_{k+1})$.

Luego, como x_k es la solución en W_k el algoritmo no vuelve a explorar W_k .

2. Cuando se tiene $d_k \neq 0$ entonces hay dos posibilidades:

- (a) Se modifica el conjunto de trabajo porque se agrega una restricción. Entonces, en este caso, en a lo sumo n iteraciones se llega a n vectores LI. Luego, se obtendrá una dirección nula.
- (b) Si no, se tiene que $t_k = 1$ entonces $x_k + d_k$ es el minimizador de $q(x)$ en W_k entonces en la próxima iteración $d_{k+1} = 0$.

Por lo tanto, en a lo sumo finitas iteraciones se encuentra el minimizador global de q en el conjunto de trabajo y no vuelve a visitar un conjunto de trabajo. Como hay solo un número finito de conjuntos de trabajo el proceso es finito.

Este método se puede extender para resolver problemas de optimización de funciones generales con restricciones lineales de desigualdad.

5.4 Ejercicios.

1. Elegir uno de los siguientes problemas y resolverlo, a partir del punto inicial mencionado, aplicando el método del gradiente proyectado con búsqueda de Armijo ($\sigma_1 = 1/2$) presentado previamente.

$$a) \begin{array}{ll} \text{Min} & x_1 - x_2 \\ \text{s.a} & 1 \leq x_1 \leq 3 \\ & 1 \leq x_2 \leq 2. \\ x_0 = (3, 1) & \end{array}$$

$$b) \begin{array}{ll} \text{Min} & x_1^2 + x_2^2 \\ \text{s.a} & 0 \leq x_1 \leq 4 \\ & 1 \leq x_2 \leq 3. \\ x_0 = (4, 3) & \end{array}$$

2. Demostrar el Teorema 5.3: Para $A \in \mathbb{R}^{m \times n}, m < n$, $\text{rango}(A) = m$, $Q \in \mathbb{R}^{n \times n}$ simétrica y $c \in \mathbb{R}^n$, si $x^T Q x > 0 \forall x \neq 0 : Ax = 0$ entonces la matriz $\begin{pmatrix} Q & A^T \\ A & 0 \end{pmatrix}$ es no singular. (Pensar en el absurdo)
3. En los siguientes problemas cuadráticos estudiar si se verifican las hipótesis del Teorema 5.3 y analizar en qué caso se puede afirmar que el problema tiene solución única.
- (a) $\text{Min } 5x_1^2 + 4x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 - x_2 + 4x_3$ sujeto a $x_1 + 2x_3 = 4, x_2 + 3x_3 = 2$.
 - (b) $\text{Min } 3x_1^2 + 6x_2^2 + 3x_3^2 - 4x_1x_2 + 8x_1x_3 + 4x_2x_3 + 2x_1 + 3x_2 + x_3$ sujeto a $x_1 + x_3 = 5, 2x_2 + x_3 = 1$.
 - (c) $\text{Min } 6x_1^2 + 6x_2^2 + 5x_3^2 - 4x_1x_2 - 2x_1x_3 - 2x_2x_3 + 3x_1 - 2x_2 - x_3$ sujeto a $x_1 + x_2 = 1$.

4. Elegir uno de los siguientes problemas y resolverlo utilizando el método de restricciones activas comenzando desde los puntos indicados. Analizar el proceso gráficamente cuando sea posible.

$$a) \begin{array}{ll} \text{Min} & (x_1 - 1)^2 + (x_2 - 2, 5)^2 \\ \text{s.a} & -x_1 + 2x_2 \leq 2 \\ & x_1 + 2x_2 \leq 6 \\ & x_1 - 2x_2 \leq 2 \\ & x_1 \geq 0, x_2 \geq 0. \end{array}$$

$$x_0 = (2, 0)$$

$$b) \begin{array}{ll} \text{Min} & x_1^2 - x_1x_2 + x_2^2 - 3x_1 \\ \text{s.a} & x_1 + x_2 \leq 4 \\ & x_1 \geq 0, x_2 \geq 0. \end{array}$$

$$x_0 = (0, 0)$$

6 Métodos numéricos para problemas generales de optimización

En esta sección presentamos algunas metodologías para resolver problemas generales del optimización.

Sabemos que un problema general de optimización no lineal se puede modelizar de la forma

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0, g(x) \leq 0 \quad (81)$$

donde las funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ son todas diferenciables hasta el orden que sea necesario.

La resolución de estos problemas trae aparejado un compromiso entre dos conflictos:

- minimizar $f(x)$ (optimalidad),
- verificar las restricciones $h(x) = 0, g(x) \leq 0$ (factibilidad).

Muchas de las metodologías y técnicas que se consideran para resolver problemas con restricciones generales introducen una función de mérito con el objetivo de balancear los dos conflictos: optimalidad y factibilidad.

Dado un parámetro p , una función $\Phi(x, p)$ es una función de mérito si los puntos $x(p)$ para cada p fijo soluciones del problema de Minimizar $\Phi(x, p)$ respecto de x tienden a un minimizador del problema original cuando $p \rightarrow \infty$.

Es decir, se busca una solución del problema original mediante la resolución de una sucesión de problemas irrestrictos en los cuales se minimiza una función de mérito que compromete los dos objetivos esenciales en la resolución del problema general con restricciones.

Los métodos llamados de penalidad, Lagrangiano Aumentado y barrera pertenecen a esta clase de metodología.

6.1 Métodos de penalidad

Este tipo de métodos busca la solución del problema con restricciones no lineales generales mediante el límite de soluciones de problemas sin restricciones (o con restricciones más fáciles).

En el caso de métodos de penalidad la función original es reemplazada por una función que se forma mediante la suma de la función objetivo más un término que penaliza el hecho de que las restricciones no se verifiquen.

En el caso de métodos de barrera, a la función objetivo se le suma un término que favorece puntos interiores de la región factible sobre los puntos que están en la frontera.

Asociados a estos métodos hay un parámetro, de penalidad o de barrera, que determina la severidad del castigo.

Para fijar ideas, consideremos de forma general el problema de Minimizar $f(x)$ sujeto a $x \in S$, donde S es el conjunto factible que puede estar formado por restricciones de igualdad y desigualdad generales como en (81).

Se reemplaza el problema original por un problema sin restricciones de la forma Minimizar $f(x) + \rho P(x)$ donde $\rho > 0$ es llamado parámetro de penalidad y P es una función de \mathbb{R}^n que tiene por objetivo penalizar las restricciones que definen al conjunto factible S . La función P debe cumplir las siguientes condiciones:

- ser continua,
- $P(x) \geq 0$,
- $P(x) = 0$ si y solo si $x \in S$.

Por ejemplo, si $S = \{x : h_i(x) = 0, i = 1, \dots, m\}$ pueden considerarse las funciones

$$1. P(x) = \sum_{i=1}^m (h_i(x))^2, \text{ llamada función de penalidad cuadrática,}$$

$$2. P(x) = \sum_{i=1}^m |h_i(x)|.$$

Si $S = \{x : g_i(x) \leq 0, i = 1, \dots, p\}$ pueden considerarse las funciones

$$1. P(x) = \sum_{i=1}^p (\max\{0, g_i(x)\})^2, \text{ llamada función de penalidad cuadrática,}$$

$$2. P(x) = \sum_{i=1}^p \max\{0, g_i(x)\}.$$

Una vez elegida un función de penalidad, se plantea la resolución de una sucesión de problemas irrestrictos:

$$\begin{aligned} & \text{Minimizar } f(x) + \rho_k P(x) \\ & \text{s. a } x \in \mathbb{R}^n \end{aligned} \tag{82}$$

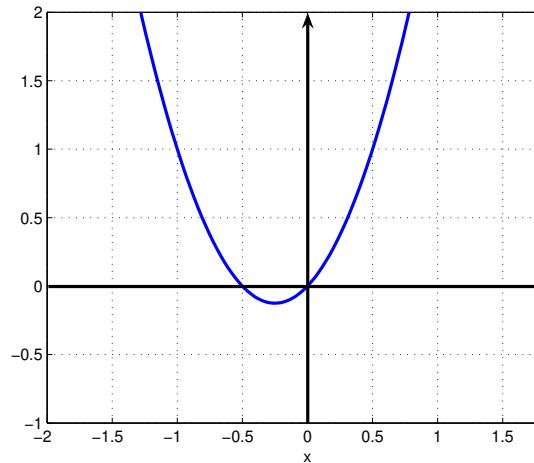
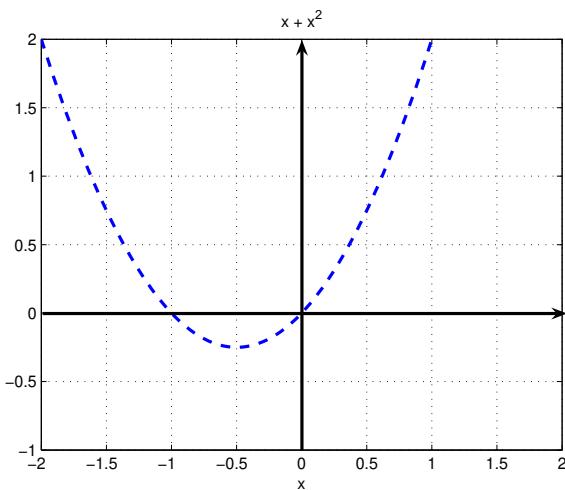
para una sucesión de parámetros diferentes $\{\rho_k\}$ que tiende a infinito. En general se considera $\{\rho_k\}$ tal que $\rho_{k+1} \geq \rho_k$, es decir, a medida que ρ_k aumenta, la no verificación de las restricciones se torna más cara para (82). Podemos esperar que las soluciones x_k de los problemas irrestrictos (82) se aproximen al conjunto factible cuando $\rho_k \rightarrow \infty$ y los puntos límites de la sucesión $\{x_k\}$, si existen, sean soluciones del problema original.

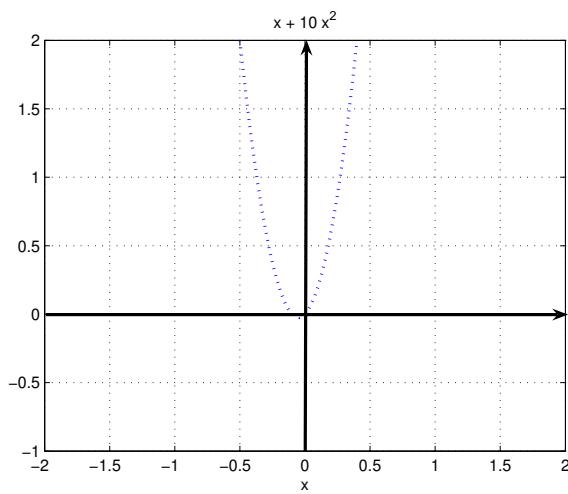
Ejemplo 1. Consideraremos el siguiente problema simple con restricciones de igualdad

$$\begin{aligned} & \text{Min } x \\ & \text{s.a } x = 0 \end{aligned}$$

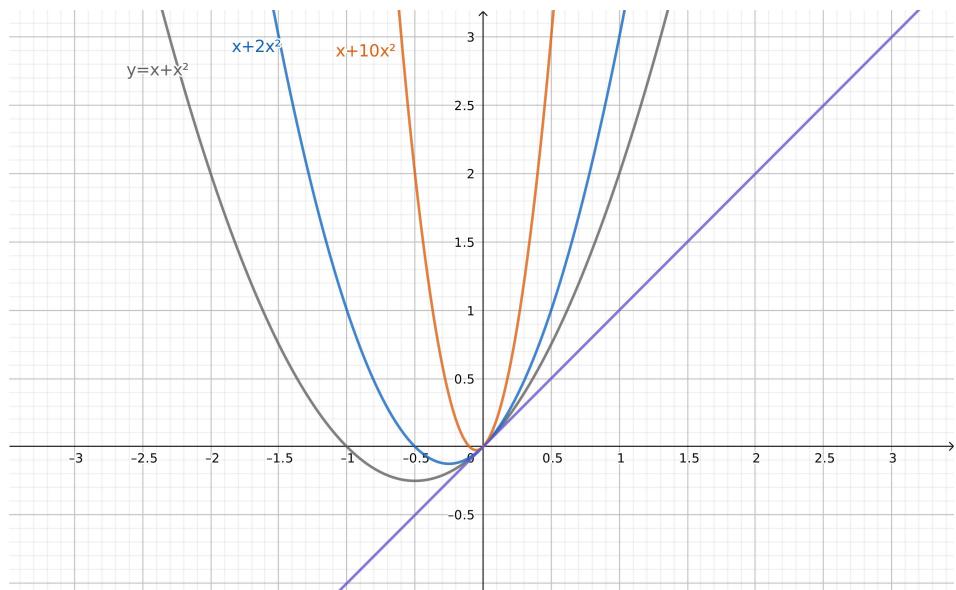
y la función de penalidad cuadrática $P(x) = x^2$.

Las siguientes son las gráficas de la función $\Phi(x, \rho) = f(x) + \rho P(x)$ para los valores de $\rho = 1, 2, 10$





Observar las tres gráficas juntas:



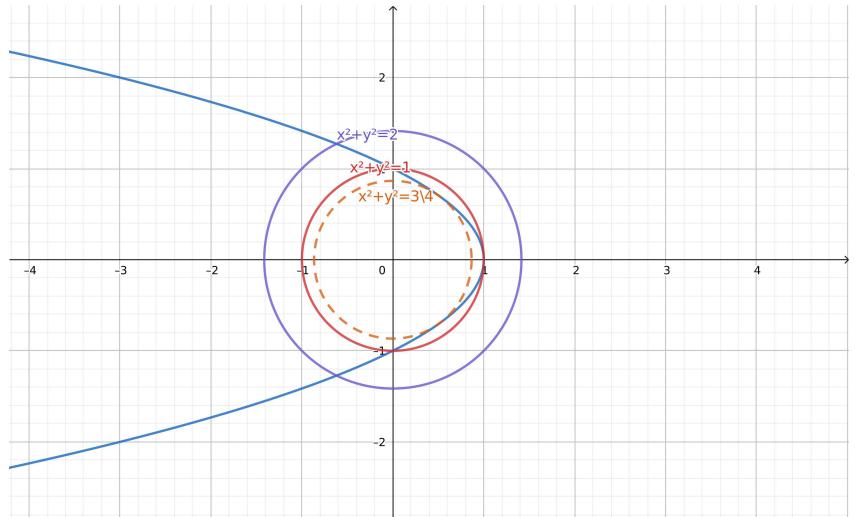
Ejemplo 2. Consideramos el problema

$$\begin{aligned} \text{Min } & x^2 + y^2 \\ \text{s. a } & x + y^2 = 1 \end{aligned}$$

El problema tiene 2 minimizadores locales:

$$\left(\frac{1}{2}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{2}, -\frac{1}{\sqrt{2}}\right)$$

(valen las condiciones suficientes de segundo orden)



En este caso:

$$\Phi((x, y), \rho) = x^2 + y^2 + \rho(x + y^2 - 1)^2$$

luego,

$$\nabla \Phi((x, y), \rho) = \begin{pmatrix} 2x + 2\rho(x + y^2 - 1) \\ 2y + 2\rho(x + y^2 - 1)2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

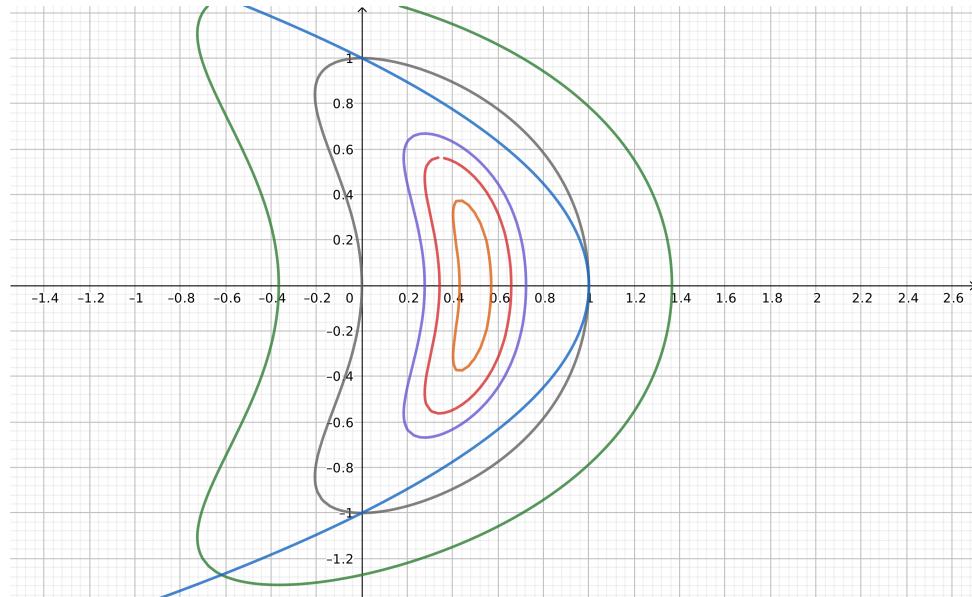
si y solo si

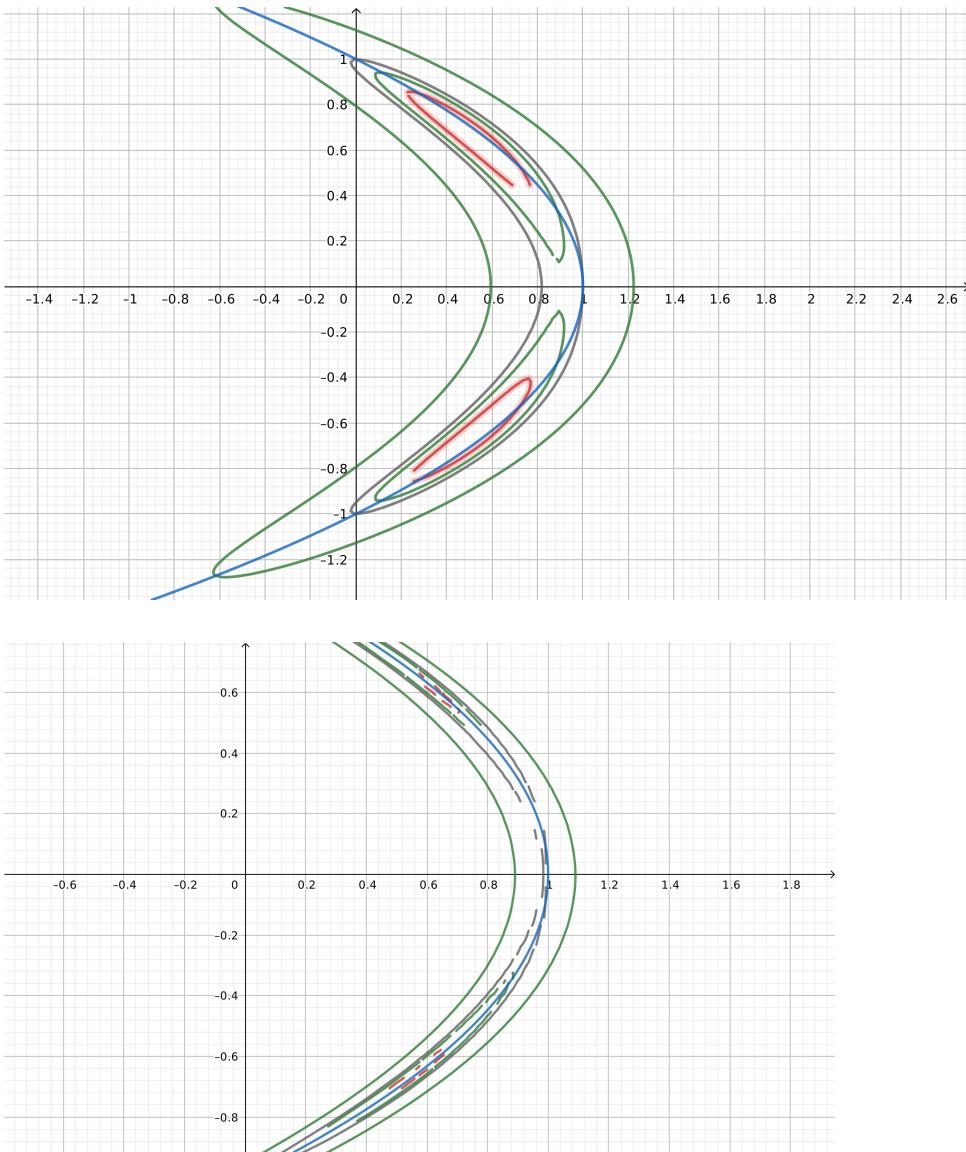
$$y = 0 \text{ o } 1 + 2\rho(x + y^2 - 1) = 0.$$

Luego, las posibles soluciones son

$$\left(1 - \frac{1}{\rho + 1}, 0\right) \text{ o } \left(\frac{1}{2}, \pm\sqrt{\frac{1}{2} - \frac{1}{2\rho}}\right).$$

Observar las curvas de nivel de Φ para $\rho = 1, 10, 100$





Analíticamente:

Si analizamos el Hessiano de $\Phi((x, y), \rho)$ en esos puntos tenemos

$$\nabla_x^2 \Phi((x, y), \rho) = \begin{pmatrix} 2 + 2\rho & 4\rho y \\ 4\rho y & 2 + 4\rho(x + y^2 - 1) + 8\rho y^2 \end{pmatrix}.$$

Luego,

$$\nabla_x^2 \Phi\left(1 - \frac{1}{\rho + 1}, 0\right) = \begin{pmatrix} 2 + 2\rho & 0 \\ 0 & \frac{2 - 2\rho}{1 + \rho} \end{pmatrix}$$

y no se cumplen las condiciones suficientes de segundo orden para $\rho \geq 1$.

$$\nabla_x^2 \Phi\left(\frac{1}{2}, \pm \sqrt{\frac{1}{2} - \frac{1}{2\rho}}\right) = \begin{pmatrix} 2 + 2\rho & 4\rho\left(\pm \sqrt{\frac{1}{2} - \frac{1}{2\rho}}\right) \\ 4\rho\left(\pm \sqrt{\frac{1}{2} - \frac{1}{2\rho}}\right) & 4\rho - 4 \end{pmatrix}.$$

$$\det(\nabla_x^2 \Phi\left(\frac{1}{2}, \pm \sqrt{\frac{1}{2} - \frac{1}{2\rho}}\right)) = 8\rho - 8 > 0 \text{ si y solo si } \rho > 1$$

entonces son minimizadores locales de Φ .

Observar que

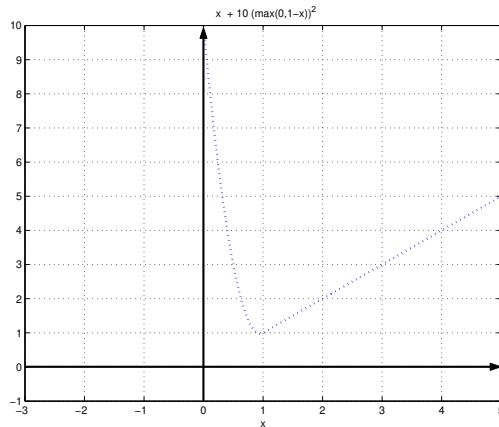
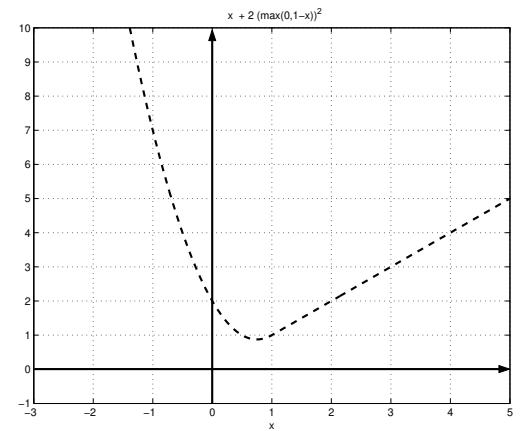
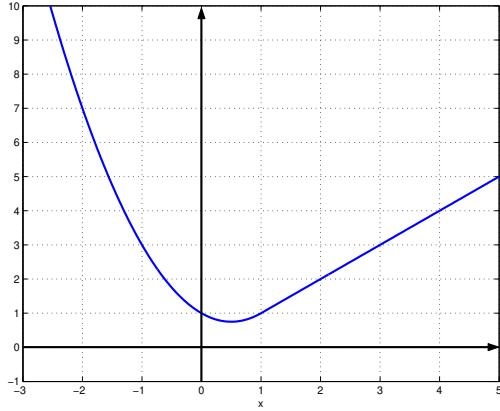
$$\lim_{\rho \rightarrow \infty} \left(\frac{1}{2}, \pm \sqrt{\frac{1}{2} - \frac{1}{2\rho}} \right) = \left(\frac{1}{2}, \pm \sqrt{\frac{1}{2}} \right). \quad (83)$$

Ejemplo 3. Consideremos el problema con restricciones de desigualdad

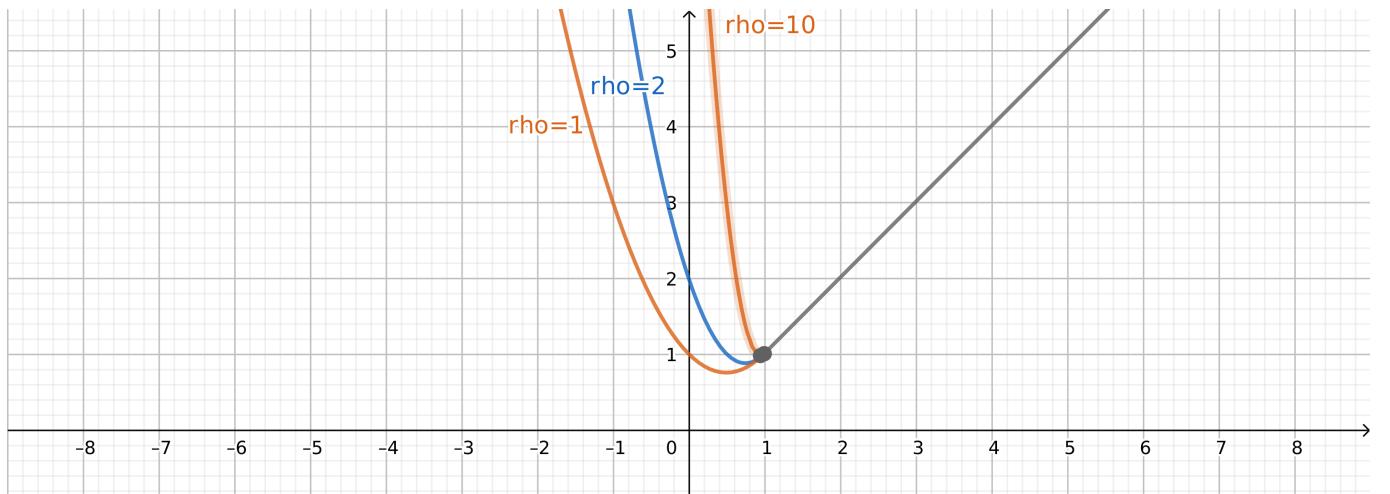
$$\begin{aligned} \text{Min } & x \\ \text{s.a } & x \geq 1 \end{aligned}$$

y la función de penalidad cuadrática $P(x) = \max\{0, 1-x\}^2$.

Las siguientes son las gráficas de la función $P(x, \rho) = f(x) + \rho P(x)$ para los valores de $\rho = 1, 2, 10$



Observar las tres gráficas juntas:



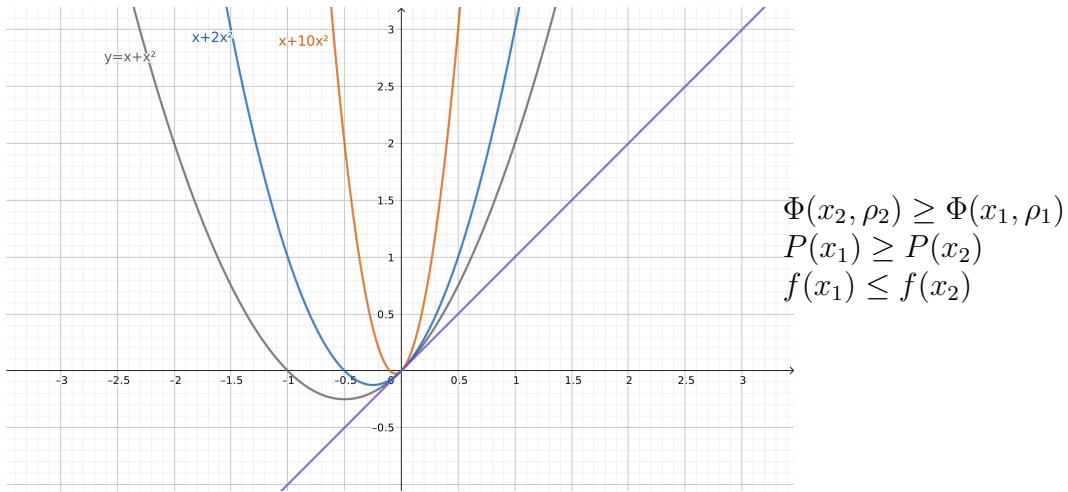
Se puede demostrar que un minimizador global de $\Phi(x, \rho)$ tiende a un minimizador global del problema original. Para probarlo necesitamos estudiar algunas propiedades de la función de penalidad.

En el ejemplo

$$\begin{aligned} \text{Min } & x \\ \text{s.a } & x = 0 \end{aligned}$$

se tiene que $\Phi(x, \rho) = x + \rho x^2$, $\Phi'(x, \rho) = 1 + 2\rho x = 0 \Leftrightarrow x = -\frac{1}{2\rho}$. Entonces:

- Para $\rho = 1$, $x_1 = -1/2$, $\Phi(x_1, \rho_1) = -1/2 + 1 \cdot (-1/2)^2 = -1/4$, $P(x_1) = 1/4$, $f(x_1) = -1/2$
- Para $\rho = 2$, $x_2 = -1/4$, $\Phi(x_2, \rho_2) = -1/4 + 2 \cdot (-1/4)^2 = -1/8$, $P(x_2) = 1/16$, $f(x_2) = -1/4$



Lema 6.1. Definimos $\Phi(x, \rho) = f(x) + \rho P(x)$. Si $\{x_k\}$ es la sucesión generada por el método de penalidad utilizando una sucesión creciente de parámetros diferentes $\rho_k \geq 0$, entonces

1. $\Phi(x_k, \rho_k) \leq \Phi(x_{k+1}, \rho_{k+1})$
2. $P(x_k) \geq P(x_{k+1})$
3. $f(x_k) \leq f(x_{k+1})$.

Demostración. 1. Se tiene:

$$\Phi(x_{k+1}, \rho_{k+1}) = f(x_{k+1}) + \underbrace{\rho_{k+1} P(x_{k+1})}_{\geq \rho_k \geq 0} \geq f(x_{k+1}) + \rho_k P(x_{k+1}) \geq f(x_k) + \rho_k P(x_k) = \Phi(x_k, \rho_k),$$

ya que x_k es el minimizador de $f(x) + \rho_k P(x)$.

2. Por definición de x_k tenemos que

$$f(x_k) + \rho_k P(x_k) \leq f(x_{k+1}) + \rho_k P(x_{k+1}).$$

Por definición de x_{k+1} tenemos que

$$f(x_{k+1}) + \rho_{k+1} P(x_{k+1}) \leq f(x_k) + \rho_{k+1} P(x_k).$$

Sumando,

$$\rho_k P(x_k) + \rho_{k+1} P(x_{k+1}) \leq \rho_k P(x_{k+1}) + \rho_{k+1} P(x_k).$$

Luego,

$$(\rho_{k+1} - \rho_k)P(x_{k+1}) \leq (\rho_{k+1} - \rho_k)P(x_k)$$

y se obtiene el resultado ya que $\rho_{k+1} - \rho_k \geq 0$.

3. Usando 2.

$$f(x_k) + \rho_k P(x_k) \underset{\text{def. de } x_k}{\leq} f(x_{k+1}) + \rho_k P(x_{k+1}) \underset{2.}{\leq} f(x_{k+1}) + \rho_k P(x_k),$$

luego se obtiene 3. \square

Lema 6.2. Si x^* es solución del problema original entonces

$$f(x^*) \geq \Phi(x_k, \rho_k) \geq f(x_k).$$

Demostración.

$$f(x^*) = f(x^*) + \rho_k P(x^*) \geq f(x_k) + \rho_k P(x_k) \geq f(x_k).$$

\square

Teorema 6.1. Sea $\{x_k\}$ una sucesión generada por el método de penalidad utilizando una sucesión creciente de parámetros diferentes $\rho_k \geq 0$. Entonces, todo punto límite de $\{x_k\}$ es solución de Minimizar $f(x)$ sujeto a $x \in S$.

Demostración. Sea x^* punto límite de la sucesión. Entonces existe $K \subset \mathbb{N}$ tal que $x_k \rightarrow_{k \in K} x^*$. Por continuidad sabemos que $f(x_k) \rightarrow_{k \in K} f(x^*)$.

Sea $f^* = \min_{x \in S} f(x)$ el valor óptimo del problema original.

Por los lemas previos, la sucesión $\{\Phi(x_k, \rho_k)\}$ es no decreciente y acotada por el valor f^* . Entonces

$$\lim_{k \in K} \Phi(x_k, \rho_k) \leq f^*.$$

Luego,

$$\lim_{k \in K} \rho_k P(x_k) = \lim_{k \in K} \Phi(x_k, \rho_k) - f(x_k) = \alpha^* - f(x^*)$$

siendo $\alpha^* = \lim_{k \in K} \Phi(x_k, \rho_k) \leq f^*$.

Como $P(x_k) \geq 0$ y $\rho_k \rightarrow \infty$ tenemos que debe ser $\lim_{k \in K} P(x_k) = 0$. Luego, por continuidad $P(x^*) = 0$ es decir, x^* es factible.

Además, tenemos que $f(x_k) \leq f^*$. Entonces, $f(x^*) = \lim_{k \in K} f(x_k) \leq f^*$. Esto implica que $f(x^*) = f^*$. \square

Observaciones:

- En el caso de restricciones de desigualdad, la función de penalidad cuadrática es de la forma

$$P(x) = \max\{0, g(x)\}^2 = \begin{cases} (g(x))^2 & \text{si } g(x) \geq 0 \\ 0 & \text{si } g(x) < 0. \end{cases}$$

Luego, $P(x)$ es una función continua con gradiente:

$$\nabla P(x) = \begin{cases} 2g(x)\nabla g(x) & \text{si } g(x) > 0 \\ 0 & \text{si } g(x) < 0 \\ 0 & \text{si } g(x) = 0, \end{cases}$$

es decir, cuya derivada es continua. Sin embargo, la función de penalidad $P^1(x) = \max\{0, g(x)\}$ es continua pero puede no tener derivada continua cuando $g(x) = 0$.

- Si se tiene el problema

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0, g(x) \leq 0, l \leq x \leq u \quad (84)$$

suele considerarse

$$P(x) = \sum_{i=1}^m (h_i(x))^2 + \sum_{i=1}^p \max\{0, g_i(x)\}^2$$

y el subproblema que se resuelve en cada iteración es:

$$\begin{aligned} &\text{Minimizar} && f(x) + \rho_k(P(x)) \\ &\text{s. a} && l \leq x \leq u. \end{aligned}$$

Luego, se puede usar el método del gradiente proyectado para encontrar una solución del subproblema.

De manera similar, cuando el problema original tiene restricciones lineales de igualdad, éstas, generalmente, no se penalizan.

Consideremos ahora la función de penalidad cuadrática para el problema con restricciones de igualdad

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0. \quad (85)$$

El subproblema que aparece en cada paso del método cuando se utiliza la función de penalidad cuadrática es

$$\text{Minimizar } f(x) + \frac{\rho_k}{2} \sum_{i=1}^m (h_i(x))^2.$$

Por un lado, un punto x_k estacionario de este problema verifica:

$$\nabla_x \Phi(x_k, \rho_k) = \nabla f(x_k) + \rho_k \sum_{i=1}^m h_i(x_k) \nabla h_i(x_k) = 0. \quad (86)$$

Por otro lado, para el problema general (85), buscamos un $\lambda \in \mathbb{R}^m$ tal que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla h_i(x^*) = 0. \quad (87)$$

Por ejemplo, si consideramos el problema

$$\begin{aligned} &\text{Min} && \frac{1}{2}(x^2 + y^2) + 2y \\ &\text{s. a} && y = 0 \end{aligned}$$

se puede demostrar que el punto $(0, 0)$ cumple las condiciones KKT con multiplicador de Lagrange $\lambda = -2$.

En este caso el subproblema del método de penalidad es

$$\text{Min} \frac{1}{2}(x^2 + y^2) + 2y + \frac{\rho}{2}y^2$$

y un punto estacionario de esta función es $x_\rho = (0, \frac{-2}{1+\rho})$ para el cual se tiene que

$$\rho h(x_\rho) = \frac{-2\rho}{1+\rho} \rightarrow_{\rho \rightarrow \infty} -2.$$

El siguiente teorema nos da condiciones necesarias bajo las cuales $\lim_{k \rightarrow \infty} \rho_k h_i(x_k) = \lambda_i$.

Teorema 6.2. Sea $\{x_k\}$ la sucesión generada por el método de penalidad con la función de penalidad cuadrática tal que $\nabla f(x_k) + \sum_{i=1}^m \lambda_k^i \nabla h_i(x_k) = 0$ siendo $\lambda_k^i = \rho_k h_i(x_k)$. Si $x_k \rightarrow x^*$ solución regular entonces $\lambda_k = \rho_k h(x_k)$ converge al multiplicador de Lagrange asociado.

Demostración. De la igualdad (86) tenemos que

$$\nabla f(x_k) + \sum_{i=1}^m \lambda_k^i \nabla h_i(x_k) = 0. \quad (88)$$

Suponemos que la sucesión $\{\lambda_k\}$ no tiene puntos de acumulación. Entonces, tenemos que $\|\lambda_k\| \rightarrow \infty$ y dividiendo por $\|\lambda_k\|$ en (88) obtenemos que

$$\frac{\nabla f(x_k)}{\|\lambda_k\|} + \sum_{i=1}^m \frac{\lambda_k^i}{\|\lambda_k\|} \nabla h_i(x_k) = 0. \quad (89)$$

Ahora, como la sucesión $u_k = \frac{\lambda_k}{\|\lambda_k\|}$ tiene norma 1, tiene una subsucesión convergente y tenemos que existen $K \subset \mathbb{N}$ y u^* tales que $\lim_{k \in K} u_k = u^*$ y $\|u^*\| = 1$. Luego, tomando límite para $k \in K$ en (89) se obtiene:

$$\sum_{i=1}^m u_i^* \nabla h_i(x^*) = 0$$

que contradice la independencia lineal.

Luego, la sucesión $\{\lambda_k\}$ debe tener al menos una subsucesión convergente y por lo tanto un punto de acumulación. Es decir, existen $K \subset \mathbb{N}$ y λ^* tales que $\lim_{k \in K} \lambda_k = \lambda^*$ y tomando límite en (88) se tiene que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0.$$

Esto implica que $\lambda_k = \rho_k h(x_k)$ converge a λ^* multiplicador de Lagrange asociado. \square

Observaciones:

- El multiplicador es único.
- El mismo resultado del Teorema 6.2 se puede demostrar si la condición (86) donde se resuelve el subproblema se reemplaza por una resolución inexacta en el siguiente sentido: x_k es "casi" estacionario, es decir, dada una sucesión $\{\varepsilon_k\}$ decreciente tal que $\varepsilon_k \rightarrow 0$, x_k cumple:

$$\|\nabla f(x_k) + \sum_{i=1}^m \lambda_k^i \nabla h_i(x_k)\| \leq \varepsilon_k. \quad (90)$$

El Teorema 6.2 se puede generalizar para desigualdades usando la función de penalidad cuadrática.

Teorema 6.3. Sea $\{x_k\}$ la sucesión tal que $\nabla f(x_k) + \sum_{i=1}^p \rho_k \max\{0, g_i(x_k)\} \nabla g_i(x_k) = 0$. Si $x_k \rightarrow x^*$ solución regular del problema de Minimizar $f(x)$ sujeto a $g_i(x) \leq 0, i = 1, \dots, p$ entonces $\mu_k = \rho_k \max\{0, g(x_k)\}$ converge al multiplicador de Lagrange asociado.

6.2 Método de Lagrangiano Aumentado

Un defecto importante de los métodos de penalidad es la necesidad de hacer $\rho_k \rightarrow \infty$ para probar convergencia a una solución. Esto puede generar problemas numéricos en la resolución de los subproblemas.

Consideremos el problema con restricciones de igualdad (85). Las condiciones KKT para este problema vienen dadas por:

$$\begin{aligned}\nabla f(x) + Jh(x)^T \lambda &= 0 \\ h(x) &= 0.\end{aligned}$$

Sabemos que, si x^* es un punto KKT y λ es el multiplicador de Lagrange asociado entonces (x^*, λ) es un punto estacionario de la función de Lagrange $l(x, \lambda)$. Lamentablemente, aunque conociéramos de antemano λ , x^* puede no ser un minimizador de $l(x, \lambda)$. Por ejemplo, si consideramos el problema de minimizar x^3 sujeto a $x + 1 = 0$. Tenemos que $x^* = -1$, $\lambda = -3$ y $l''(x^*, \lambda) < 0$.

Sin embargo se puede demostrar el siguiente resultado.

Teorema 6.4. *Si x^* cumple las condiciones suficientes de segundo orden para el problema (85) y λ es el multiplicador de Lagrange asociado entonces existe $\bar{\rho} \geq 0$ tal que la función*

$$\bar{l}(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \frac{\rho}{2} \sum_{i=1}^m (h_i(x))^2$$

tiene un minimizador local en x^ para todo $\rho \geq \bar{\rho}$.*

Lema 6.3. *Si $G = G^T \in \mathbb{R}^{n \times n}$, $z^T G z > 0 \forall z \in \text{Nu}(A) - \{0\}$ entonces existe $\bar{c} \geq 0$ tal que $G + cA^T A$ es definida positiva $\forall c \geq \bar{c}$.*

Demostración. Suponemos que no, entonces, para todo k , $G + kA^T A$ no es definida positiva. Entonces, para cada $k \in \mathbb{N}$ existe $x_k : \|x_k\| = 1$ y

$$x_k^T (G + kA^T A) x_k \leq 0. \quad (91)$$

Como $\|x_k\| = 1$, $\{x_k\}$ tiene una subsucesión convergente. Entonces existen $K \subset \mathbb{N}$ y $\bar{x} \in \mathbb{R}^n$ tales que $\lim_{k \in K} x_k = \bar{x}$. Luego, tomando límite superior en (91) se tiene:

$$\bar{x}^T G \bar{x} + \limsup_{k \in K} k x_k^T A^T A x_k \leq 0.$$

Como $x_k^T A^T A x_k = \|Ax_k\|^2 \geq 0, \forall k$ entonces $\bar{x}^T A^T A \bar{x} = 0$. Luego, $\bar{x} \in \text{Nu}(A) - \{0\}$. Entonces por hipótesis $\bar{x}^T G \bar{x} > 0$. Pero, esto es una contradicción ya que (91) implica que

$$\underbrace{\bar{x}^T G \bar{x}}_{> 0} + \underbrace{\limsup_{k \in K} k x_k^T A^T A x_k}_{\geq 0} \leq 0.$$

□

Demostración del Teorema 6.4. Tenemos que

$$\nabla \bar{l}(x) = \nabla f(x) + (Jh(x))^T \lambda + \rho(Jh(x))^T h(x).$$

Luego, $\nabla \bar{l}(x^*) = 0$, es decir, x^* es un punto estacionario de \bar{l} . Si calculamos $\nabla^2 \bar{l}$ tenemos que

$$\nabla^2 \bar{l}(x) = \nabla^2 f(x) + \sum_{i=1}^m \lambda_i \nabla^2 h_i(x) + \rho((Jh(x))^T Jh(x) + \sum_{i=1}^m h_i(x) \nabla^2 h_i(x)).$$

Entonces

$$\nabla^2 \bar{l}(x^*) = \nabla^2 l(x^*) + \rho(Jh(x^*))^T Jh(x^*).$$

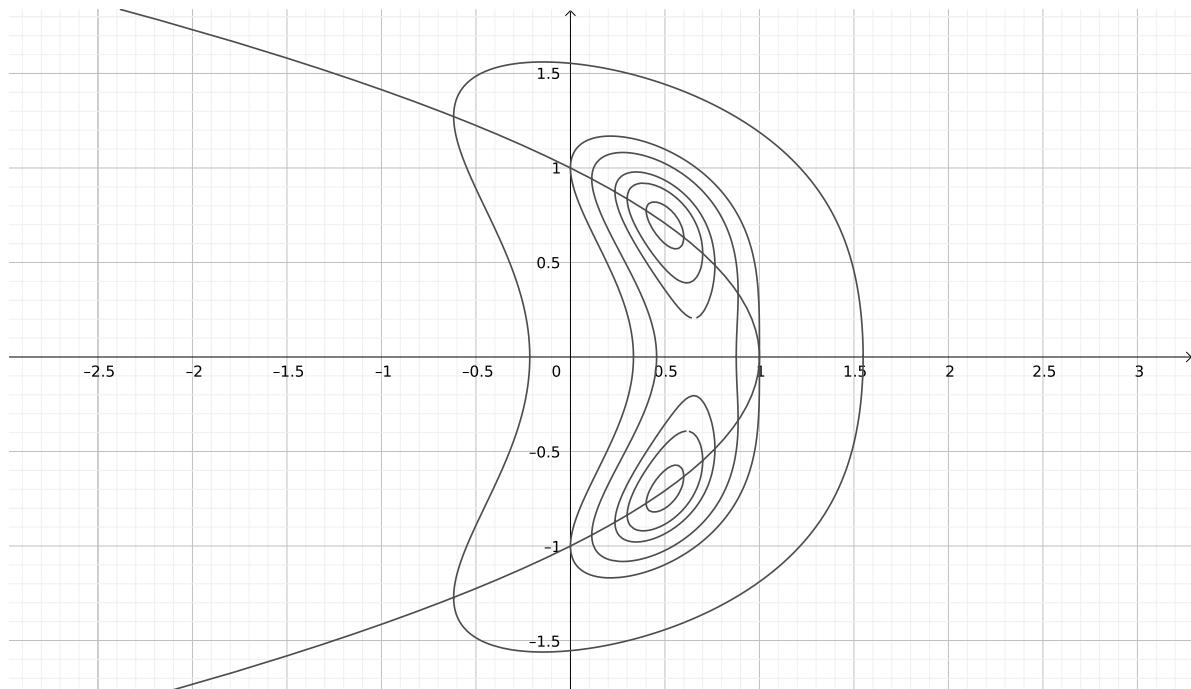
Luego, utilizando el Lema 6.3 tenemos que existe $\bar{\rho} \geq 0$ tal que para todo $\rho \geq \bar{\rho}$, $\nabla^2 \bar{l}(x^*)$ es definida positiva lo que implica que x^* es un minimizador local de \bar{l} . \square

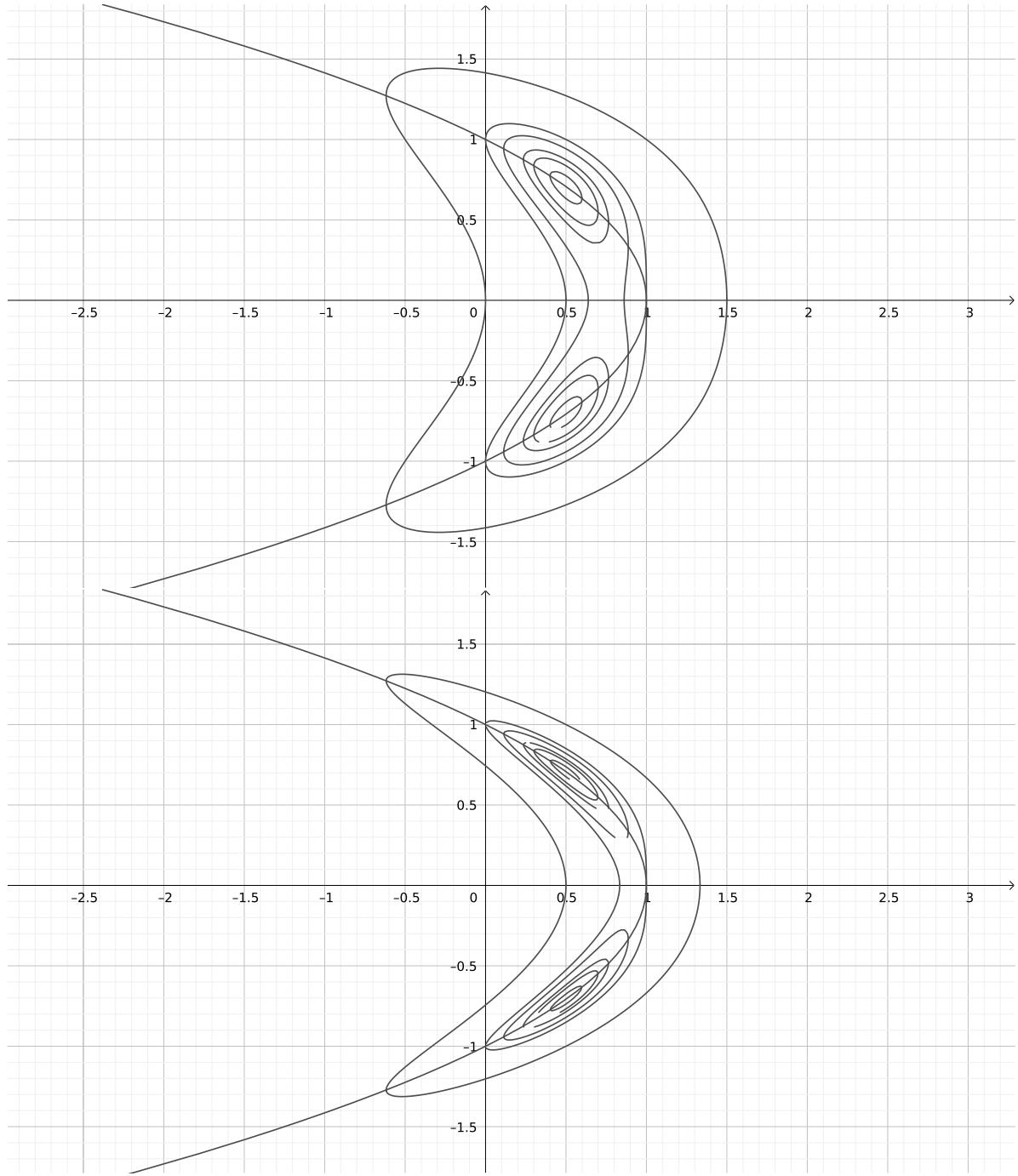
En el ejemplo

$$\begin{aligned} & \text{Min } x^2 + y^2 \\ \text{s. a } & x + y^2 = 1, \end{aligned}$$

el problema tiene dos minimizadores locales: $(\frac{1}{2}, \frac{1}{\sqrt{2}})$, $(\frac{1}{2}, -\frac{1}{\sqrt{2}})$ para los cuales $\lambda = -1$ y se cumplen las condiciones suficientes de segundo orden. Se puede demostrar que para todo $\rho > 0$, los puntos $(\frac{1}{2}, \frac{1}{\sqrt{2}})$, $(\frac{1}{2}, -\frac{1}{\sqrt{2}})$ son minimizadores locales de la función \bar{l} del Teorema 6.4 con $\lambda = -1$, $\forall \rho > 0$.

Las siguientes son las curvas de nivel de \bar{l} para $\rho = 1, 2, 10$.





Luego, si conociéramos de antemano los multiplicadores, bastaría un valor finito de ρ para transformar el problema original en un problema sin restricciones.

El problema con restricciones de igualdad (85) es equivalente, para $\lambda \in \mathbb{R}^m$ fijo, al problema

$$\text{Minimizar } f(x) + \sum_{i=1}^m \lambda_i h_i(x) \text{ sujeto a } h(x) = 0. \quad (92)$$

Si aplicamos el método de penalidad cuadrática a (92), en cada iteración tenemos un subproblema de la forma

$$\text{Minimizar } f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \frac{\rho}{2} \sum_{i=1}^m (h_i(x))^2, \quad (93)$$

que, para cada valor diferente de λ es un subproblema diferente. Si buscamos un punto estacionario

de (93) tenemos que

$$\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla h_i(x) + \rho \sum_{i=1}^m h_i(x) \nabla h_i(x) = 0$$

que equivale a

$$\nabla f(x) + \sum_{i=1}^m (\lambda_i + \rho h_i(x)) \nabla h_i(x) = 0.$$

Comparando con el sistema KKT del problema original, se puede deducir que, para λ fijo, el escalar $\lambda + \rho h(x)$ podría ser una buena estimativa para llegar a calcular el multiplicador asociado. Esto sugiere el siguiente método.

Método de Lagrangiano Aumentado.

Dados x_0 inicial, $\rho_1 > 0$, $\lambda_1 \in \mathbb{R}^m$.

1. Calcular x_k minimizador local de $f(x) + \sum_{i=1}^m [\lambda_i]_k h_i(x) + \frac{\rho_k}{2} \sum_{i=1}^m (h_i(x))^2$.
2. Si $\|h(x_k)\| > 0.1 \|h(x_{k-1})\|$ entonces $\rho_{k+1} = 10\rho_k$.
3. Definir $\lambda_{k+1} = \lambda_k + \rho_k h(x_k)$ y repetir.

La función

$$L(x, \lambda, \rho) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \frac{\rho}{2} \sum_{i=1}^m (h_i(x))^2$$

se denomina *función Lagrangiano Aumentado*.

La solución de cada subproblema del paso 1. es un punto x_k que cumple

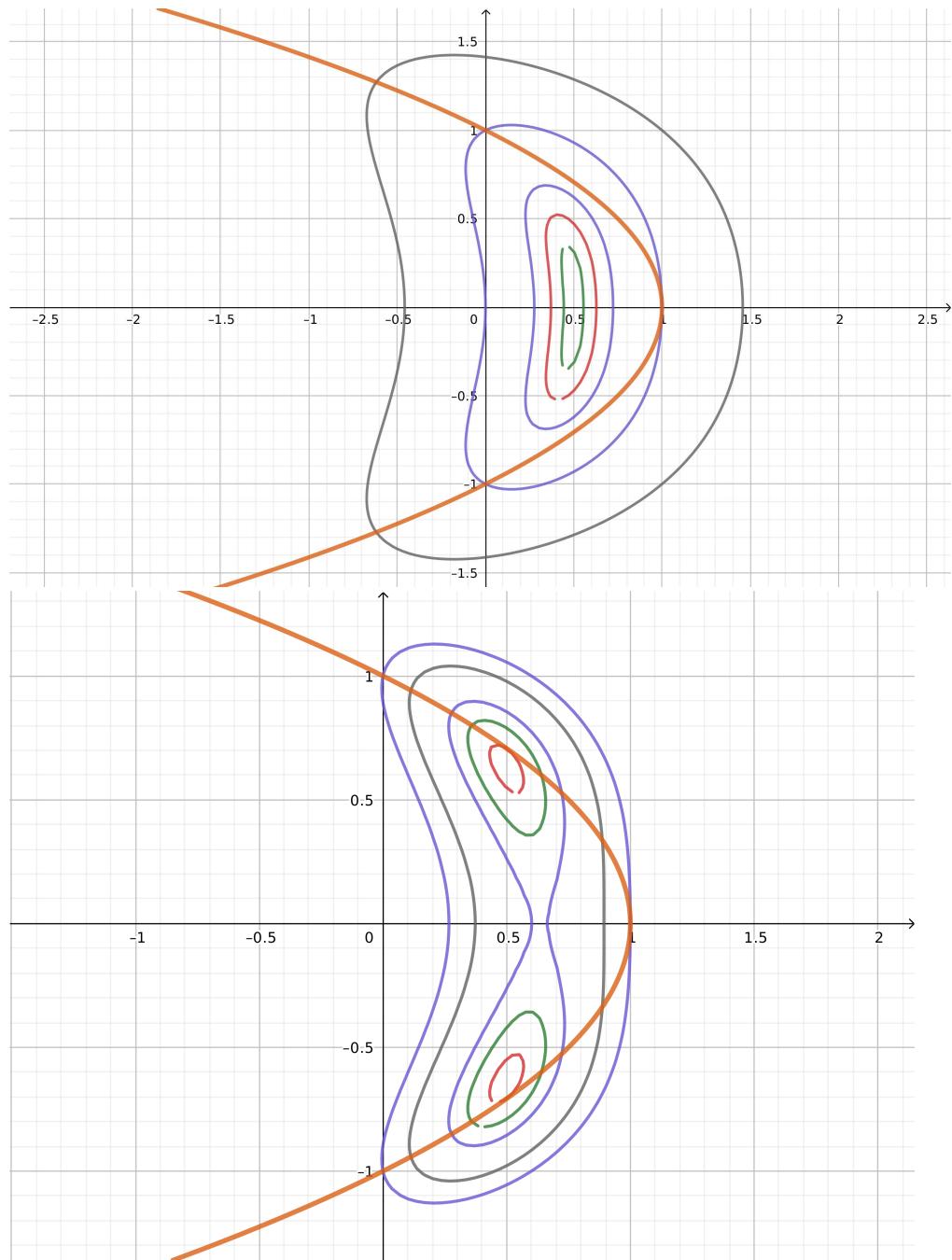
$$\nabla f(x_k) + \sum_{i=1}^m ([\lambda_k]_i + \rho_k h_i(x_k)) \nabla h_i(x_k) = 0$$

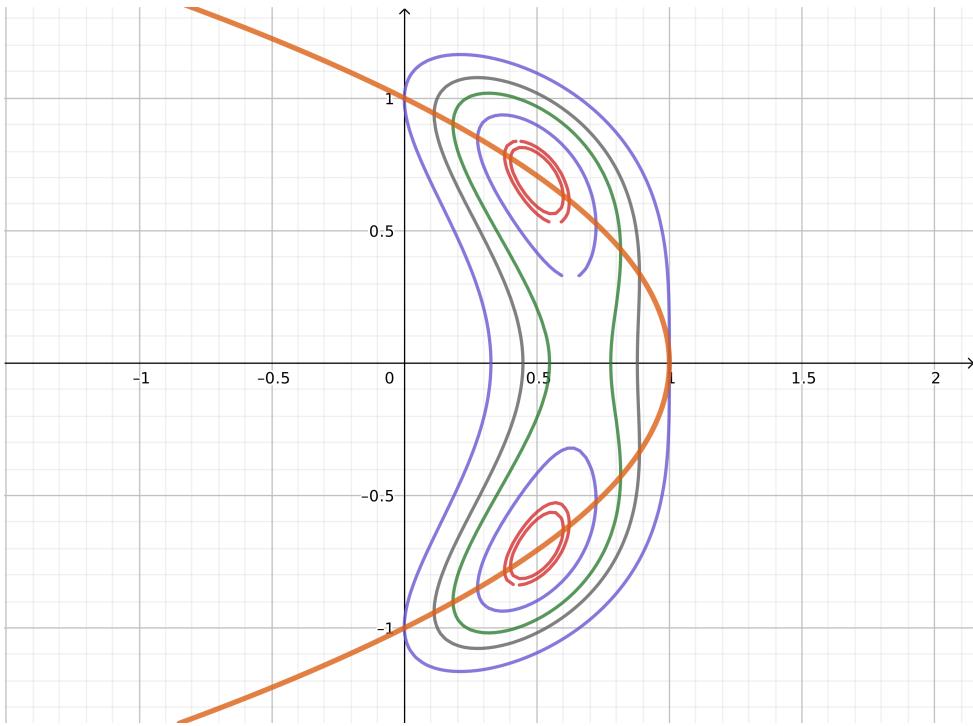
pero puede no ser factible, o puede, no ser “tan factible” como lo era x_{k-1} . Por eso, en el paso 2. comparamos si $\|h(x_k)\|$ decrece comparada con el valor previo $\|h(x_{k-1})\|$.

Para el ejemplo

$$\begin{aligned} \text{Min } & x^2 + y^2 \\ \text{s. a } & x + y^2 = 1 \end{aligned}$$

Las siguientes son las curvas de nivel de la función Lagrangiano aumentado con valor fijo $\rho = 1$ y valores de $\lambda = -0.5, -0.9, -0.99$



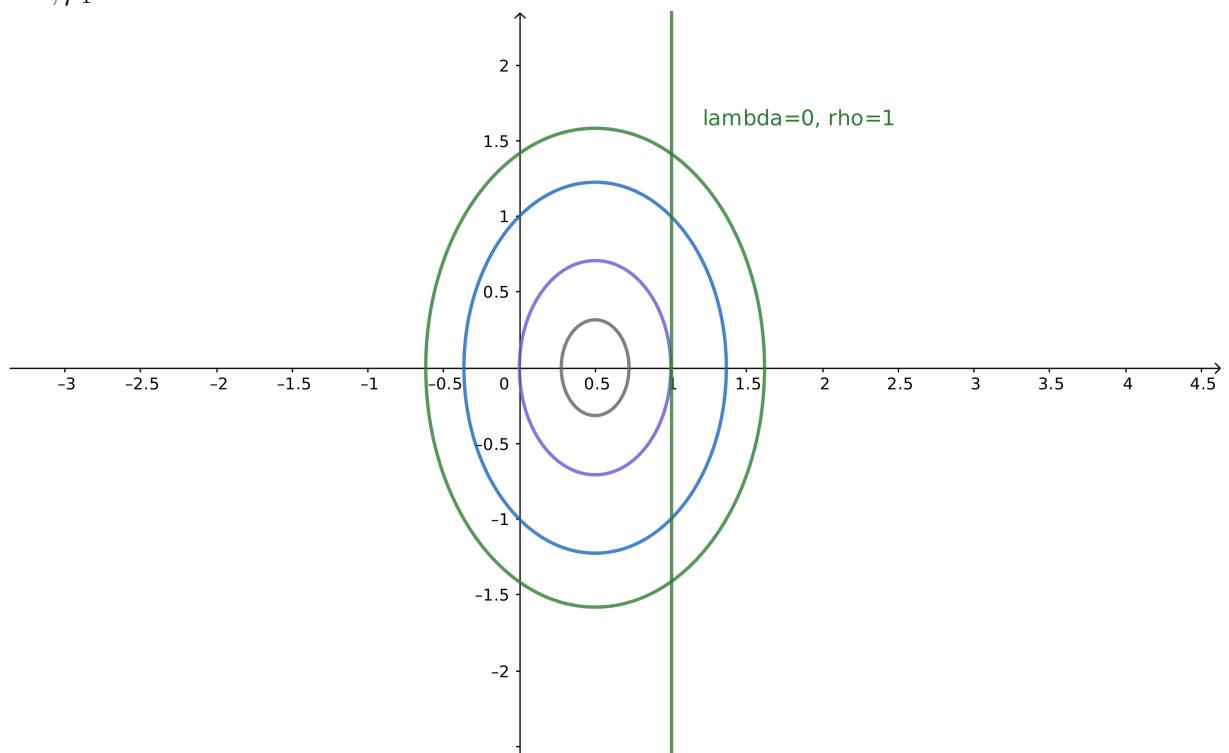


Consideramos el problema

$$\begin{aligned} \text{Min } & 1/2(x^2 + y^2) \\ \text{s. a } & x = 1 \end{aligned}$$

Tenemos que $x^* = (1, 0)$. Veamos la aplicación del método de Lagrangiano Aumentado iniciando con $x_0 = (0, 0)$, $\lambda_1 = 0$, $\rho_1 = 1$.

Las siguientes son las curvas de nivel de la función Lagrangiano aumentado para los valores de $\lambda_1 = 0$, $\rho_1 = 1$



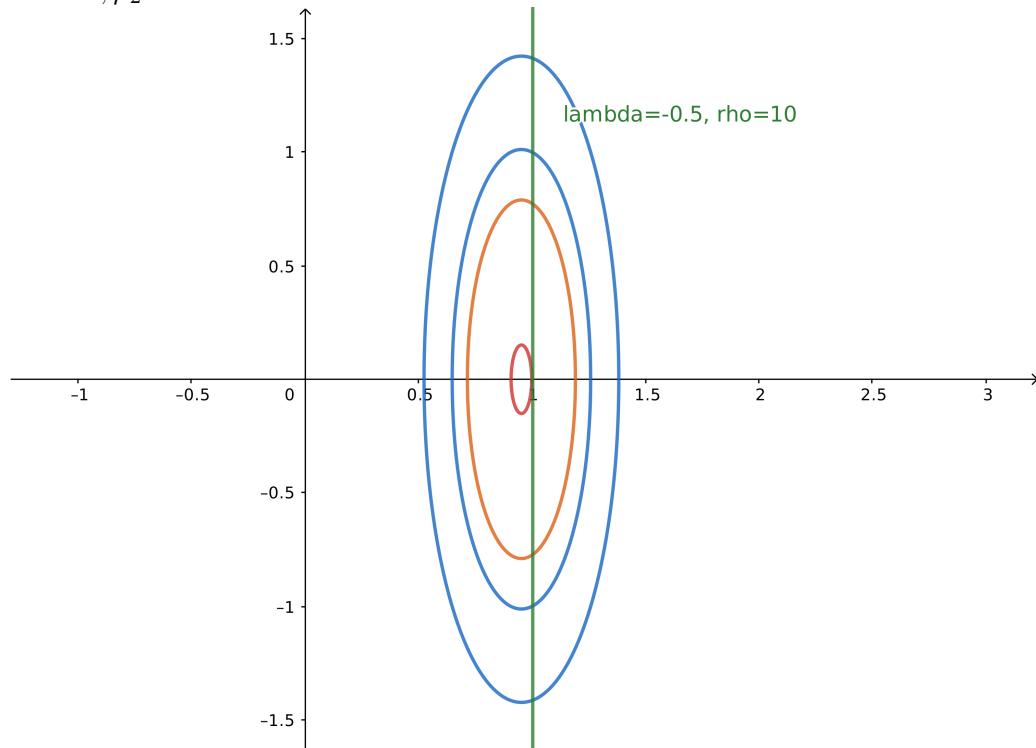
Se puede ver que, la función Lagrangiano aumentado es:

$$L((x, y), \lambda) = \frac{1}{2}(x^2 + y^2) + \lambda(x - 1) + \frac{\rho}{2}(x - 1)^2$$

y se puede ver que un punto estacionario de L es $(\frac{\rho - \lambda}{1 + \rho}, 0)$.

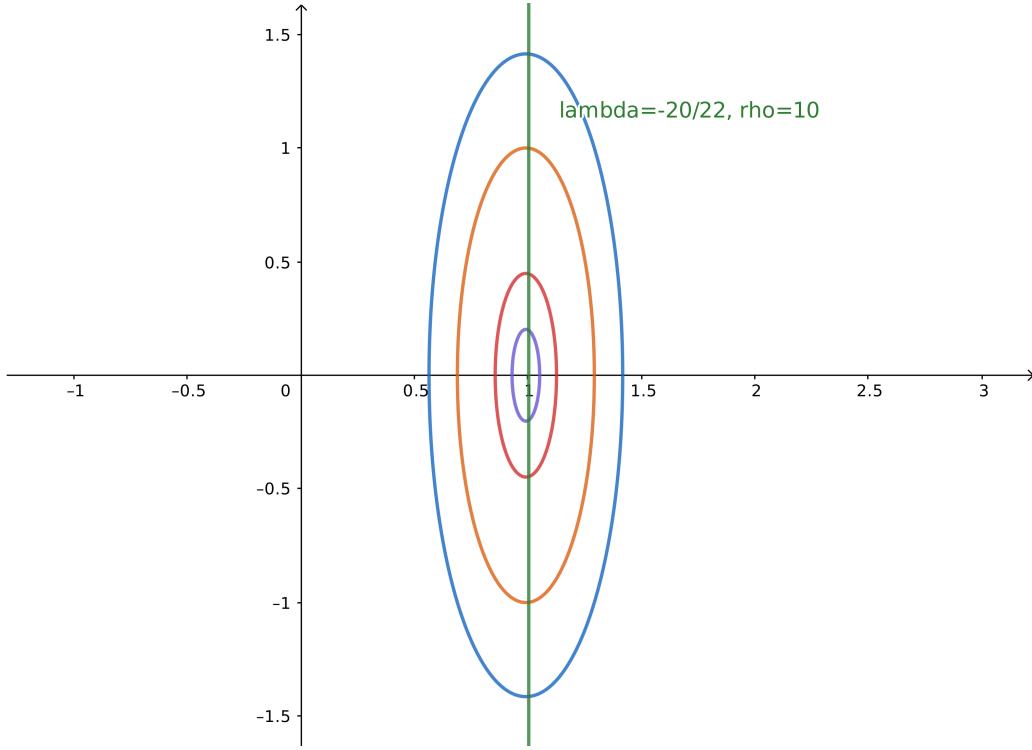
Luego, se tiene que $\|h(x_0)\| = 1, x_1 = (\frac{1}{2}, 0), \|h(x_1)\| = \frac{1}{2} > \frac{1}{10}\|h(x_0)\|$ entonces $\rho_2 = 10, \lambda_2 = \lambda_1 + \rho_1 h(x_1) = -\frac{1}{2}$.

Las siguientes son las curvas de nivel de la función Lagrangiano aumentado para los valores de $\lambda_2 = -0.5, \rho_2 = 10$.



Luego, se tiene que $x_2 = (\frac{21}{22}, 0), \|h(x_2)\| = \frac{1}{22} \leq \frac{1}{10} \frac{1}{2}$ entonces $\rho_3 = \rho_2 = 10$ y $\lambda_3 = \lambda_2 + \rho_2 h(x_2) = -\frac{20}{22}$.

Las siguientes son las curvas de nivel de la función Lagrangiano aumentado para los valores de $\lambda_3 = -20/22, \rho_3 = 10$.



Observación: En este caso se tiene que

$$\lambda_{k+1} = \lambda_k + \rho_k h(x_k) = \lambda_k + \rho_k(x_k - 1) = \lambda_k + \rho_k\left(\frac{\rho_k - \lambda_k}{1 + \rho_k} - 1\right) = \frac{\lambda_k - \rho_k}{1 + \rho_k}.$$

Luego,

$$\|\lambda_{k+1} - \lambda^*\| = \|\lambda_{k+1} + 1\| = \left|\frac{\lambda_k - \rho_k}{1 + \rho_k} + 1\right| = \frac{|\lambda_k + 1|}{1 + \rho_k} \leq \dots \leq \frac{|\lambda_1 + 1|}{(1 + \rho_k)(1 + \rho_{k-1}) \dots (1 + \rho_1)}$$

por lo tanto:

- si $\rho_k \rightarrow \infty$ se tiene que $\lambda_{k+1} \rightarrow \lambda^*$;
- si $\{\rho_k\}$ es una sucesión acotada entonces $\rho_{k+1} = \rho_k, \forall k \geq k_0$ y en ese caso

$$\|\lambda_{k+1} - \lambda^*\| \leq \frac{|\lambda_1 + 1|}{(1 + \rho_{k_0})^{k-k_0}(1 + \rho_{k_0-1}) \dots (1 + \rho_1)}$$

y se tiene que $\lambda_{k+1} \rightarrow \lambda^*$ cuando $k \rightarrow \infty$.

Veamos qué se puede demostrar teóricamente.

Teorema 6.5. *Sea x_k minimizador global de la función $L(x, \lambda_k, \rho_k)$. Suponemos que la sucesión $\{\lambda_k\}$ es acotada y que $\{\rho_k\}$ es una sucesión creciente tal que $\rho_k \rightarrow \infty$. Entonces, todo punto límite de $\{x_k\}$ es minimizador global de f sujeto a $h(x) = 0$.*

Demostración. Sea \bar{x} un punto límite de $\{x_k\}$. Entonces existe $K \subset \mathbb{N}$ tal que $\lim_{k \rightarrow \infty} x_k = \bar{x}$.

$$\text{Sea } f^* = \min_{h(x)=0} f(x) = \min_{h(x)=0} L(x, \lambda_k, \rho_k).$$

Como

$$\begin{aligned} L(x_k, \lambda_k, \rho_k) &\leq L(x, \lambda_k, \rho_k) \\ L(x_k, \lambda_k, \rho_k) &\leq \min_{h(x)=0} L(x, \lambda_k, \rho_k) = f^*. \end{aligned}$$

Es decir,

$$f(x_k) + \sum_{i=1}^m \lambda_i^k h_i(x_k) + \frac{\rho_k}{2} \sum_{i=1}^m (h_i(x_k))^2 \leq f^*. \quad (94)$$

Como $\{\lambda_k\}_{k \in K}$ es acotada tiene una subsucesión convergente. Entonces existen $K_1 \subset K$ y $\bar{\lambda}$ tales que

$$\lim_{k \in K_1} \lambda_k = \bar{\lambda}.$$

Tomando límite superior en (94) para $k \in K_1$ se obtiene que

$$f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i h_i(\bar{x}) + \limsup_{k \in K_1} \frac{\rho_k}{2} \sum_{i=1}^m (h_i(x_k))^2 \leq f^*.$$

Como $\rho_k \rightarrow \infty$ debe ser $\|h(x_k)\| \rightarrow_{k \in K_1} 0$. Entonces $h(\bar{x}) = 0$ y \bar{x} es factible.

Luego, se tiene que $f^* \leq f(\bar{x})$.

Esto implica que,

$$f^* + \limsup_{k \in K_1} \frac{\rho_k}{2} \sum_{i=1}^m (h_i(x_k))^2 \leq f(\bar{x}) + \limsup_{k \in K_1} \frac{\rho_k}{2} \sum_{i=1}^m (h_i(x_k))^2 \leq f^*,$$

luego,

$$\limsup_{k \in K_1} \frac{\rho_k}{2} \sum_{i=1}^m (h_i(x_k))^2 = 0$$

y $f^* = f(\bar{x})$. □

Observación:

- Se asume que $L(x, \lambda_k, \rho_k)$ tiene un minimizador global. Esto puede no suceder. Por ejemplo para el problema de Minimizar $-x^4$ sujeto a $x = 0$ la función Lagrangiano aumentado es $L(x, \lambda_k, \rho_k) = -x^4 + \lambda_k x + \rho_k x^2$ que no es acotada inferiormente y no tiene minimizador global finito.

La tesis del teorema es fuerte porque se basa en una hipótesis fuerte, de hecho, se pide demasiado con buscar un minimizador global del subproblema.

Proposición 6.1. *Sea x_k tal que $\nabla_x L(x_k, \lambda_k, \rho_k) = 0$ y asumimos que $\{\rho_k\}$ es una sucesión creciente de escalares positivos. Si $\{x_k\}_{k \in K}$ es una subsucesión tal que $x_k \rightarrow_{k \in K} x^*$ para x^* regular entonces x^* es un punto KKT del problema original y $\lambda_k + \rho_k h(x_k) \rightarrow_{k \in K} \lambda$ multiplicador de Lagrange asociado.*

Demostración. Como $\nabla_x L(x_k, \lambda_k, \rho_k) = 0$ y $\lambda_{k+1} = \lambda_k + \rho_k h(x_k)$ se tiene que

$$\nabla f(x_k) + \sum_{i=1}^m \lambda_{k+1}^i \nabla h_i(x_k) = 0. \quad (95)$$

Suponemos que la sucesión $\{\lambda_{k+1}\}$ no tiene puntos de acumulación. Entonces, tenemos que $\|\lambda_{k+1}\| \rightarrow \infty$ y dividiendo por $\|\lambda_{k+1}\|$ en (95) obtenemos que

$$\frac{\nabla f(x_k)}{\|\lambda_{k+1}\|} + \sum_{i=1}^m \frac{\lambda_{k+1}^i}{\|\lambda_{k+1}\|} \nabla h_i(x_k) = 0. \quad (96)$$

Ahora, definimos como $u_k = \frac{\lambda_{k+1}}{\|\lambda_{k+1}\|}$. Luego, la sucesión $\{u_k\}$ tiene norma 1 por lo tanto tiene una subsucesión convergente. Es decir, existen $K \subset \mathbb{N}$ y $u^* \in \mathbb{R}^m$ tales que $\lim_{k \in K} u_k = u^*$ y $\|u^*\| = 1$. Luego, tomando límite para $k \in K$ en (96) se obtiene:

$$\sum_{i=1}^m u_i^* \nabla h_i(x^*) = 0$$

que contradice la independencia lineal.

Luego, la sucesión $\{\lambda_{k+1}\}$ debe tener al menos una subsucesión convergente y por lo tanto un punto de acumulación. Es decir, existen $K \subset \mathbb{N}$ y $\lambda^* \in \mathbb{R}^m$ tales que $\lim_{k \in K} \lambda_{k+1} = \lambda^*$ y tomando límite en (95) se tiene que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0$$

Esto implica que x^* es un punto KKT y se tiene que λ_{k+1} converge a λ^* multiplicador de Lagrange asociado. \square

Observaciones: 1. Se puede demostrar un resultado análogo al Teorema anterior si se considera una resolución inexacta del subproblema:

- x_k tal que $\|\nabla_x L(x_k, \lambda_k, \rho_k)\| \leq \varepsilon_k$ siendo $\{\varepsilon_k\}$ una sucesión decreciente no negativa tal que $\varepsilon_k \rightarrow 0$.

2. Se puede demostrar el mismo resultado de convergencia pero considerando condiciones de calidad más débiles en la literatura.

Este método se puede extender para problemas que además tienen restricciones de desigualdad. Consideremos un problema general con restricciones de igualdad y de desigualdad de la forma:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } h_i(x) = 0, \quad i = 1, \dots, m \\ & g_j(x) \leq 0, \quad j = 1, \dots, p. \end{aligned} \quad (97)$$

Mediante la introducción de una variable adicional z_j asociada a cada restricción de desigualdad, se puede convertir el problema (97) en un problema con restricciones de igualdad al que se le puede aplicar el estudio anterior:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeto a } h_i(x) = 0, \quad i = 1, \dots, m, \\ & g_j(x) + z_j^2 = 0, \quad j = 1, \dots, p. \end{aligned} \quad (98)$$

Se puede demostrar lo siguiente: x^* es un minimizador local de (97) si y solo si (x^*, z^*) es minimizador local de (98) donde $z^* = \sqrt{-g_i(x^*)}$.

Basados en esa conversión usamos que el problema (98) tiene restricciones de igualdad y sabemos que la función Lagrangiano aumentado que utiliza la función de penalidad cuadrática se define como:

$$\bar{L}(x, z, \lambda, \mu, \rho) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \frac{\rho}{2} \sum_{i=1}^m (h_i(x))^2 + \sum_{j=1}^p \mu_j [g_j(x) + z_j^2] + \frac{\rho}{2} \sum_{j=1}^p [g_j(x) + z_j^2]^2.$$

Claramente, la función \bar{L} depende de la variable $x \in \mathbb{R}^n$ y de la variable adicional $z \in \mathbb{R}^p$.

En el primer paso del método de Lagrangiano aumentado se resuelve el subproblema de encontrar un minimizador de la función Lagrangiano aumentado:

$$\text{Minimizar } \bar{L}(x, z, \lambda_k, \mu_k, \rho_k). \quad (99)$$

Observando la función \bar{L} vemos que es cuadrática en la variable z_j^2 para cada $j = 1, \dots, p$, luego, la minimización (99) se puede hacer en cada z_j independientemente de x . La minimización con respecto a z_j es equivalente a

$$\text{Minimizar}_{z_j} \mu_j [g_j(x) + z_j^2] + \frac{\rho}{2} [g_j(x) + z_j^2]^2. \quad (100)$$

La función es una cuadrática en la variable z_j^2 y la minimización con respecto a z_j es equivalente a

$$\text{Minimizar}_{u_j \geq 0} \mu_j [g_j(x) + u_j] + \frac{\rho}{2} [g_j(x) + u_j]^2. \quad (101)$$

que es una cuadrática con ramas hacia arriba. El minimizador global irrestringido de la cuadrática en u_j es el escalar \hat{u}_j que anula la derivada:

$$\mu_j + \rho [g_j(x) + \hat{u}_j] = 0$$

para el cual

$$\hat{u}_j = -\frac{\mu_j}{\rho} - g_j(x).$$

Ahora hay dos posibilidades, $\hat{u}_j \geq 0$ y en este caso es la solución de (101); o bien la solución de (101) es $u_j^* = 0$. Es decir, la solución de (101) es

$$u_j^* = \max\left\{0, -\frac{\mu_j}{\rho} - g_j(x)\right\}$$

Así, la solución del problema (100) es:

$$(z_j^*)^2 = \max\left\{0, -\frac{\mu_j}{\rho} - g_j(x)\right\}$$

y entonces

$$z_j^* = \sqrt{\max\left\{0, -\frac{\mu_j}{\rho} - g_j(x)\right\}}. \quad (102)$$

Luego,

$$g_j(x) + (z_j^*)^2 = \max\left\{g_j(x), -\frac{\mu_j}{\rho}\right\},$$

y

$$\mu_j [g_j(x) + (z_j^*)^2] + \frac{\rho}{2} [g_j(x) + (z_j^*)^2]^2 = \frac{1}{2\rho} ([\max\{0, \mu_j + \rho g_j(x)\}]^2 - \mu_j^2).$$

De esta manera, se obtiene la siguiente función Lagrangiano aumentado que ya no depende las variables auxiliares:

$$L(x, \lambda, \mu, \rho) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \frac{\rho}{2} \sum_{i=1}^m (h_i(x))^2 + \frac{1}{2\rho} \sum_{i=1}^p ([\max\{0, \mu_i + \rho g_i(x)\}]^2 - \mu_i^2), \quad (103)$$

para $x \in \mathbb{R}^n, \rho \geq 0, \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$.

Observaciones.

- La actualización del multiplicador asociado a las restricciones de desigualdad es:

$$\mu_{k+1}^i = \mu_k^i + \rho_k(g_i(x_k) + (z_i^k)^2) = \max\{0, \mu_k^i + \rho_k g_i(x_k)\}$$

- Para ver si se adapta el parámetro de penalidad o no se considera la siguiente medida:

$$\max\{\|h(x_k)\|, \|g(x_k) + z_k^2\|\} = \max\{\|h(x_k)\|, \|\max\{g(x_k), -\frac{\mu_k}{\rho_k}\}\|\}$$

Se puede definir entonces el método de Lagrangiano aumentado para el problema (97):

Método de Lagrangiano Aumentado.

Dados x_0 inicial, $\rho_1 > 0$, $\lambda_1 \in \mathbb{R}^m$, $\mu_1 \in \mathbb{R}^p$, $\mu_1 \geq 0$, $0 < \tau < 1$, $\gamma > 1$.

Paso 1 Calcular x_k minimizador local de la función $L(x, \lambda_k, \mu_k, \rho_k)$ dada por (103).

Paso 2. Definir

$$V_j^k = \max \left\{ g_j(x_k), -\frac{\mu_k^j}{\rho_k} \right\}$$

Si $\max\{\|h(x_k)\|, V^k\} \leq \tau \max\{\|h(x_{k-1})\|, V^{k-1}\}$ entonces $\rho_{k+1} = \rho_k$.

Si no, $\rho_{k+1} = \gamma \rho_k$

Paso 3. Definir $\lambda_{k+1} = \lambda_k + \rho_k h(x_k)$, $\mu_{k+1} = \max\{0, \mu_k + \rho_k g(x_k)\}$, hacer $k = k + 1$ y repetir.

Se puede demostrar el siguiente resultado:

Lema 6.4. Sea $\{x_k\}$ la sucesión generada por el Algoritmo de Lagrangiano Aumentado y asumimos que la sucesión $\{\mu_k\}$ es acotada. Sea $K \subset \mathbb{N}$ tal que $\lim_{k \in K} x_k = x^*$. Entonces, para $k \in K$ suficientemente grande,

$$\mu_{k+1}^i = 0 \quad \text{para todo } i = 1, \dots, p \text{ tal que } g_i(x^*) < 0.$$

Demostración. Por definición se tiene $\mu_{k+1} \geq 0$ para todo k .

Supongamos que $g_i(x^*) < 0$ y sea $k_1 \in K$ y $c < 0$ tales que

$$g_i(x_k) < c < 0 \quad \text{para todo } k \in K, k \geq k_1.$$

Consideramos dos casos:

- La sucesión $\{\rho_k\}$ tiende a infinito.
- La sucesión $\{\rho_k\}$ está acotada.

En el primer caso, dado que $\{\mu_k^i\}$ está acotada, existe $k_2 \geq k_1$ tal que, para todo $k \in K, k \geq k_2$,

$$\mu_k^i + \rho_k g_i(x_k) < 0.$$

Esto implica que

$$\mu_{k+1}^i = 0 \quad \text{para todo } k \in K, k \geq k_2.$$

Ahora supongamos que $\{\rho_k\}$ está acotada. Entonces se cumple para todo k suficientemente grande $\max\{\|h(x_k)\|, V^k\} \leq \tau \max\{\|h(x_{k-1})\|, V^{k-1}\}$ y, en consecuencia,

$$\lim_{k \rightarrow \infty} V^k = 0.$$

Por lo tanto,

$$\lim_{k \rightarrow \infty} \max \left\{ g_i(x_k), -\frac{\mu_k^i}{\rho_k} \right\} = 0.$$

Como $g_i(x_k) < c < 0$ para $k \in K$ suficientemente grande, se tiene que debe ser $\lim_{k \in K} \frac{\mu_k^i}{\rho_k} = 0$.

Por lo tanto, como la sucesión $\{\rho_k\}$ está acotada, debe ser $\lim_{k \in K} \mu_k^i = 0$. Entonces, para $k \in K$ suficientemente grande, $\mu_k^i + \rho_k g_i(x_k) < 0$. Por la definición de μ_{k+1} , esto implica que $\mu_{k+1}^i = 0$ para $k \in K$ suficientemente grande, como queríamos demostrar. \square

El resultado anterior establece que vale la condición de complementariedad siempre y cuando la sucesión $\{\mu_k\}$ sea acotada. Esta hipótesis de acotación es usual en métodos de Lagrangiano Aumentado sin embargo se puede evitar considerando una secuencia auxiliar $\{\bar{\mu}_k\}$ en el Algoritmo obtenida a partir de $\{\mu_k\}$ de la siguiente forma:

$$\bar{\mu}_k = \mathcal{P}_{[0, \mu_{max}]}(\mu_k).$$

Si bien utilizando esta sucesión auxiliar que proyecta los estimados de los multiplicadores se pueden obtener los mismos resultados que veremos a continuación, asumimos la acotación de la secuencia $\{\mu_k\}$ por practicidad.

Se tiene entonces el siguiente resultado de convergencia global:

Teorema 6.6. *Sea $\{x_k\}$ una sucesión tal que $\|\nabla_x L(x_k, \lambda_k, \mu_k, \rho_k)\| \leq \varepsilon_k$ siendo L la función Lagrangiano Aumentada definida en (103), $\{\varepsilon_k\}$ una sucesión decreciente de escalares positivos que tiende a cero y $\{\rho_k\}$ una sucesión creciente de escalares positivos. Asumimos que la sucesión $\{\mu_k\}$ es acotada. Entonces, todo punto límite x^* regular de $\{x_k\}$ es un punto KKT de (97) y además $\lambda_{k+1} \rightarrow \lambda^*, \mu_{k+1} \rightarrow \mu^*$ donde $\lambda_{k+1} = \lambda_k + \rho_k h(x_k)$ y $\mu_{k+1} = \max\{0, \mu_k + \rho_k g(x_k)\}$.*

Demostración. La demostración es similar a la de la Proposición 6.1. Sea x^* un punto límite regular de $\{x_k\}$. Como $\|\nabla_x L(x_k, \lambda_k, \mu_k, \rho_k)\| \leq \varepsilon_k$, por el Paso 3. se tiene que existe $v_k \in \mathbb{R}^n$ tal que $\|v_k\| \leq \varepsilon_k$ y

$$\nabla f(x_k) + \sum_{i=1}^m \lambda_{k+1}^i \nabla h_i(x_k) + \sum_{j=1}^p \mu_{k+1}^j \nabla g_j(x_k) = v_k.$$

Por el Lema 6.4 tenemos que

$$\nabla f(x_k) + \sum_{i=1}^m \lambda_{k+1}^i \nabla h_i(x_k) + \sum_{j:g_j(x^*)=0} \mu_{k+1}^j \nabla g_j(x_k) = v_k. \quad (104)$$

Suponemos que la sucesión $\{(\lambda_{k+1}, \mu_{k+1})\}$ no tiene puntos de acumulación. Entonces, tenemos que $\|(\lambda_{k+1}, \mu_{k+1})\| \rightarrow \infty$ y dividiendo por $\|(\lambda_{k+1}, \mu_{k+1})\|$ en (104) obtenemos que

$$\frac{\nabla f(x_k)}{\|(\lambda_{k+1}, \mu_{k+1})\|} + \sum_{i=1}^m \frac{\lambda_{k+1}^i}{\|(\lambda_{k+1}, \mu_{k+1})\|} \nabla h_i(x_k) + \sum_{j:g_j(x^*)=0} \frac{\mu_{k+1}^j}{\|(\lambda_{k+1}, \mu_{k+1})\|} \nabla g_j(x_k) = \frac{v_k}{\|(\lambda_{k+1}, \mu_{k+1})\|}. \quad (105)$$

Pero la sucesión $\left\{ \frac{(\lambda_{k+1}, \mu_{k+1})}{\|(\lambda_{k+1}, \mu_{k+1})\|} \right\}$ tiene norma 1 por lo tanto tiene una subsucesión convergente.

Es decir, existen $K \subset \mathbb{N}$ y un vector $u^* = (\lambda^*, \mu^*) \in \mathbb{R}^{m+p}$ tales que, $\lim_{k \in K} \frac{(\lambda_{k+1}, \mu_{k+1})}{\|(\lambda_{k+1}, \mu_{k+1})\|} = u^*$,

$\|u^*\| = 1$ y $\mu^* \geq 0$ (ya que $\mu_{k+1} \geq 0$). Luego, tomando límite para $k \in K$ en (105) se obtiene:

$$\sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j:g_j(x^*)=0} \mu_j^* \nabla g_j(x^*) = 0$$

que contradice la independencia lineal.

Luego, la sucesión $\{(\lambda_{k+1}, \mu_{k+1})\}$ debe tener al menos una subsucesión convergente y por lo tanto un punto de acumulación. Es decir, existen $K \subset \mathbb{N}$ y $(\lambda^*, \mu^*) \in \mathbb{R}^{m+p}$ tales $\lim_{k \in K} (\lambda_{k+1}, \mu_{k+1}) = (\lambda^*, \mu^*)$ con $\mu^* \geq 0$ (ya que $\mu_{k+1} \geq 0$). Luego, tomando límite en (104) se tiene que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j:g_j(x^*)=0} \mu_j^* \nabla g_j(x^*) = 0$$

y se obtiene el resultado deseado. \square

Observación: La demostración anterior se puede utilizar para probar la convergencia del método de Lagrangiano Aumentado considerando la condición de calidad de Mangasarian-Fromovitz (Definición 4.18) en lugar de regularidad (independencia lineal de los gradientes en el punto límite).

La convergencia se puede demostrar también considerando la condición de calidad de rango constante (Definición 4.19) aunque para ello es necesario usar un lema clásico en optimización que es el lema de Caratheodory:

Lema 6.5. [15] *Lema de Caratheodory. Si*

$$x = \sum_{i=1}^m \alpha_i v_i$$

con $x_i \in \mathbb{R}^n$ y $\alpha_i \neq 0$ para todo i , entonces existen $J \subset \{1, \dots, m\}$ y escalares $\bar{\alpha}_j$ para cada $j \in J$ tales que $x = \sum_{j \in J} \bar{\alpha}_j v_j$, $\bar{\alpha}_j > 0$ para todo $j \in J$ y el conjunto de vectores $\{v_j\}_{j \in J}$ es linealmente independiente.

Demostración. Asumimos que $\{v_i\}_{i=1}^m$ es linealmente dependiente, de lo contrario el resultado trivialmente. Entonces, existe $\beta \in \mathbb{R}^m$, $\beta \neq 0$, tal que

$$\sum_{i=1}^m \beta_i v_i = 0.$$

Por lo tanto, podemos escribir

$$x = \sum_{i=1}^m (\alpha_i - \gamma \beta_i) v_i,$$

para todo $\gamma \in \mathbb{R}$. Sea $\bar{\gamma}$ el menor valor de γ que hace que uno de los coeficientes $\alpha_i - \gamma \beta_i$ sea nulo. Entonces, para el valor de $\bar{\gamma}$ podemos expresar x como una combinación lineal de un subconjunto de vectores estrictamente menor que el conjunto original. Este proceso puede repetirse hasta que todos los vectores en la combinación lineal sean linealmente independientes. \square

El Lema anterior se puede extender a la siguiente versión:

Lema 6.6. Sean u_1, \dots, u_m , v_1, \dots, v_p vectores en \mathbb{R}^n . Sean $x \in \mathbb{R}^n$, $\lambda_1, \dots, \lambda_m$, μ_1, \dots, μ_p escalares tales que $\mu_j \geq 0$ para todo $j = 1, \dots, p$ y

$$x = \sum_{i=1}^m \lambda_i u_i + \sum_{j=1}^p \mu_j v_j.$$

Entonces, existen subconjuntos $I \subset \{1, \dots, m\}$, $J \subset \{1, \dots, p\}$ y escalares $\{\lambda_i^0\}_{i \in I}$, $\{\mu_j^0\}_{j \in J}$, con $\mu_j^0 \geq 0$ para todo $j \in J$, tales que

$$x = \sum_{i \in I} \lambda_i^0 u_i + \sum_{j \in J} \mu_j^0 v_j,$$

y los vectores $\{u_i\}_{i \in I} \cup \{v_j\}_{j \in J}$ son linealmente independientes.

Entonces, se puede demostrar el siguiente teorema:

Teorema 6.7. Sea $\{x_k\}$ una sucesión tal que tal que $\|\nabla_x L(x_k, \lambda_k, \mu_k, \rho_k)\| \leq \varepsilon_k$ siendo L la función Lagrangiano Aumentada definida en (103) siendo $\{\varepsilon_k\}$ una sucesión decreciente de escalares positivos que tiende a cero y $\{\rho_k\}$ una sucesión creciente de escalares positivos. Asumimos que la sucesión $\{\mu_k\}$ es acotada. Entonces, todo punto límite x^* de $\{x_k\}$ que cumple la condición de rango constante (Definición 4.19) es un punto KKT de (97).

Demostración. Sea x^* un punto límite de $\{x_k\}$. Como $\|\nabla_x L(x_k, \lambda_k, \mu_k, \rho_k)\| \leq \varepsilon_k$, por el Paso 3. se tiene que existe $v_k \in \mathbb{R}^n$ tal que $\|v_k\| \leq \varepsilon_k$ y

$$\nabla f(x_k) + \sum_{i=1}^m \lambda_{k+1}^i \nabla h_i(x_k) + \sum_{j=1}^p \mu_{k+1}^j \nabla g_j(x_k) = v_k.$$

Luego, por la factibilidad de x^* y el Lema 6.4 tenemos que

$$\nabla f(x_k) + \sum_{i=1}^m \lambda_{k+1}^i \nabla h_i(x_k) + \sum_{j: g_j(x^*)=0} \mu_{k+1}^j \nabla g_j(x_k) = v_k. \quad (106)$$

Así, por el Lema 6.6 sabemos que, para cada $k \in \mathbb{N}$, existen $I_k \subset \{1, \dots, m\}$, $J_k \subset I(x^*)$ y sucesiones de vectores $\{\alpha_k\}$, $\{\beta_k\}$ tales que $\beta_j^k \geq 0$ para todo k ,

$$\nabla f(x_k) + \sum_{i \in I_k} \alpha_i^k \nabla h_i(x_k) + \sum_{j \in J_k} \beta_j^k \nabla g_j(x_k) = v_k,$$

y los vectores

$$\{\nabla h_i(x_k)\}_{i \in I_k} \cup \{\nabla g_j(x_k)\}_{j \in J_k}$$

son linealmente independientes.

Como el número posible de subconjuntos I_k , J_k es finito, existe $K \subset \mathbb{N}$ tal que, para todo $k \in K$, $I_k = I \subset \{1, \dots, m\}$, $J_k = J \subset I(x^*)$,

$$\nabla f(x_k) + \sum_{i \in I} \alpha_i^k \nabla h_i(x_k) + \sum_{j \in J} \beta_j^k \nabla g_j(x_k) = v_k, \quad (107)$$

y los vectores

$$\{\nabla h_i(x_k)\}_{i \in I} \cup \{\nabla g_j(x_k)\}_{j \in J}$$

son linealmente independientes.

Asumimos que las sucesiones $\{\alpha_k\}$, $\{\beta_k\}$ no son acotadas. Definimos, para todo k ,

$$M_k = \max\{|\alpha_i^k|, \beta_j^k : i \in I, j \in J\}.$$

Tenemos que, $\lim_{k \rightarrow \infty} M_k = \infty$. Por lo tanto, por (107) tenemos que

$$\frac{\nabla f(x_k)}{M_k} + \sum_{i \in I} \frac{\alpha_i^k}{M_k} \nabla h_i(x_k) + \sum_{j \in J} \frac{\beta_j^k}{M_k} \nabla g_j(x_k) = \frac{v_k}{M_k},$$

y entonces,

$$\lim_{k \rightarrow \infty} \sum_{i \in I} \frac{\alpha_i^k}{M_k} \nabla h_i(x_k) + \sum_{j \in J} \frac{\beta_j^k}{M_k} \nabla g_j(x_k) = 0.$$

Como el número de restricciones es finito ($\leq m + p$), existe un índice que realiza el máximo M_k infinitas veces. Es decir, existe un subconjunto $K_1 \subset K$ tal que las sucesiones

$$\left\{ \frac{\alpha_i^k}{M_k} \right\}_{k \in K_1}, \quad i \in I, \quad \text{y} \quad \left\{ \frac{\beta_j^k}{M_k} \right\}_{k \in K_1}, \quad j \in J$$

convergen a un conjunto de coeficientes no todos nulos, pues al menos uno de ellos tiene módulo siempre igual a 1. Por lo tanto, tomando límite para $k \in K_1$, existen $\tilde{\alpha} \in \mathbb{R}^m$, $\tilde{\beta} \in \mathbb{R}^p$, $\tilde{\beta} \geq 0$ con componentes no todas nulas tales que

$$\sum_{i \in I} \tilde{\alpha}_i \nabla h_i(x^*) + \sum_{j \in J} \tilde{\beta}_j \nabla g_j(x^*) = 0.$$

Luego, los gradientes $\{\nabla h_i(x^*)\}_{i \in I} \cup \{\nabla g_j(x^*)\}_{j \in J}$ son linealmente dependientes, y existen puntos en un entorno de x^* donde estos gradientes son linealmente independientes, lo que contradice la hipótesis de rango constante.

Luego, las sucesiones $\{\alpha^k\}$, $\{\beta^k\}$ son acotadas y podemos asumir que existe $K_1 \subset K$, $\alpha \in \mathbb{R}^m$, $\beta \in \mathbb{R}^p$ tales que

$$\lim_{k \in K_1} \alpha^k = \alpha, \quad \text{y} \quad \lim_{k \in K_1} \beta^k = \beta \geq 0.$$

Luego, tomando límite en (107) obtenemos que

$$\nabla f(x^*) + \sum_{i \in I} \alpha_i \nabla h_i(x^*) + \sum_{j \in J} \beta_j \nabla g_j(x^*) = 0,$$

y el teorema está demostrado. □

Observaciones.

1. Existen condiciones de calidad aún más débiles que rango constante y Mangasarian Fromovitz para las cuales se puede demostrar la convergencia del algoritmo de Lagrangiano Aumentado.
2. Se puede considerar un parámetro diferente por cada restricción y adaptar el parámetro o no dependiendo de cómo cambia la restricción asociada. En este caso se pueden obtener los mismos resultados de convergencia.

3. Si se tiene el problema

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0, g(x) \leq 0, l \leq x \leq u \quad (108)$$

se penalizan solo las restricciones $h(x) = 0$ y $g(x) \leq 0$ y en cada paso se resuelve el subproblema

$$\begin{aligned} \text{Minimizar} & \quad L(x, \lambda_k, \mu_k, \rho_k) \\ \text{s. a} & \quad l \leq x \leq u. \end{aligned}$$

De manera similar, cuando el problema original tiene restricciones lineales de igualdad, éstas, generalmente, no se penalizan.

- 4. Bajo ciertas hipótesis se puede demostrar que el parámetro de penalidad no crece indefinidamente: existen $\bar{\rho} > 0$ y k_0 tales que $\forall k \geq k_0, \rho_{k+1} = \rho_k$.
- 5. Existen otras variantes del método en las cuales el multiplicador de Lagrange no se actualiza en todas las iteraciones.
- 6. La acotación de la sucesión de estimados de los multiplicadores de Lagrange se ha estudiado en los últimos años y puede evitarse, bajo ciertas condiciones, en las hipótesis de los teoremas de convergencia global.

6.3 Método de programación cuadrática sucesiva

El método de programación cuadrática sucesiva (SQP, por sus siglas en inglés) es una técnica popular y útil para resolver problemas con restricciones. La idea básica es obtener una solución del problema original resolviendo una sucesión de problemas cuadráticos (función objetivo cuadrática y restricciones lineales), ver [4].

Para dar una idea general de estos métodos consideramos el problema con restricciones de igualdad:

$$\text{Minimizar } f(x) \text{ sujeto a } c(x) = 0. \quad (109)$$

El método SQP es iterativo y la idea básica es, en un iterado x_k considerar un modelo cuadrático que aproxima a (109) en el punto y utilizar la solución de ese subproblema para construir una aproximación mejor x_{k+1} . El proceso genera una sucesión de aproximaciones que, se espera, sea convergente a una solución del problema original, bajo ciertas hipótesis.

Sabemos que la función de Lagrange asociada al problema (109) es $l(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i c_i(x)$.

En una iteración k , teniendo el iterado x_k y un estimado del multiplicador de Lagrange λ_k un algoritmo básico SQP consiste en resolver en cada iteración un problema cuadrático con restricciones lineales de la forma:

$$\begin{aligned} & \text{Minimizar } l(x_k, \lambda_k) + \nabla_x l(x_k, \lambda_k)^T d + \frac{1}{2} d^T B(x_k, \lambda_k) d \\ & \text{s. a } c_i(x_k) + \nabla c_i(x_k)^T d = 0, \forall i = 1, \dots, m. \end{aligned} \quad (110)$$

Cuando $B(x_k, \lambda_k) = \nabla_x^2 l(x_k, \lambda_k)$ se minimiza la aproximación cuadrática de $l(x, \lambda_k)$ alrededor de x_k sujeto a la aproximación lineal de $c(x) = 0$ alrededor de x_k .

Bajo ciertas hipótesis se puede demostrar que el subproblema (110) tiene solución única (d_k, ξ_k) siendo ξ_k el multiplicador de Lagrange asociado. En este caso se define $x_{k+1} = x_k + d_k$ y se calcula un nuevo estimado del multiplicador de Lagrange λ_{k+1} para repetir el proceso con un nuevo modelo cuadrático. Se espera obtener una sucesión $\{(x_k, \lambda_k)\}$ que tenga algún punto de acumulación (x^*, λ^*) que cumpla las condiciones KKT para (109):

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla c_i(x^*) &= 0 \\ c(x^*) &= 0. \end{aligned}$$

Definimos

$$\nabla l(x, \lambda) = \begin{pmatrix} \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla c_i(x) \\ c(x) \end{pmatrix} \quad (111)$$

Algoritmo básico SQP. Dados $(x_0, \lambda_0) \in \mathbb{R}^n \times \mathbb{R}^m$, mientras $\nabla_x l(x_k, \lambda_k) \neq 0$:

Paso 1. Resolver el subproblema (110) obteniendo una solución (d_k, ξ_k) , llamada solución primal-dual.

Paso 2. Hacer $x_{k+1} = x_k + d_k$.

Paso 3. Definir $\lambda_{k+1} = \lambda_k + \xi_k$.

Paso 4. Hacer $k = k + 1$ e ir a 1.

A simple vista hay que tener en cuenta algunas consideraciones:

- El subproblema (110) tiene que tener puntos factibles. (es de esperar que cerca de la solución esto suceda)
- Los subproblemas deben tener una solución óptima. (esto se puede asegurar mediante una elección apropiada de las matrices $B(x_k, \lambda_k)$).
- El algoritmo debe ser convergente a un punto deseable. En este caso, hay dos aspectos de la convergencia que se debe considerar considerar: la convergencia local y la convergencia global. La convergencia local se refiere al comportamiento de la sucesión que genera el algoritmo cuando el punto inicial está cerca del óptimo. Los resultados de convergencia global describen el comportamiento de la sucesión cuando se parte de cualquier punto inicial. Este análisis se realiza típicamente por medio de una función de mérito Φ , es decir, una función cuya reducción implica avanzar hacia una solución.

En el paso 1 del Algoritmo básico, pensamos en una solución que cumpla las condiciones KKT del subproblema, es decir, una solución (d_k, ξ_k) del sistema lineal:

$$\begin{aligned} B(x_k, \lambda_k)d + \nabla_x l(x_k, \lambda_k) + \sum_{i=1}^m \xi_i \nabla c_i(x_k) &= 0 \\ c_i(x_k) + \nabla c_i(x_k)^T d &= 0 \end{aligned} \quad (112)$$

que equivale a

$$\begin{aligned} B(x_k, \lambda_k)d + Jc(x_k)^T \xi &= -\nabla_x l(x_k, \lambda_k) \\ Jc(x_k)d &= -c(x_k). \end{aligned} \quad (113)$$

Luego

$$\begin{pmatrix} B(x_k, \lambda_k) & Jc(x_k)^T \\ Jc(x_k) & 0 \end{pmatrix} \begin{pmatrix} d \\ \xi \end{pmatrix} = \begin{pmatrix} -\nabla_x l(x_k, \lambda_k) \\ -c(x_k) \end{pmatrix}. \quad (114)$$

Observar que, cuando $B(x_k, \lambda_k) = \nabla_x^2 l(x_k, \lambda_k)$, el punto (d_k, ξ_k) definido por el algoritmo es la dirección de Newton para la función $(x, \lambda) \mapsto l(x, \lambda)$ alrededor de (x_k, λ_k) . La matriz Jacobiana, respecto de las dos variables de $l(x, \lambda)$ es la matriz (111) y tiene matriz Hessiana:

$$\begin{pmatrix} \nabla^2 f(x) + \sum_{i=1}^m \lambda_i \nabla^2 c_i(x) & Jc(x)^T \\ Jc(x) & 0 \end{pmatrix}. \quad (115)$$

Se puede demostrar que cuando (x_k, λ_k) está en un entorno de un punto estacionario que cumple las condiciones suficientes de segundo orden, el subproblema está bien definido y se genera una sucesión que converge cuadráticamente a la solución (convergencia local).

Hipótesis H. El par (x^*, λ^*) es una solución de (109) que cumple que $\nabla^2 f$ y $\nabla^2 c_i$ son Lipschitz continuos en un entorno de x^* . Además, x^* es regular y $\nabla_x^2 l(x^*, \lambda^*)$ es definido positivo en el espacio tangente $T^{lin}(x^*) - \{0\}$.

Lema 6.7. Sea (x^*, λ^*) una solución de (109) y suponemos que vale la hipótesis (H). Entonces existe un entorno V_1 de (x^*, λ^*) tal que si $(x_k, \lambda_k) \in V_1$ entonces el sistema (114) tiene única solución (d_k, ξ_k) .

Demostración. Por hipótesis se tiene que la matriz

$$\begin{pmatrix} \nabla_x^2 l(x^*, \lambda^*) & Jc(x^*)^T \\ Jc(x^*) & 0 \end{pmatrix}$$

es no singular y por continuidad existe un entorno V_1 de (x^*, λ^*) tal que $\forall (x, \xi) \in V_1$

$$\begin{pmatrix} \nabla_x^2 l(x, \xi) & Jc(x)^T \\ Jc(x) & 0 \end{pmatrix}$$

es no singular.

Por lo tanto, si $(x_k, \lambda_k) \in V_1$ la matriz

$$\begin{pmatrix} \nabla_x^2 l(x_k, \lambda_k) & Jc(x_k)^T \\ Jc(x_k) & 0 \end{pmatrix}$$

es no singular y el sistema (114) tiene única solución (d_k, ξ_k) como queríamos demostrar. \square

Luego, bajo las hipótesis del Lema anterior, valen las condiciones suficientes de segundo orden para el subproblema (110) y d_k es un minimizador del subproblema.

Teorema 6.8. Sea (x^*, λ^*) una solución de (109) y suponemos que vale la hipótesis (H). Entonces existe un entorno V de (x^*, λ^*) tal que si $(x_0, \lambda_0) \in V$ el algoritmo está bien definido y genera una sucesión $\{(x_k, \lambda_k)\}$ que converge a (x^*, λ^*) .

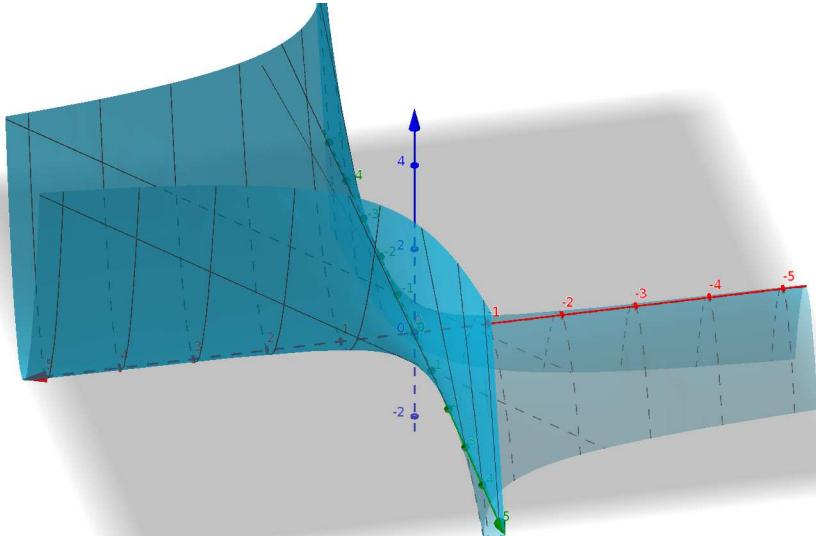
Demostración. El punto (d_k, ξ_k) definido por el algoritmo es la dirección de Newton para la función $(x, \lambda) \mapsto l(x, \lambda)$ alrededor de (x_k, λ_k) . Luego, si $(x_k, \lambda_k) \in V_1$ dado en el Lema 6.7 entonces el paso de Newton está bien definido y la convergencia se obtiene a partir de la convergencia del método de Newton puro y la misma es cuadrática. \square .

Observaciones.

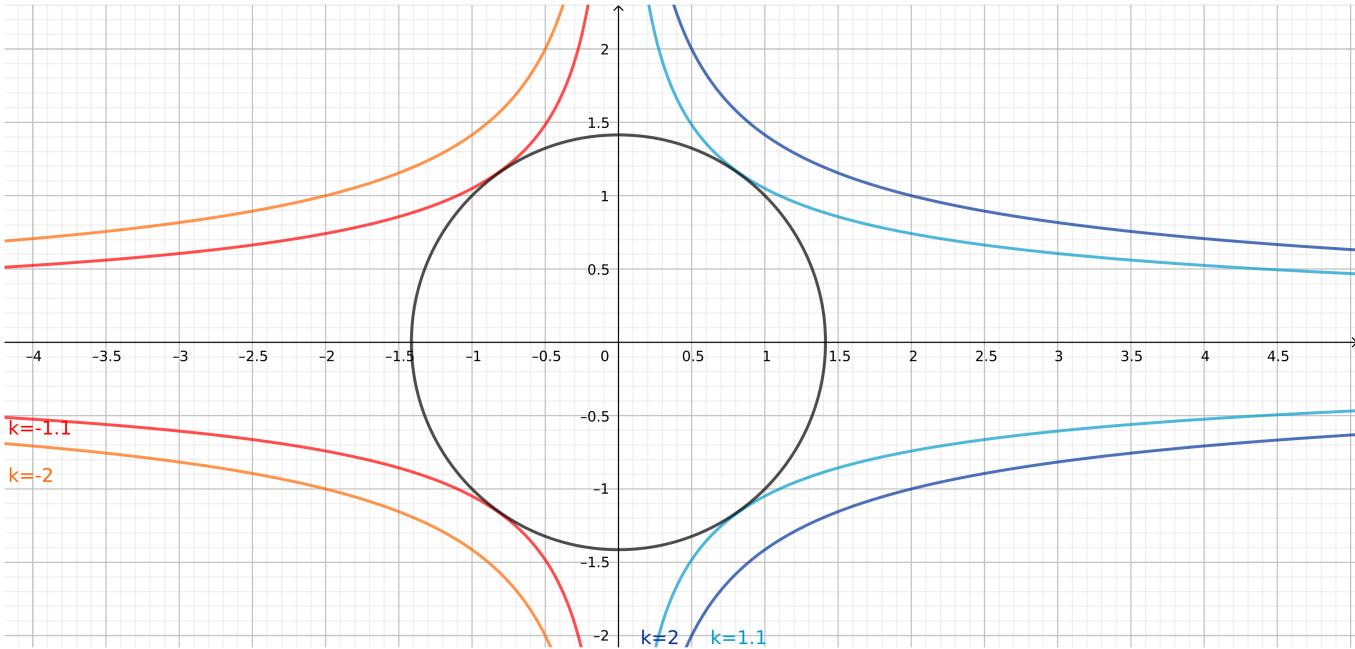
1. El Teorema evidencia la estrecha relación entre el método SQP y el método de Newton (aplicado a la función $(x, \lambda) \mapsto l(x, \lambda)$). Cuando se utiliza $B(x_k, \lambda_k) = \nabla_x^2 l(x_k, \lambda_k)$ el método SQP se interpreta como el método de Newton para minimizar la función de Lagrange asociada al problema original.
2. Si bien se puede demostrar la convergencia cuadrática de la sucesión $\{(x_k, \lambda_k)\}$, no implica directamente la convergencia cuadrática de $\{x_k\}$. Se puede modificar el algoritmo para transformarlo en un algoritmo puramente primal y obtener la convergencia cuadrática de $\{x_k\}$.

Ejemplo: Consideramos el problema

$$\begin{aligned} \text{Min } & xy^2 \\ \text{s. a } & x^2 + y^2 = 2 \end{aligned}$$



El problema tiene 2 minimizadores locales: $(-\sqrt{\frac{2}{3}}, -\sqrt{\frac{4}{3}})$, $(-\sqrt{\frac{2}{3}}, \sqrt{\frac{4}{3}})$ y en esos puntos se tiene que el multiplicador de Lagrange es $\lambda = \sqrt{\frac{2}{3}}$.



Iniciando con $(x_0, y_0) = (-1.5, -1.6)$, $\lambda_0 = 0$ se genera el subproblema cuadrático (110) donde la restricción lineal está dada por

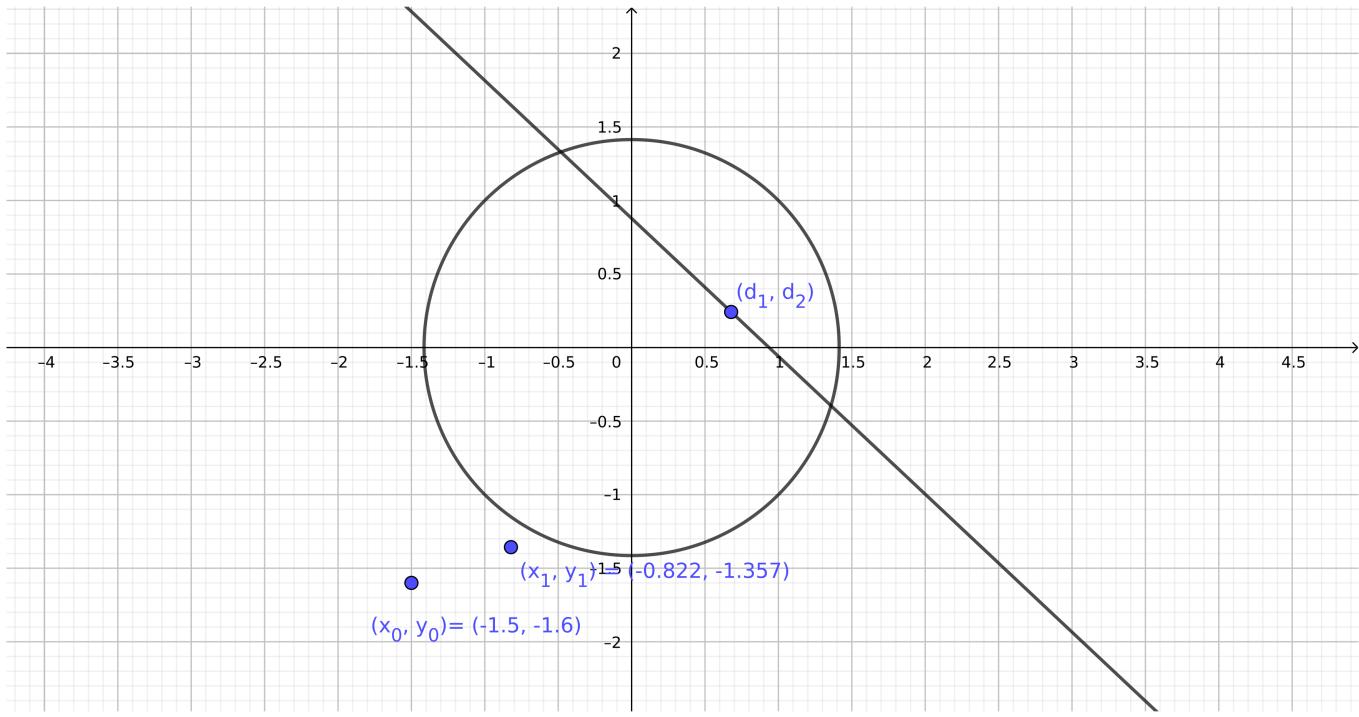
$$c(x_0, y_0) + \nabla c(x_0, y_0)^T d = 0, \text{ que es la recta}$$

$$2.81 + (-3 \quad -3.2) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 2.81 - 3d_1 - 3.2d_2 = 0.$$

El sistema (114) es el sistema de ecuaciones lineales:

$$\begin{pmatrix} 0 & -3.2 & -3 \\ -3.2 & -3 & -3.2 \\ -3 & -3.2 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \xi \end{pmatrix} = \begin{pmatrix} -2.56 \\ -4.8 \\ -2.81 \end{pmatrix}.$$

La solución del subproblema es: $d = (0.678, 0.242)$, $\xi = 0.594$. Luego: $(x_1, y_1) = (-0.822, -1.357)$, $\lambda_1 = 0.594$

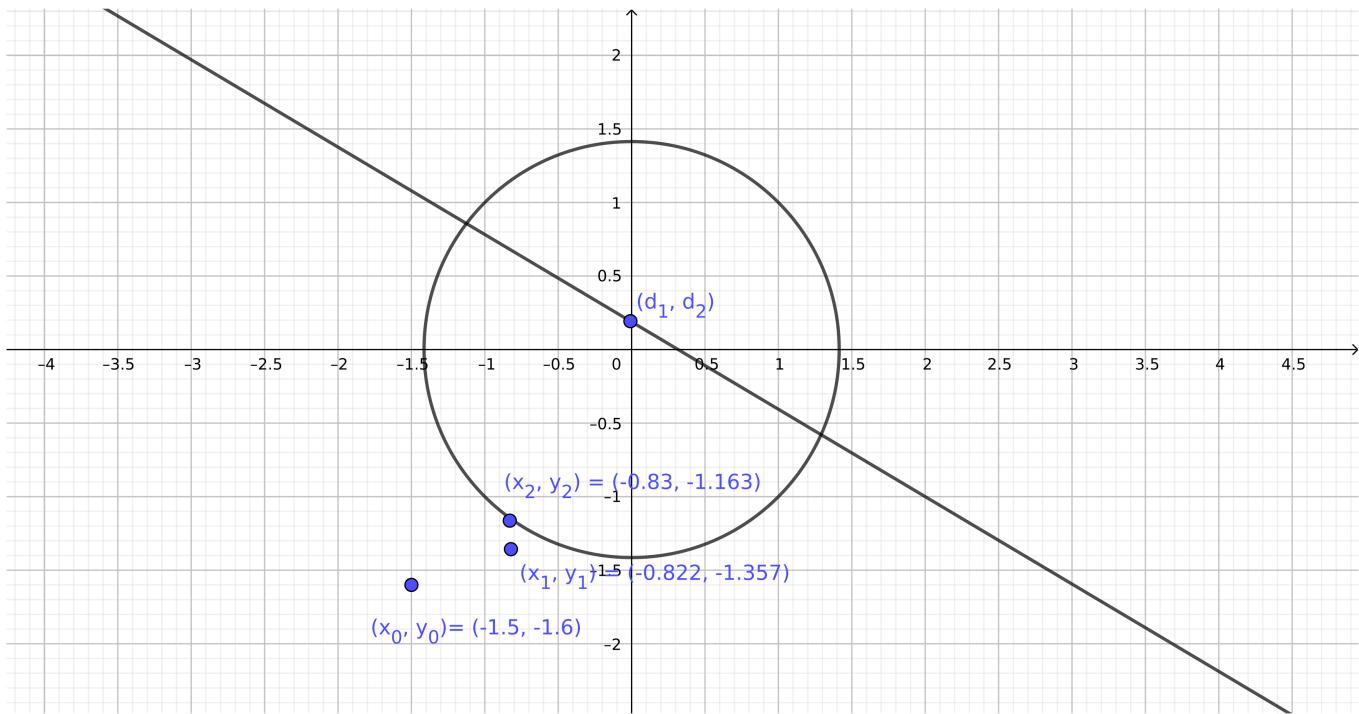


Se genera un nuevo subproblema cuadrático correspondiente (110) donde la restricción lineal está dada por

$$c(x_1, y_1) + \nabla c(x_1, y_1)^T d = 0, \text{ que es la recta}$$

$$0.517133 - 1.644d_1 - 2.714d_2 = 0.$$

Se resuelve el sistema (114) correspondiente, la solución del nuevo subproblema es: $d = (-0.0083, 0.193)$, $\xi = 0.205$. Luego: $(x_2, y_2) = (-0.83, -1.163)$, $\lambda_2 = 0.799$.



Observar que, si llamamos $\mu = \xi + \lambda_k$, el sistema (113) se puede reescribir como:

$$\begin{aligned} B(x_k, \lambda_k)d + \nabla f(x_k) + Jc(x_k)^T\mu &= 0 \\ Jc(x_k)d + c(x_k) &= 0 \end{aligned} \quad (116)$$

que representa las condiciones KKT para el problema cuadrático:

$$\begin{aligned} \text{Minimizar } & f(x_k) + \nabla f(x_k)^T d + \frac{1}{2}d^T B(x_k, \lambda_k)d \\ \text{s. a } & c_i(x_k) + \nabla c_i(x_k)^T d = 0, \forall i = 1, \dots, m. \end{aligned} \quad (117)$$

Es decir, el algoritmo SQP puede ser visto como una metodología que en cada paso minimiza un modelo cuadrático de la función objetivo del problema original sujeto a la linealización de las restricciones. Sin embargo, en el modelo cuadrático se utiliza una matriz de derivadas segundas que no es el hessiano de la función objetivo sino de la función de Lagrange.

Es interesante observar que el modelo cuadrático usual de la función objetivo, el polinomio de Taylor de segundo orden:

$$f(x_k) + \nabla f(x_k)^T d + \frac{1}{2}d^T \nabla^2 f(x_k)d \quad (118)$$

puede no funcionar bien para aproximar una solución del problema (109). Es imprescindible incorporar la información de la curvatura de las restricciones en el modelo cuadrático.

Observaciones

1. Una desventaja del método es que requiere la evaluación de derivadas de segundo orden en cada iteración, lo que puede ser computacionalmente costoso. Otra cuestión es que las hessianas $\nabla_x^2 l(x_k, \lambda_k)$ podrían no ser definidas positivas y, como resultado, el subproblema puede ser difícil de resolver. Estos dos problemas se pueden superar utilizando aproximaciones definidas positivas de las hessianas verdaderas.

Es deseable que estas matrices sean definidas positivas (al menos en el tangente linealizado). También sería bueno que estas matrices sean buenas aproximaciones de $\nabla_x^2 l(x_k, \lambda_k)$ en el límite (al menos, de nuevo, al actuar sobre vectores del tangente linealizado). Una de las formas de generar las matrices B_k es usar los métodos quasi-Newton que discutimos en el marco de la optimización sin restricciones. La principal diferencia es el uso de gradientes de l (en lugar de simplemente f) para construir las actualizaciones de las matrices. Por ejemplo, podríamos usar el método BFGS y actualizar las matrices de la siguiente manera:

$$B_{k+1} = B_k + \frac{y_k^T y_k}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}$$

donde $s_k = x_{k+1} - x_k$, $y_k = \nabla l(x_{k+1}, \lambda_{k+1}) - \nabla l(x_k, \lambda_k)$.

2. El método se puede extender para problemas con restricciones de cotas en las variables,

$$\text{Minimizar } f(x) \text{ sujeto a } h(x) = 0, l \leq x \leq u$$

manteniendo las restricciones $l \leq x \leq u$ sin linealizar, [16].

3. Se puede extender también para problemas con restricciones de desigualdad $g(x) \leq 0$ incorporando la restricción lineal de desigualdad:

$$g_i(x_k) + \nabla g_i(x_k)^T d \leq 0.$$

6.4 Métodos de barrera

Los métodos de barrera se aplican exclusivamente a problemas con restricciones de desigualdad.

Consideramos el problema

$$\text{Minimizar } f(x) \text{ sujeto a } g_i(x) \leq 0, i = 1, \dots, p \quad (119)$$

y suponemos que el interior relativo del conjunto factible es no vacío: $\text{int}(S) = \{x \in \mathbb{R}^n : g_i(x) < 0\} \neq \emptyset$.

En este caso la función objetivo se reemplaza por una función de la forma $f(x) + \mu B(x)$ donde μ es el parámetro de barrera y la función $B(x)$ definida para $x \in \text{int}(S)$ debe cumplir:

1. B debe ser continua en $\text{int}(S)$,
2. $B(x) \geq 0, \forall x \in \text{int}(S)$,
3. Si $\{x_k\} \subset S, g_i(x_k) < 0$ y, para algún $i \in \{1, \dots, p\}$ $\lim_{k \rightarrow \infty} g_i(x_k) = 0$ entonces $\lim_{k \rightarrow \infty} B(x_k) = \infty$.

La última condición establece que la barrera se acerca a infinito cuando un punto se acerca a la frontera de la región.

Algunas funciones de barrera conocidas en la literatura:

1. $B(x) = -\sum_{i=1}^p \frac{1}{g_i(x)}$ es llamada *barrera inversa*.
2. $B(x) = -\sum_{i=1}^p \ln(-g_i(x))$ es llamada *barrera logarítmica*.

La idea es, una vez elegida una función barrera, resolver una sucesión de problemas de la forma

$$\text{Minimizar } f(x) + \mu_k B(x), \text{ sujeto a } x \in \text{int}(S)$$

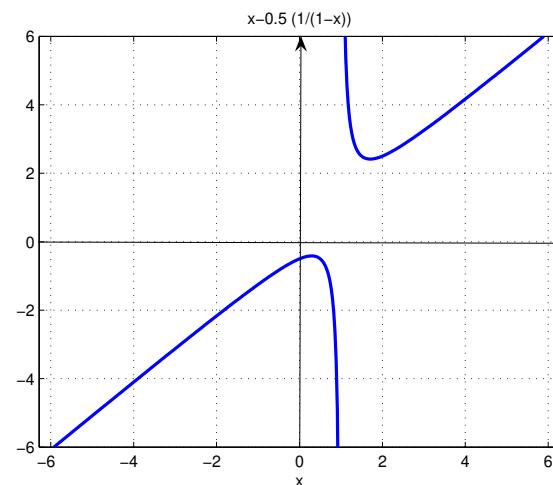
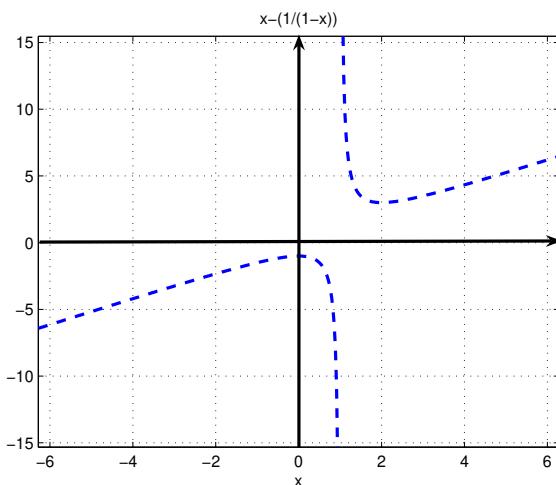
donde $\{\mu_k\}$ es una sucesión decreciente de escalares no negativos que tiende a cero.

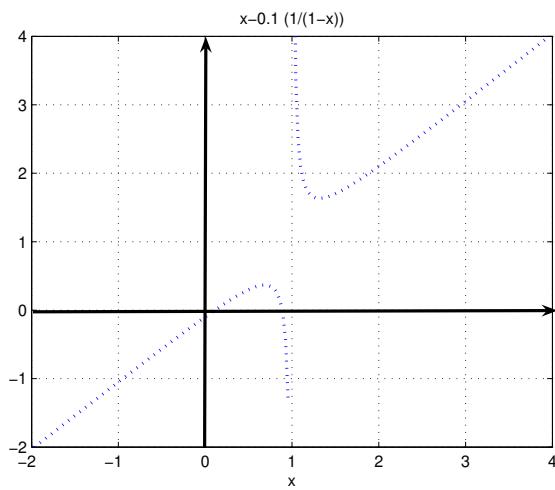
Consideremos el problema

$$\begin{aligned} \text{Min} \quad & x \\ \text{s.a} \quad & x \geq 1. \end{aligned}$$

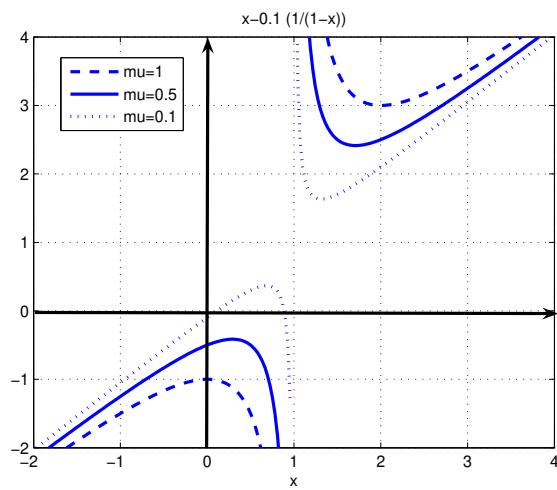
Consideramos la función barrera inversa $B(x) = -\frac{1}{1-x}$.

Las siguientes son la gráficas de la función $Q(x, \mu) = f(x) + \mu B(x)$ para los valores de $\mu = 1; 0.5; 0.1$



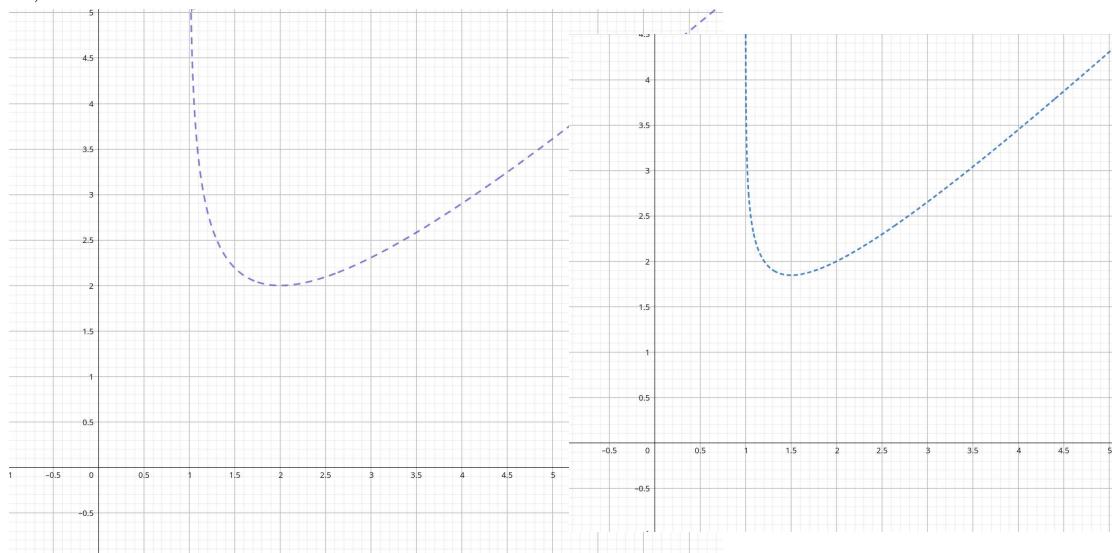


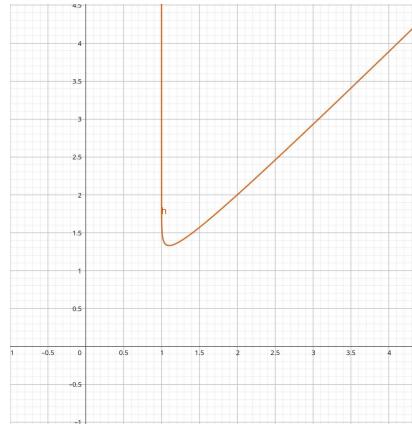
Observar las tres gráficas juntas:



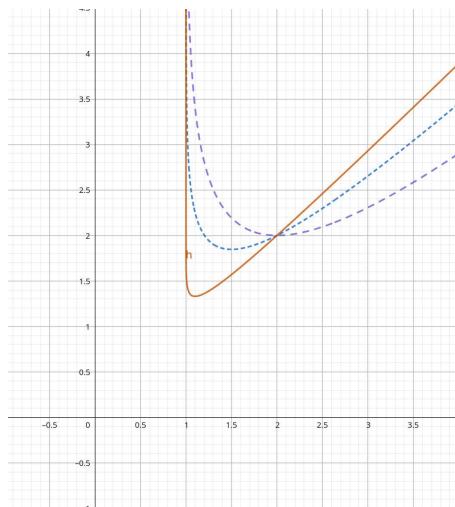
Veamos las gráficas para el mismo problema utilizando la función barrera logarítmica. $B(x) = -\ln(x - 1)$.

Las siguientes son la gráficas de la función $Q(x, \mu) = f(x) + \mu B(x)$ para los valores de $\mu = 1, 0.5, 0.1$





Observar las tres gráficas juntas:

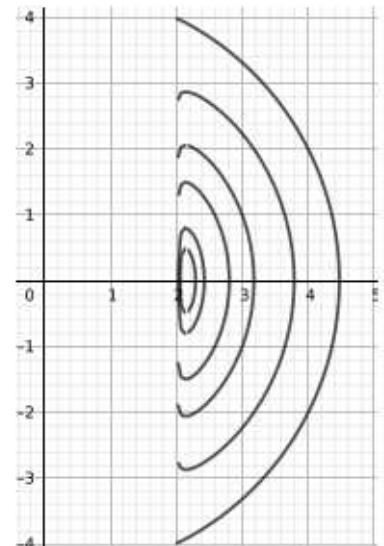
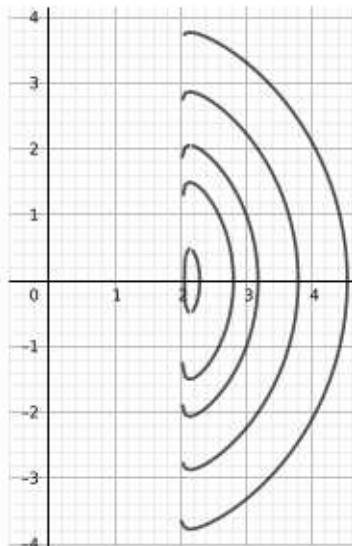
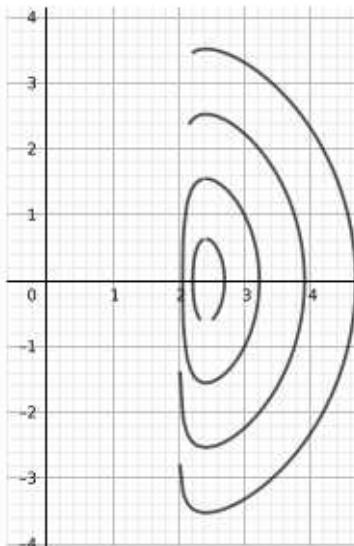


Consideremos el problema

$$\begin{aligned} \text{Min } & 0.5(x^2 + y^2) \\ \text{s.a. } & x \geq 2. \end{aligned}$$

y la función barrera logarítmica $B(x) = -\ln(x-2)$.

Las siguientes son las curvas de nivel de la función $Q(x, \mu) = f(x) + \mu B(x)$ para los valores de $\mu = 1, 0.3, 0.01$:



De manera similar al método de penalidad se pueden demostrar los siguientes resultados. Para un referencia ver [5].

Lema 6.8. Definimos $Q(x, \mu) = f(x) + \mu B(x)$. Si $\{x_k\}$ es la sucesión generada por el método de barrera utilizando una sucesión decreciente de parámetros μ_k , entonces

1. $Q(x_{k+1}, \mu_{k+1}) \leq Q(x_k, \mu_k)$
2. $B(x_k) \leq B(x_{k+1})$
3. $f(x_{k+1}) \leq f(x_k)$.

Teorema 6.9. Sea $\{x_k\}$ una sucesión generada por el método de barrera utilizando una sucesión decreciente de parámetros μ_k . Entonces, todo punto límite de $\{x_k\}$ es solución de (119).

Se puede analizar la convergencia del método cuando se busca un punto estacionario del subproblema. Supongamos que se considera la barrera logarítmica:

$$Q(x, \mu) = f(x) - \mu \sum_{i=1}^p \ln(-g_i(x)).$$

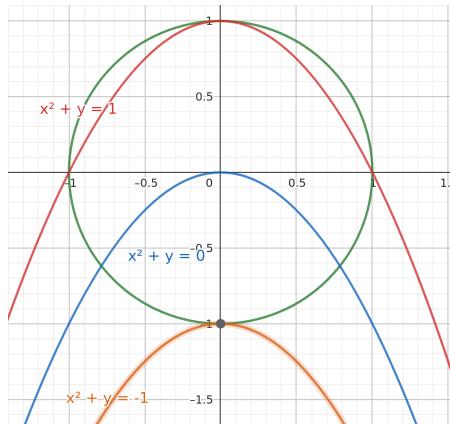
Luego, un punto estacionario de $Q(x, \mu)$ cumple:

$$\nabla f(x) + \mu \sum_{i=1}^p \frac{\nabla g_i(x)}{-g_i(x)} = 0.$$

Teorema 6.10. Sea $\{x_k\}$ la sucesión tal que $\nabla f(x_k) + \mu_k \sum_{i=1}^p \frac{\nabla g_i(x_k)}{-g_i(x_k)} = 0$. Si $x_k \rightarrow x^*$ solución regular del problema de Minimizar $f(x)$ sujeto a $g_i(x) \leq 0, i = 1, \dots, p$ entonces $\lambda_k^i = \frac{-\mu_k}{g_i(x_k)}$ converge al multiplicador de Lagrange asociado.

Ejemplo: Consideramos el problema

$$\begin{aligned} \text{Min } & x^2 + y \\ \text{s. a } & x^2 + y^2 \leq 1 \end{aligned}$$



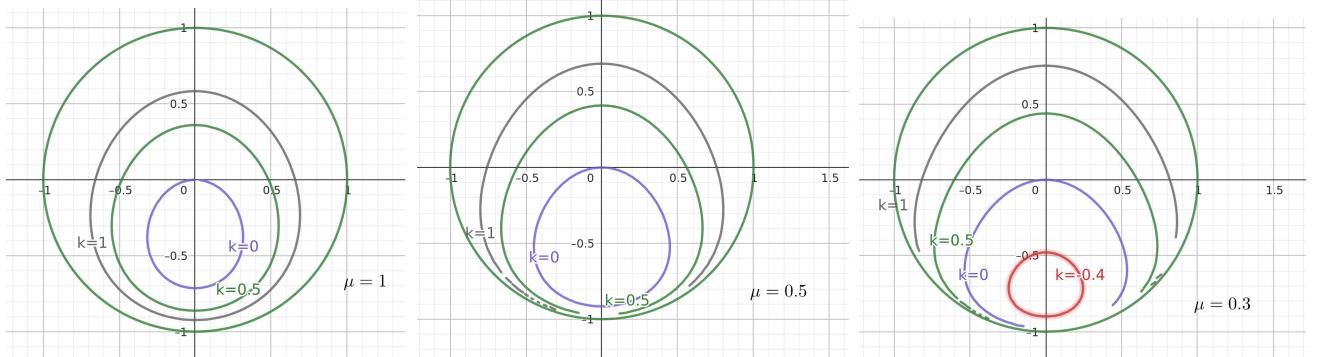
Gráficamente el problema tiene un minimizador local: $x^* = (0, -1)$ y en ese punto el multiplicador de Lagrange es $\mu^* = \frac{1}{2}$.

Si se considera la barrera logarítmica:

$$Q(x, y, \mu) = x^2 + y - \mu \ln(1 - x^2 - y^2)$$

cuyo dominio es el conjunto de puntos tales que $\{(x, y) : x^2 + y^2 < 1\}$.

Observar las curvas de nivel de Q para los diferentes valores de μ :



Si buscamos los puntos estacionarios de Q se obtiene:

$$\nabla Q(x, y, \mu) = \begin{pmatrix} 2x + \frac{\mu}{1-x^2-y^2}2x \\ 1 + \frac{\mu}{1-x^2-y^2}2y \end{pmatrix} = 0$$

si y solo si $(x_\mu, y_\mu) = (0, \mu - \sqrt{\mu^2 + 1})$ o $(\tilde{x}_\mu, \tilde{y}_\mu) = (0, \mu + \sqrt{\mu^2 + 1})$.

Observar que solo (x_μ, y_μ) es factible y se puede ver que

$$\lim_{\mu \rightarrow 0} (x_\mu, y_\mu) = \lim_{\mu \rightarrow 0} (0, \mu - \sqrt{\mu^2 + 1}) = (0, -1)$$

que es el minimizador local.

En este punto se tiene que

$$\frac{-\mu}{g(x_\mu, y_\mu)} = \frac{-\mu}{x_\mu^2 + y_\mu^2 - 1} = \frac{-\mu}{(\mu - \sqrt{\mu^2 + 1})^2 - 1} = \frac{-\mu}{2\mu^2 - 2\mu\sqrt{\mu^2 + 1}} \xrightarrow{\mu \rightarrow 0} \frac{1}{2}.$$

6.5 Ejercicios.

1. Demostrar el Teorema 6.3 (la demostración es similar a la del Teorema 6.2).
2. Considere el problema

$$\begin{aligned} \text{Min } & x_1^2 - x_2 \\ \text{s.a } & x_1 + x_2 = 6, x_1 \geq 0. \end{aligned}$$

Mostrar que si se utiliza el método de penalidad con la función de penalidad cuadrática, entonces $(x_1^k, x_2^k) \rightarrow x^*$ cuando $\rho_k \rightarrow \infty$.

3. En el problema

$$\begin{aligned} \text{Min } & x^2 + y^2 \\ \text{s. a } & x + y^2 = 1. \end{aligned}$$

Demostrar que para todo $\rho > 0$, los puntos $(\frac{1}{2}, \frac{1}{\sqrt{2}})$, $(\frac{1}{2}, -\frac{1}{\sqrt{2}})$ son minimizadores locales de la función \bar{l} del Teorema 6.4 con $\lambda = -1$.

4. Demostrar que x^* es un minimizador local de (97) si y solo si (x^*, z^*) es minimizador local de (98) donde $z_i^* = \sqrt{-g_i(x^*)}$.

5. Considere el problema

$$\begin{aligned} \text{Min } & 2x_1^2 + 2x_1x_2 + x_2^2 - 2x_2 \\ \text{s.a } & x_1 = 0. \end{aligned}$$

(a) Calcule la solución (x^*, λ^*) .

(b) Puede afirmar que existe $\bar{\rho}$ para el cual x^* es minimizador de la función Lagrangiano aumentada $L(x, \lambda^*, \rho)$ para todo $\rho \geq \bar{\rho}$?

(c) Hacer 3 iteraciones el método de Lagrangiano aumentado comenzando con $x_0 = (0, 0)$, $\rho_1 = 1$, $\lambda_1 = 0$.

6. [4] **Optativo** Considere el problema

$$\begin{aligned} \text{Min } & -\frac{x_1^2}{2} + 2x_2 \\ \text{s. a } & x_1^2 + x_2^2 = 1. \end{aligned}$$

Mostrar que $x^* = (0, -1)$ es la única solución del problema con $\lambda^* = 1$. Suponga que en alguna iteración se tiene el punto $x_k = (\delta, -\sqrt{1-\delta^2})$ con $\delta > 0$ pequeño. Mostrar que si a partir de x_k el paso es calculado usando el modelo (118), entonces el nuevo punto está más lejos de la solución que el anterior. ($\|x_k + d_k - x^*\| \approx 2\|x_k - x^*\|$). Que sucede con el paso obtenido mediante el modelo (113)?.

7. Considere el problema

$$\begin{aligned} \text{Min } & 2x_1^2 + 9x_2 \\ \text{s.a } & x_1 + x_2 \geq 4. \end{aligned}$$

Mostrar que si se utiliza el método de barrera inversa, entonces $(x_1^k, x_2^k) \rightarrow x^*$ cuando $\mu_k \rightarrow 0$.

8. Considere el problema

$$\begin{aligned} \text{Min } & -30x_1 + 3x_1^2 - 8x_2 + 2x_2^2 \\ \text{s.a } & 3x_1 + 2x_2 \leq 6. \end{aligned}$$

Calcular la solución mediante la aplicación del método de barrera logarítmica. Calcular el multiplicador de Lagrange asociado.

References

- [1] N. Patriksson, M.; Andréasson and A. Evgrafov. *An Introduction to Continuous Optimization - Foundations and Fundamental Algorithm*. Courier Dover Publications., 2016.
- [2] S.G. Griva, I.; Nash and A Sofer. *Linear and Nonlinear Optimization*. Society for Industrial and Applied Mathematics., 2009.
- [3] M. Ehrgott. *Multicriteria Optimization*. Springer-Verlag Berlin Heidelberg., 2005.
- [4] A. A. Ribeiro and E. W. Karas. *Otimização contínua*. Cengage Learing, 2013.
- [5] D. G. Luenberger. *Linear and Nonlinear Programming, 2nd Edition*. Addison- Wesley Inc., Reading, Massachusetts, 1984.
- [6] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer, 1999.
- [7] A. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust region methods*. MPS-SIAM Series on Optimization, 2000.
- [8] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Wiley, third edition, 2006.
- [9] M. Guignard. Generalized kuhn-tucker conditions for mathematical programming problems in banach space. *SIAM Journal on Control*, 7:232–241, 1969.
- [10] F.J. Gould and J.W. Tolle. A necessary and sufficient qualification for constrained optimization. *SIAM Journal on Applied Mathematics*, 20:164–172, 1969.
- [11] O.L. Mangasarian and S. Fromovitz. The Fritz John optimality conditions in the presence of equality and inequality constraints. *Journal of Mathematical Analysis and Applications*, 17:37–47, 1967.
- [12] R. Janin. Directional derivative of the marginal function in nonlinear programming. *Mathematical Programming Study*, 21:110–126, 1984.
- [13] A. V. Arutyunov. Perturbations of extremum problems with constraints and necessary optimality conditions. *Journal of Soviet Mathematics*, 54:1342–1400, 1991.
- [14] R. Andreani, C.E. Echagüe, and M.L. Schuverdt. Constant-rank condition and second-order constraint qualification. *Journal of Optimization Theory and Applications*, 146:155–266, 2009.
- [15] D.P. Bertsekas. *Nonlinear Programming: 2nd Edition*. Athena Scientific, 1999.
- [16] J. M. Martínez and S. A. Santos. *Métodos computacionais de Otimização*. Sociedade Brasileira de Matemática, 1995.