

Métodos Numéricos de Optimización con restricciones.

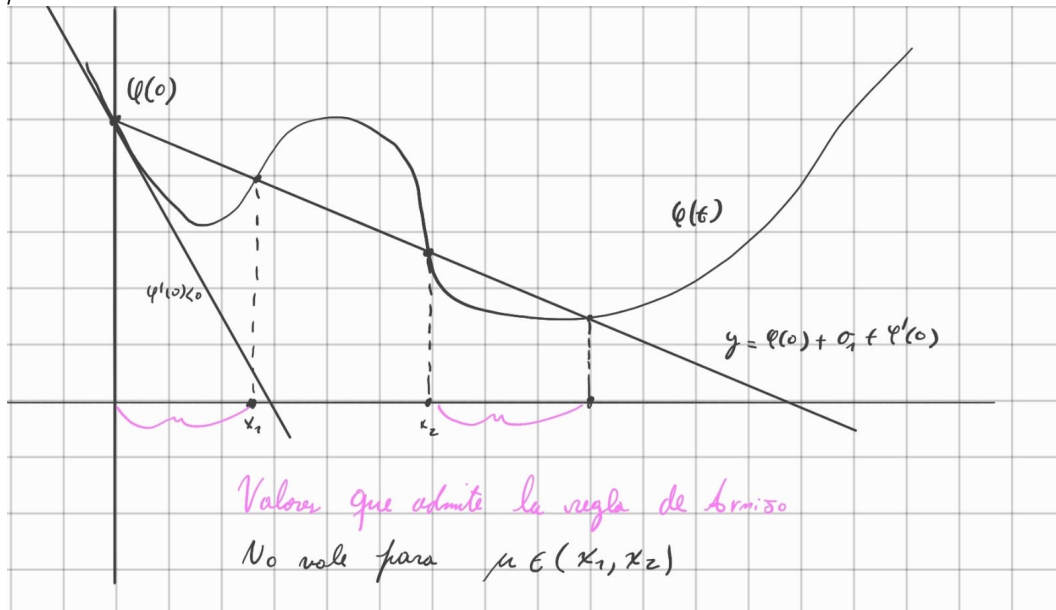
Bustos Jordi
Práctica II

September 24, 2025

Sección 3.2

Ejercicio 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? Pruébalo o dé un contraejemplo que puede ser gráfico.

Proof. Análogamente al ejemplo visto en clase podemos ver que no siempre se cumple la condición de Armijo para $0 < \mu < \lambda$ pues en este caso, si $\mu \in (x_1, x_2)$ no se cumple la regla de armijo y sin embargo se vale que $\mu < \lambda$.



□

Sección 3.3

Ejercicio 2. Considere la función

$$f(x, y) = x - y + 2x^2 + 2xy + y^2.$$

- Muestre que $d = (-1, 0)$ es una dirección de descenso para f en $(0, 0)$. Analizar cuál es el paso óptimo que se puede dar en esa dirección para hacer decrecer el valor de f utilizando búsqueda exacta.
- 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 $\sigma_1 = 1/4$.

Proof. Para demostrar (a) notemos que f es diferenciable y por lo tanto si $\nabla f(0,0)^T \cdot d < 0 \implies d$ es una dirección de descenso. En efecto, sea $d = (-1, 0)$:

$$\begin{aligned}\nabla f(x, y) &= \begin{pmatrix} 1 + 4x + 2y \\ -1 + 2x + 2y \end{pmatrix} \\ \nabla f(0, 0) &= \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ \nabla f(0, 0)^T \cdot d &= -1 < 0\end{aligned}$$

Para hallar la longitud del paso óptimo, definimos $\phi(t) = f((0, 0) + t \cdot d) = f(-t, 0) = -t + 2t^2$, luego $\phi'(t) = -1 + 4t = 0 \iff t = \frac{1}{4}$ que es la longitud de paso óptima en la dirección d . Para la parte (b) consideremos la regla de Armijo:

$$f(x + td) \leq f(x) + \sigma_1 t \nabla f(x)^T d$$

La dirección de máximo decrecimiento está dada por $-\nabla f(0, 0) = (-1, 1)$, si $\sigma_1 = \frac{1}{4}$ la regla de Armijo se traduce en:

$$\begin{aligned}f((0, 0) + t \cdot (-1, 1)) &\leq f(0, 0) + \frac{1}{4} t \nabla f(0, 0)^T \cdot (-1, 1) \\ f(-t, t) &\leq \frac{1}{4} t \cdot (-2) \\ t^2 - 2t &\leq -\frac{1}{2} t \\ t^2 - \frac{3}{2} t &\leq 0 \\ t(t - \frac{3}{2}) &\leq 0\end{aligned}$$

Por lo tanto, el intervalo de paso máximo es $(0, \frac{3}{2}]$. □

Ejercicio 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?

Proof. Análogamente al ejercicio anterior, para (a) se tiene que, dado $d = (0, 1)$:

$$\begin{aligned}\nabla f(x, y) &= \begin{pmatrix} 4x - 2y + 6x^2 + 4x^3 \\ 2y - 2x \end{pmatrix} \\ \nabla f(0, -2) &= \begin{pmatrix} 4 \\ -4 \end{pmatrix} \\ \nabla f(0, -2)^T \cdot d &= -4 < 0\end{aligned}$$

Luego, d es una dirección de descenso para f en $(0, -2)$.

Para (b) consideremos la regla de Armijo con $\sigma_1 = \frac{4}{5}$ y $t = 1$:

$$f((0, -2) + 1 \cdot (0, 1)) \leq f(0, -2) + \frac{4}{5} \cdot 1 \cdot \nabla f(0, -2)^T \cdot (0, 1)$$

$$f(0, -1) \leq f(0, -2) + \frac{4}{5} \cdot (-4)$$

$$1 \leq 4 - \frac{16}{5}$$

$$1 \leq \frac{4}{5}$$

Absurdo, luego $t = 1$ no verifica la regla de Armijo con $\sigma_1 = \frac{4}{5}$.

Para hallar los valores de σ_1 para los cuales $t = 1$ verifica la regla de Armijo, consideremos:

$$1 \leq 4 + \sigma_1 \cdot (-4)$$

$$\sigma_1 \leq \frac{3}{4}$$

Por lo tanto, $t = 1$ verifica la regla de Armijo para $\sigma_1 \in (0, \frac{3}{4}]$. □

Ejercicio 4. Sea f una función diferenciable tal que $\nabla f(x) \neq 0$. 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 x .

Proof. Como la matriz $H(x)$ es definida positiva y $\nabla f(x) \neq 0$ se tiene que

$$\nabla f(x)^T \cdot H(x) \cdot \nabla f(x) > 0$$

$$-\nabla f(x)^T \cdot H(x) \cdot \nabla f(x) < 0$$

$$\nabla f(x)^T \cdot d < 0$$

y como f es diferenciable $d = -H(x)\nabla f(x)$ es una dirección de descenso. □

Sección 3.4

Ejercicio 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 $\sigma_1 = 1/5$?

Proof. Por definición de dirección de Newton buscamos:

$$\nabla^2 f(2, 1) \cdot d = -\nabla f(2, 1)$$

Donde

$$\nabla f(x, y) = \begin{pmatrix} 2(x - 2y) + 4x^3 \\ -4(x - 2y) \end{pmatrix}$$

$$\nabla f(2, 1) = \begin{pmatrix} 16 \\ 0 \end{pmatrix}$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 2 + 12x^2 & -4 \\ -4 & 8 \end{pmatrix}$$

$$\nabla^2 f(2, 1) = \begin{pmatrix} 50 & -4 \\ -4 & 8 \end{pmatrix}$$

Por lo tanto,

$$\begin{pmatrix} 50 & -4 \\ -4 & 8 \end{pmatrix} \cdot d = \begin{pmatrix} -16 \\ 0 \end{pmatrix}$$

$$d = \begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{6} \end{pmatrix}$$

Una vez más, recordemos la regla de Armijo:

$$f(x + t \cdot d) \leq f(x) + \sigma_1 t \nabla f(x)^T \cdot d$$

En particular,

$$f((2, 1) + 1 \cdot (-1/3, -1/6)) \leq f(2, 1) + \frac{1}{5} \cdot 1 \cdot \nabla f(2, 1)^T \cdot (-1/3, -1/6)$$

$$\left(\frac{5}{3}\right)^4 \leq 16 - \frac{16}{15}$$

Que es verdadero, luego $t = 1$ cumple la regla de Armijo con $\sigma_1 = \frac{1}{5}$. □

Ejercicio 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 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.

Proof. Para (b) consideremos la dirección de Cauchy i.e la dirección de máximo decrecimiento dada por $-\nabla f(x_k)$. Luego, siguiendo el método propuesto:

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

$$\nabla f(2, 0) = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$$

$$d_0 = -\nabla f(2, 0) = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$$

Reemplazando en la ecuación:

$$f((2, 0) + 1 \cdot (-4, 2)) \stackrel{?}{\leq} f(2, 0) + \frac{1}{2} \cdot 1 \cdot \nabla f(2, 0)^T \cdot (-4, 2)$$

$$f(-2, 2) \stackrel{?}{\leq} f(2, 0) - 10$$

$$8 \stackrel{?}{\leq} -6$$

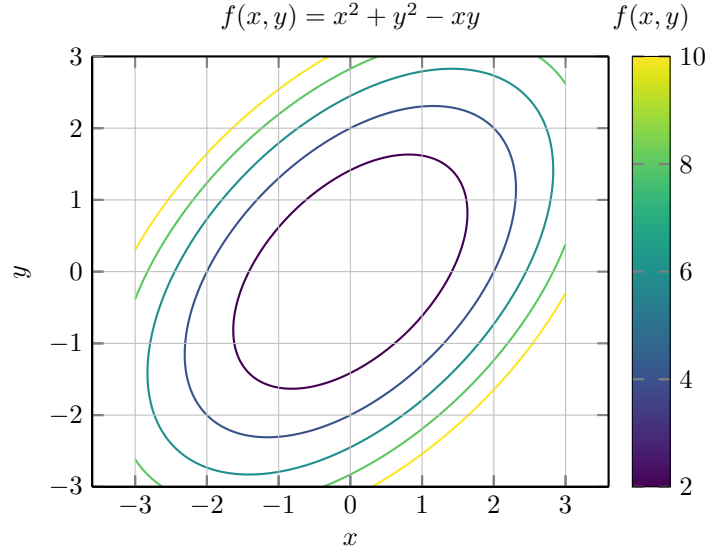


Figure 1: (a) Curvas de nivel de $f(x, y) = x^2 + y^2 - xy$.

que es falso, luego $t := \frac{1}{2}$ y repetimos:

$$\begin{aligned} f((2, 0) + \frac{1}{2} \cdot (-4, 2)) &\stackrel{?}{\leq} f(2, 0) + \frac{1}{4} \cdot \nabla f(2, 0)^T \cdot (-4, 2) \\ f(0, 1) &\stackrel{?}{\leq} f(2, 0) - 5 \\ 1 &\stackrel{?}{\leq} -1 \end{aligned}$$

que también es falso, luego $t := \frac{1}{4}$ y de nuevo:

$$\begin{aligned} f((2, 0) + \frac{1}{4} \cdot (-4, 2)) &\stackrel{?}{\leq} f(2, 0) + \frac{1}{8} \cdot \nabla f(2, 0)^T \cdot (-4, 2) \\ \iff f(1, \frac{1}{2}) &\stackrel{?}{\leq} f(2, 0) - \frac{5}{2} \\ \iff \frac{3}{4} &\leq \frac{3}{2} \end{aligned}$$

que es verdadero, entonces definimos $x_1 = (2, 0) + \frac{1}{4} \cdot (-4, 2) = (1, \frac{1}{2})$.
Repetimos el proceso para x_1 :

$$\begin{aligned} \nabla f(1, \frac{1}{2}) &= \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \\ d_1 &= -\nabla f(1, \frac{1}{2}) = \begin{pmatrix} -\frac{1}{2} \\ 0 \end{pmatrix} \end{aligned}$$

Luego, reiniciando $t = 1$:

$$\begin{aligned} f((1, \frac{1}{2}) + (-\frac{1}{2}, 0)) &\stackrel{?}{\leq} f(1, \frac{1}{2}) + \frac{1}{2} \nabla f(1, \frac{1}{2})^T \cdot (-\frac{1}{2}, 0) \\ \iff f(\frac{1}{2}, \frac{1}{2}) &\stackrel{?}{\leq} f(1, \frac{1}{2}) - \frac{1}{8} \\ \iff \frac{1}{4} &\stackrel{?}{\leq} \frac{3}{4} - \frac{1}{8} \\ \iff \frac{1}{4} &\leq \frac{5}{8} \end{aligned}$$

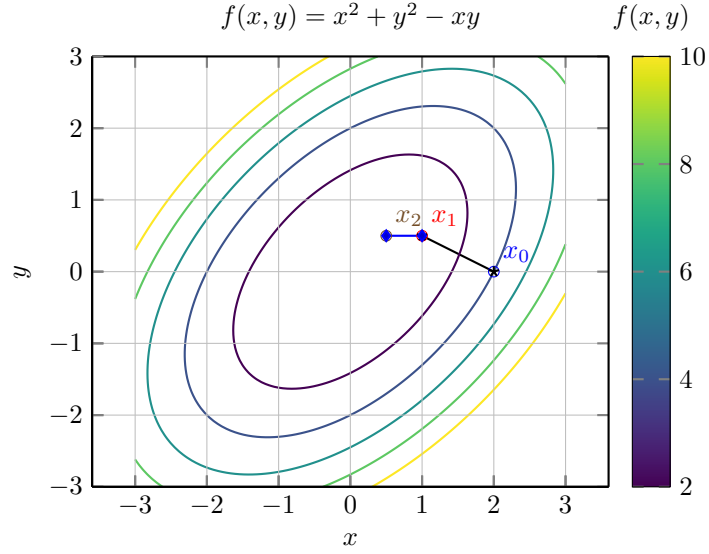


Figure 2: Curvas de nivel de $f(x, y) = x^2 + y^2 - xy$ con los puntos x_0, x_1, x_2 y las direcciones de descenso.

que es verdadero, luego $x_2 = (1, \frac{1}{2}) + 1 \cdot (-\frac{1}{2}, 0) = (\frac{1}{2}, \frac{1}{2})$.

Por lo tanto los puntos sobre la curva de nivel quedarían así:

Por último, para el inciso (c) utilizando la dirección de Newton dada por $\nabla^2 f(x_k) \cdot d_k = -\nabla f(x_k)$, si consideramos $x_0 = (2, 0)$:

$$\begin{aligned}\nabla^2 f(x, y) &= \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \\ \nabla^2 f(2, 0) &= \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \\ \nabla^2 f(2, 0) \cdot d &= -\nabla f(2, 0)\end{aligned}$$

Resolviendo el sistema obtenemos $d = (-2, 0) \implies x_1 = x_0 + d = (0, 0)$ que es el mínimo de la función. \square

Sección 3.6

Lemma Determinante de una matriz. Supongamos que A es una matriz no singular y $u, v \in \mathbb{R}^n$ no nulos, entonces:

$$\det(A + uv^T) = \det(A)(1 + v^T A^{-1}u)$$

Ejercicio 7 Fórmula de Sherman-Morrison. Sean $u, v \in \mathbb{R}^n$ no nulos y $A \in \mathbb{R}^{n \times n}$ 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

$$B^{-1} = A^{-1} - \frac{1}{\sigma} A^{-1} uv^T A^{-1}.$$

Proof. Para la ida consideremos el contrarrecíproco, si $1 + v^T A^{-1}u = 0 \implies \det(A)(1 + v^T A^{-1}u) = 0 \implies \det(A + uv^T) = 0 \implies A + uv^T$ es singular por el lema anterior.

Para la vuelta, si $1 + v^T A^{-1} u \neq 0$, $X = A + uv^T$ e $Y = A^{-1} - \frac{A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u}$ entonces:

$$\begin{aligned}
XY &= (A + uv^T) \left(A^{-1} - \frac{A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u} \right) \\
&= AA^{-1} + uv^T A^{-1} - \frac{AA^{-1} uv^T A^{-1} + uv^T A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u} \\
&= I + uv^T A^{-1} - \frac{uv^T A^{-1} + uv^T A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u} \\
&= I + uv^T A^{-1} - \frac{u(1 + v^T A^{-1} u) v^T A^{-1}}{1 + v^T A^{-1} u} \\
&= I + uv^T A^{-1} - uv^T A^{-1} \\
&= I
\end{aligned}$$

Análogamente se puede probar que $YX = I$, luego $Y = X^{-1}$ y por lo tanto B es no singular y su inversa es la indicada. \square

Ejercicio 8. 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.

Proof. Sea $\rho_k := \frac{1}{y_k^T s_k} > 0$, definimos la actualización BFGS para la inversa como:

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

Veamos primero que es simétrica:

$$\begin{aligned}
H_{k+1}^T &= ((I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T)^T \\
&= (I - \rho_k y_k s_k^T)^T H_k^T (I - \rho_k s_k y_k^T)^T + \rho_k s_k s_k^T \\
&= (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T \\
&= H_{k+1}
\end{aligned}$$

Usando que si A, B, C matrices entonces $(ABC)^T = C^T B^T A^T$ y que tanto H_k como $\rho_k s_k s_k^T$ son simétricas. Veamos ahora que está definida positiva, sea $z \in \mathbb{R}^n \setminus \{0\}$ quiero ver que $z^T H_{k+1} z > 0$. Sea

$$w := (I - \rho_k y_k s_k^T) z = z - \rho_k y_k (s_k^T z)$$

Calculemos $z^T H_{k+1} z$:

$$\begin{aligned}
z^T H_{k+1} z &= z^T (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) z + z^T \rho_k s_k s_k^T z \\
&= w^T H_k w + \rho_k (s_k^T z)^2 \geq 0 \quad \text{pues } H_k \text{ es definida positiva y } \rho_k > 0
\end{aligned}$$

Analicemos si la desigualdad es estricta, si $z \neq 0$ entonces:

- Si $z^T s_k \neq 0 \implies \rho_k (z^T s_k)^2 > 0 \implies z^T H_{k+1} z > 0$.
- Si $z^T s_k = 0 \implies w = z - \rho_k y_k (s_k^T z) = z \implies z^T H_{k+1} z = z^T H_k z > 0$ pues $z \neq 0$ y H_k es definida positiva.

\square

Ejercicio 9. 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.

Proof. Recordemos que $s_k = x_{k+1} - x_k$ e $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$. Si la búsqueda lineal es exacta, definamos:

- $g_k = \nabla f(x_k)$.
- $p_k = -H_k g_k$ la dirección de descenso.
- $s_k = x_{k+1} - x_k = t_k p_k$ con t_k la longitud del paso óptimo.
- $y_k = g_{k+1} - g_k$.
- $\phi_k(t) = f(x_k + t p_k)$.

Notemos que $-g_k^T H_k g_k < 0$. Por lo tanto existe un t_k que minimiza la función. Luego,

$$\begin{aligned} \phi'_k(t_k) &= 0 \quad \text{y} \quad g_{k+1}^T p_k = 0 \\ \implies y_k^T s_k &= (g_{k+1} - g_k)^T (x_{k+1} - x_k) = g_{k+1}^T s_k - g_k^T s_k \end{aligned}$$

Como $s_k = t_k p_k$ y $g_{k+1}^T p_k = 0$ se tiene que:

$$\begin{aligned} g_{k+1}^T s_k &= t_k g_{k+1}^T p_k = 0 \\ \implies y_k^T s_k &= -g_k^T s_k = -t_k g_k^T p_k > 0 \quad \text{pues } t_k > 0 \text{ y } g_k^T p_k < 0 \end{aligned}$$

Si la búsqueda es de Wolfe se tiene que dado $\sigma_2 \in (0, 1)$:

$$g_{k+1}^T p_k \geq \sigma_2 g_k^T p_k$$

Nuevamente $s_k = t_k p_k$ así:

$$y_k^T s_k = (g_{k+1} - g_k)^T s_k = t_k (g_{k+1}^T p_k - g_k^T p_k)$$

Aplicando la desigualdad en el primer termino se obtiene:

$$g_{k+1}^T p_k - g_k^T p_k \geq \sigma_2 g_k^T p_k - g_k^T p_k = (\sigma_2 - 1) g_k^T p_k$$

Entonces,

$$y_k^T s_k \geq t_k (\sigma_2 - 1) g_k^T p_k > 0 \quad \text{pues } t_k > 0, \sigma_2 - 1 < 0 \text{ y } g_k^T p_k < 0$$

□