

Métodos Numéricos de Optimización con restricciones.

Bustos Jordi
Práctica III

23 de octubre de 2025

Capítulo IV

Ejercicio 1. Si x^* es un minimizador local del problema

$$\begin{aligned} &\text{mín } f(x) \\ &\text{s.a. } h(x) = 0 \end{aligned}$$

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

$$-\nabla f(x^*) = \sum_{i=1}^m \lambda_i \nabla h_i(x^*).$$

Demostración. La existencia de los escalares λ_i se sigue del Teorema 4.5 del apunte, veamos la unicidad. Para ello, supongamos que existen dos conjuntos de escalares $\{\lambda_i\}$ y $\{\mu_i\}$ que cumplen la condición del enunciado, es decir,

$$-\nabla f(x^*) = \sum_{i=1}^m \lambda_i \nabla h_i(x^*) = \sum_{i=1}^m \mu_i \nabla h_i(x^*).$$

Restando ambas expresiones obtenemos

$$0 = \sum_{i=1}^m (\lambda_i - \mu_i) \nabla h_i(x^*).$$

Dado que los vectores $\{\nabla h_1(x^*), \dots, \nabla h_m(x^*)\}$ son linealmente independientes por hipótesis, la única solución a esta ecuación es que todos los coeficientes sean cero, es decir, $\lambda_i - \mu_i = 0$ para todo $i = 1, \dots, m$. Por lo tanto, $\lambda_i = \mu_i$ para todo i , lo que demuestra la unicidad de los escalares λ_i . \square

Ejercicio 2. Sea x^* un minimizador local de f sujeto a $g_i(x) \leq 0$, $i = 0, \dots, p$ y $A(x^*)$ es el conjunto de restricciones activas entonces x^* es un minimizador local del siguiente problema con restricciones de igualdad

$$\begin{aligned} &\text{mín } f(x) \\ &\text{s.a. } g_i(x) = 0, i \in A(x^*). \end{aligned}$$

Demostración. En efecto, si x^* es un minimizador local de f sujeto a las restricciones $g_i(x) \leq 0$, entonces existe un entorno U de x^* tal que para todo $x \in U$ que cumple las restricciones $g_i(x) \leq 0$, se tiene $f(x) \geq f(x^*)$. Ahora, consideremos el problema con restricciones de igualdad $g_i(x) = 0$ para $i \in A(x^*)$. Dado que x^* satisface estas restricciones (ya que $g_i(x^*) = 0$ para $i \in A(x^*)$), podemos tomar el mismo entorno U alrededor de x^* . Cualquier punto x en este entorno que satisfaga las restricciones de igualdad $g_i(x) = 0$ para $i \in A(x^*)$ también satisface las restricciones $g_i(x) \leq 0$ del problema original, ya que $g_i(x) = 0$ implica $g_i(x) \leq 0$. Por lo tanto, para cualquier x en este entorno que satisfaga las restricciones de igualdad, se cumple $f(x) \geq f(x^*)$. Esto demuestra que x^* es un minimizador local del problema con restricciones de igualdad. \square

Ejercicio 3. Análogamente al caso de restricciones de igualdad, demostrar que si

$$\{\nabla g_i(x^*) : i \in A(x^*)\}$$

Son l.i entonces $T_{x^*} = T_{x^*}^{\text{lin}}$ donde

$$T_{x^*}^{\text{lin}} = \{d \in \mathbb{R}^n : \nabla g_i(x^*)^T d = 0, i \in A(x^*)\}.$$

Demostración. En efecto, para demostrar que $T_{x^*} = T_{x^*}^{\text{lin}}$, debemos mostrar que ambos conjuntos son iguales. Primero, demostraremos que $T_{x^*} \subseteq T_{x^*}^{\text{lin}}$. Sea $d \in T_{x^*}$. Por definición de T_{x^*} , existe una curva $\gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^n$ tal que $\gamma(0) = x^*$, $\gamma'(0) = d$ y $g_i(\gamma(t)) \leq 0$ para todo t en un entorno de 0 y para todo $i \in A(x^*)$. Dado que $g_i(x^*) = 0$ para $i \in A(x^*)$, podemos aplicar la regla de la cadena para obtener

$$\left. \frac{d}{dt} g_i(\gamma(t)) \right|_{t=0} = \nabla g_i(x^*)^T \gamma'(0) = \nabla g_i(x^*)^T d.$$

Pero la función $t \mapsto g_i(\gamma(t))$ tiene un máximo local en $t = 0$ (ya que $g_i(\gamma(t)) \leq 0$ y $g_i(\gamma(0)) = 0$), por lo que su derivada en $t = 0$ debe ser cero. Por lo tanto, $\nabla g_i(x^*)^T d = 0$ para todo $i \in A(x^*)$, lo que implica que $d \in T_{x^*}^{\text{lin}}$. Así, hemos demostrado que $T_{x^*} \subseteq T_{x^*}^{\text{lin}}$.

Por otro lado, demostraremos que $T_{x^*}^{\text{lin}} \subseteq T_{x^*}$. Sea $d \in T_{x^*}^{\text{lin}}$. Por definición de $T_{x^*}^{\text{lin}}$, tenemos que $\nabla g_i(x^*)^T d = 0$ para todo $i \in A(x^*)$. Consideremos la curva $\gamma(t) = x^* + td$. Para t suficientemente pequeño, tenemos que $g_i(\gamma(t)) = g_i(x^* + td)$. Usando la expansión en serie de Taylor, obtenemos

$$g_i(\gamma(t)) = g_i(x^*) + t \nabla g_i(x^*)^T d + o(t).$$

Dado que $g_i(x^*) = 0$ para $i \in A(x^*)$ y $\nabla g_i(x^*)^T d = 0$, se sigue que $g_i(\gamma(t)) = o(t)$. Por lo tanto, para t suficientemente pequeño, $g_i(\gamma(t)) \leq 0$. Esto implica que $d \in T_{x^*}$ y así $T_{x^*}^{\text{lin}} \subseteq T_{x^*}$. \square

Ejercicio 4. Dados los vectores $v_1, \dots, v_m \in \mathbb{R}^n \setminus \{0\}$ demostrar que

$$C = \left\{ \sum_{i=1}^m y_i v_i : y_i \geq 0 \quad \forall i = 1, \dots, m \right\}$$

Es un cono.

Demostración. Un conjunto es un cono si para cualquier vector x en el conjunto y cualquier escalar $\alpha \geq 0$ el vector $x\alpha$ también está en el conjunto. Por lo tanto, sea $x \in C$, veamos que $\forall \alpha \geq 0$, $\alpha x \in C$. Por definición de C , existen escalares $y_1, \dots, y_m \geq 0$ tales que

$$x = \sum_{i=1}^m y_i v_i.$$

Multiplicando ambos lados por $\alpha \geq 0$ obtenemos

$$\alpha x = \sum_{i=1}^m (\alpha y_i) v_i.$$

Dado que $\alpha \geq 0$ y $y_i \geq 0$ para todo i , se sigue que $\alpha y_i \geq 0$ para todo i . Por lo tanto, αx puede ser expresado como una combinación lineal de los vectores v_i con coeficientes no negativos, lo que implica que $\alpha x \in C$. Así, hemos demostrado que C es un cono. \square

Ejercicio 5. Sea $S \subseteq \mathbb{R}^n$ el polar de S definido como $S^\circ = \{p \in \mathbb{R}^n : p^T x \leq 0 \quad \forall x \in S\}$ es un cono convexo y cerrado. Sea $C \subset \mathbb{R}^n$ un cono convexo cerrado. Entonces para $z \in \mathbb{R}^n$, sea $\bar{z} = P_C(z)$ se tiene que $z - \bar{z} \in \{z\}^\perp \cap C^\circ$.

Demostración. Veamos que S° es un cono convexo y cerrado. En efecto, para demostrar que S° es un cono, trivialmente tomemos $p \in S^\circ$, $\alpha \geq 0$ y entonces

$$(\alpha p)^T x = \alpha(p^T x) \leq 0 \quad \forall x \in S,$$

por lo que $\alpha p \in S^\circ$. Para ver que es convexo tomemos dos puntos $x, y \in S^\circ$ y $\lambda \in [0, 1]$ quiero ver que $\lambda x + (1 - \lambda)y \in S^\circ$. En efecto,

$$(\lambda x + (1 - \lambda)y)^T z = \lambda(x^T z) + (1 - \lambda)(y^T z) \leq 0 \quad \forall z \in S,$$

por lo que $\lambda x + (1 - \lambda)y \in S^\circ$. Finalmente, para ver que es cerrado, tomemos una sucesión $\{x_k\} \subset S^\circ$ que converge a $x \in \mathbb{R}^n$. Entonces, para todo $z \in S$

$$x_k^T z \leq 0 \quad \forall k \in \mathbb{N}.$$

Tomando el límite cuando $k \rightarrow \infty$ obtenemos

$$x^T z = \lim_{k \rightarrow \infty} x_k^T z \leq 0 \quad \forall z \in S,$$

por lo que $x \in S^\circ$ y así S° es cerrado.

Para la otra parte tomemos $\bar{z} = P_C(z)$. Entonces, por la propiedad de la proyección se tiene que

$$(z - \bar{z})^T (x - \bar{z}) \leq 0 \quad \forall x \in C.$$

Tomando $x = 0$ (que pertenece a C ya que es un cono) obtenemos

$$(z - \bar{z})^T (-\bar{z}) \leq 0,$$

y tomando $x = 2\bar{z}$ (que también pertenece a C) obtenemos

$$(z - \bar{z})^T (\bar{z}) \leq 0.$$

Esto nos dice que $(z - \bar{z})^T \bar{z} = 0$, es decir, $z - \bar{z} \in \{z\}^\perp$. □

Ejercicio 6. Mostrar con un ejemplo numérico que las condiciones de calidad de Mangasarian-Fromovitz es estrictamente más débil que regularidad.

Demostración. Recordemos que la condición de regularidad nos dice que $\{\nabla h_i(x^*) : i = 1, \dots, m\} \cup \{\nabla g_i(x^*) : i \in A(x^*)\}$ son l.i, mientras que la condición de Mangasarian-Fromovitz que si x^* es solución factible entonces

$$\sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*) = 0, \quad \mu_i \geq 0 \implies \lambda_i = 0, \mu_i = 0 \quad \forall i.$$

Siguiendo estas definiciones consideremos el siguiente problema de optimización sacado del **Ejemplo 7.44** del *Ribeiro*:

$$\begin{aligned} \text{mín } f(x, y) &= x \\ \text{s.a. } g_1(x, y) &= -x^3 + y \leq 0 \\ g_2(x, y) &= -x^3 - y \leq 0. \\ g_3(x, y) &= -x \leq 0. \end{aligned}$$

Es claro que $(0, 0)$ es un punto factible. Veamos que no cumple la condición de regularidad. En efecto,

$$\nabla g_1(0, 0) = (0, 1), \quad \nabla g_2(0, 0) = (0, -1), \quad \nabla g_3(0, 0) = (-1, 0).$$

Claramente estos vectores no son l.i ya que $\nabla g_1(0, 0) = -\nabla g_2(0, 0)$. Ahora veamos que sí cumple la condición de Mangasarian-Fromovitz. En este caso, $h(x, y)$ no existe, por lo que la condición se reduce a ver que

$$\sum_{i \in A(x^*)} \mu_i \nabla g_i(x^*) = 0, \quad \mu_i \geq 0 \implies \mu_i = 0 \quad \forall i.$$

En nuestro caso, $A(x^*) = \{1, 2, 3\}$ y por lo tanto

$$\mu_1 \nabla g_1(0, 0) + \mu_2 \nabla g_2(0, 0) + \mu_3 \nabla g_3(0, 0) = 0.$$

Esto ocurre si $\mu_1 = -\mu_2$ y $\mu_3 = 0$ ya que los vectores $\nabla g_1(0, 0)$ y $\nabla g_2(0, 0)$ son linealmente independientes del vector $\nabla g_3(0, 0)$. Por lo tanto, $\mu_1 \leq 0 \implies \mu_2 \geq 0$ y viceversa, por lo tanto no estamos en las condiciones enunciadas y hemos encontrado un ejemplo donde Mangasarian-Fromovitz se cumple, pero la regularidad no. \square

Ejercicio 7. Mostrar con un ejemplo numérico que la condición de calidad de rango constante es estrictamente más débil que regularidad.

Demostración. Decimos que un punto factible x^* del problema

$$\text{mín } f(x) \text{ sujeto a } h_i(x) = 0, \quad g_i(x) \leq 0$$

cumple con la condición de rango constante si existe un entorno $B(x^*, \epsilon) = U$ tal que para todo

$$I_0 = \{1, \dots, m\} \quad J_0 \subseteq A(x^*)$$

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 $x \in U$. Consideremos el problema con las siguientes restricciones de desigualdad:

$$\begin{aligned} \text{mín } f(x, y) &= x \\ \text{s.a } g_1(x, y) &= x \leq 0 \\ g_2(x, y) &= 2x \leq 0. \end{aligned}$$

Luego, el punto $(0, 0)$ es un punto factible. Veamos que no cumple la condición de regularidad. En efecto,

$$\nabla g_1(0, 0) = (1, 0), \quad \nabla g_2(0, 0) = (2, 0).$$

Claramente estos vectores no son l.i ya que $\nabla g_2(0, 0) = 2\nabla g_1(0, 0)$. Ahora veamos que sí cumple la condición de rango constante. Para todo $\emptyset, \{1\}, \{2\}, \{1, 2\} \subseteq A(x^*)$ el rango de los gradientes es constante en un entorno de $(0, 0)$ ya que en todos los casos el rango es 1. Pues, dado $x \in U$ las matrices serán de la forma:

$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \end{pmatrix}$$

Por lo tanto, hemos encontrado un ejemplo donde la condición de rango constante se cumple, pero la regularidad no. La diferencia entre LICQ y CRCQ radica en que LICQ requiere que los gradientes de la restricciones sean l.i en el punto x^* , mientras que CRCQ admite cierta dependencia lineal entre los gradientes, siempre y cuando la dependencia lineal se mantenga en un entorno de x^* , intuitivamente notamos que CRCQ es una condición más débil. \square

Ejercicio 8. Mostrar con un ejemplo numérico que las condiciones de calidad de rango constante y Mangasarian-Fromovitz no se implican mutuamente.

Demostración. Para ver que Mangasarian-Fromovitz no implica condición de rango constante consideremos de nuevo el ejemplo dado en el ejercicio 6:

$$\begin{aligned} \min f(x, y) &= x \\ \text{s.a. } g_1(x, y) &= -x^3 + y \leq 0 \\ g_2(x, y) &= -x^3 - y \leq 0. \\ g_3(x, y) &= -x \leq 0. \end{aligned}$$

Ya vimos que este ejemplo cumple la condición de Mangasarian-Fromovitz pero no la de regularidad. Veamos que no cumple la condición de rango constante. En efecto,

$$\nabla g_1(x, y) = (-3x^2, 1), \quad \nabla g_2(x, y) = (-3x^2, -1), \quad \nabla g_3(x, y) = (-1, 0).$$

En un entorno de $(0, 0)$ tomando únicamente los gradientes de las restricciones activas g_1 y g_2 el rango es 1 pues son l.d., pero si tomamos g_1 y g_3 el rango es 2 ya que son l.i. Por lo tanto, no existe un entorno donde el rango sea constante y no se cumple la condición de rango constante.

Para ver que la condición de rango constante no implica Mangasarian-Fromovitz consideremos el siguiente problema:

$$\begin{aligned} \min f(x, y) \\ \text{s.a. } g_1(x, y) &= x \leq 0 \\ g_2(x, y) &= -x \leq 0. \end{aligned}$$

Donde ambas restricciones están activas en el punto $(0, 0)$. Es fácil ver que la condición de rango constante se cumple ya que los gradientes de las restricciones son $\nabla g_1(0, 0) = (1, 0)$ y $\nabla g_2(0, 0) = (-1, 0)$ y en un entorno de $(0, 0)$ el rango de los gradientes es siempre 1. Ahora veamos que no se cumple la condición de Mangasarian-Fromovitz. En efecto,

$$\mu_1 \nabla g_1(0, 0) + \mu_2 \nabla g_2(0, 0) = 0,$$

se cumple si $\mu_1 = \mu_2$. Por lo tanto, si tomamos $\mu_1 = \mu_2 = 1$ tenemos una solución no trivial con $\mu_i \geq 0$ para todo i y no se cumple la condición de Mangasarian-Fromovitz. \square

Ejercicio 9. Considere el problema

$$\begin{aligned} \min f(x, y) &= x^2 + 4xy + y^2 \\ \text{s.a. } g(x, y) &= x^2 + y^2 = 1 \end{aligned}$$

- Usando las condiciones KKT, encontrar una solución del problema. Interpretar las condiciones KKT geoméricamente en x^* . ¿Cómo son las curvas de nivel de f ?
- Analizar si se verifican las condiciones de segundo orden.
- ¿Tiene el problema una única solución?

Demostración. Por el **Teorema 4.13** del apunte el minimizador debe satisfacer:

$$\begin{aligned} -\nabla f(x, y) &= \mu \nabla g(x, y) \\ -\begin{pmatrix} 2x + 4y \\ 4x + 2y \end{pmatrix} &= \mu \begin{pmatrix} 2x \\ 2y \end{pmatrix} \\ \begin{pmatrix} -2x - 4y \\ -4x - 2y \end{pmatrix} &= \begin{pmatrix} 2\mu x \\ 2\mu y \end{pmatrix} \end{aligned}$$

Para algún $\mu \geq 0$. Resolviendo el sistema obtenemos que:

$$\begin{aligned} \begin{pmatrix} -2x - 4y - 2\mu x \\ -4x - 2y - 2\mu y \end{pmatrix} = 0 &\implies \begin{pmatrix} x + 2y + \mu x \\ 2x + y + \mu y \end{pmatrix} = 0 \\ (1 + \mu)x + 2y &= 0 \quad (1) \\ 2x + (1 + \mu)y &= 0 \quad (2) \\ \implies (1 + \mu)^2 - 4 &= 0 \implies \mu = 1 \text{ o } \mu = -3. \end{aligned}$$

Como debe ser $\mu \geq 0$ tomamos $\mu = 1$ y de ahí obtenemos que $x = y = 0$ es solución, pero la descartamos pues no es factible. Tomando $\mu = 1$ en (1) obtenemos $2x + 2y = 0 \implies y = -x$. Usando la restricción $g(x, y) = 0$ tenemos que

$$x^2 + (-x)^2 = 1 \implies 2x^2 = 1 \implies x = \pm \frac{1}{\sqrt{2}} \implies y = \mp \frac{1}{\sqrt{2}}.$$

Por lo tanto, las soluciones candidatas son $x_1^* = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ y $x_2^* = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ donde $f(x_1^*) = f(x_2^*)$, por lo que la solución no es única. Geométricamente, las condiciones KKT nos dicen que en el punto x^* , el gradiente de la función objetivo f es proporcional al gradiente de la restricción g . Esto implica que las curvas de nivel de f son tangentes a la curva definida por $g(x, y) = 0$ en el punto x^* . Las curvas de nivel de f son hipérbolas centradas en el origen.

Veamos finalmente si cumple las condiciones de segundo orden:

$$\begin{aligned} \nabla^2 f(x, y) &= \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}, \\ \nabla^2 g(x, y) &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \end{aligned}$$

Luego, la matriz de Lagrange es

$$L(x, y, \mu) = \nabla^2 f(x, y) + \mu \nabla^2 g(x, y) = \begin{pmatrix} 2 + 2\mu & 4 \\ 4 & 2 + 2\mu \end{pmatrix}.$$

Con el mismo $\mu = 1$ que antes y sea $d \in T_{x^*}^{\text{lin}}$ entonces

$$\begin{aligned} d^T L(x^*, y^*, \mu) d &= d^T \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} d \\ &= 4d_1^2 + 8d_1d_2 + 4d_2^2 \\ &= 4(d_1 + d_2)^2 \geq 0. \end{aligned}$$

Por lo tanto, la condición de segundo orden se cumple en x^* . □