Semana6 - Concavidad y Optimización

Abelardo Jordán Liza

Maestría en Matemáticas Aplicadas
PUCP

Lima, Setiembre 19, 2023

Sea $\emptyset \neq C \subset \mathbb{R}^n$ convexo y abierto, f una función de clade C^2 en C. Entonces f es fuertemente convexa de módulo $c > 0$ en C, si y solo si

$$d' Hf(x)d \geq c\|d\|^2, \qquad \forall x \in C, d \in \mathbb{R}^n \tag{1}$$

Sea $\emptyset \neq C \subset \mathbb{R}^n$ convexo y abierto, f una función de clade C^2 en C. Entonces f es fuertemente convexa de módulo $c > 0$ en C, si y solo si

$$d' Hf(x)d \geq c\|d\|^2, \qquad \forall x \in C, d \in \mathbb{R}^n \tag{1}$$

Dem: f es fuertemenete convexa de módulo $c > 0$ en C, si y solo si, la función g definida por $g(x) = f(x) - \frac{c}{2}\|x\|^2$ es convexa en C, donde $Hg(x) = Hf(x) - cI$ la cual debe ser positiva semidefinida. Esto significa

$$d' Hf(x) \geq c\|d\|^2, \qquad \forall d \in \mathbb{R}^n$$

Sea $\emptyset \neq C \subset \mathbb{R}^n$ convexo y abierto, f una función de clade C^2 en C. Entonces f es fuertemente convexa de módulo $c > 0$ en C, si y solo si

$$d' Hf(x)d \geq c\|d\|^2, \qquad \forall x \in C, d \in \mathbb{R}^n \tag{1}$$

Dem: f es fuertemenete convexa de módulo $c > 0$ en C, si y solo si, la función g definida por $g(x) = f(x) - \dfrac{c}{2}\|x\|^2$ es convexa en C, donde $Hg(x) = Hf(x) - cI$ la cual debe ser positiva semidefinida. Esto significa

$$d' Hf(x) \geq c\|d\|^2, \qquad \forall d \in \mathbb{R}^n$$

NOTA:
Sea A simétrica, $b \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, entonces la función $f(x) = \dfrac{1}{2}x'Ax + b'x + \alpha$ es fuertemente convexa, si y solo si A es positiva definida.

Sean A matriz simétrica de orden n, $b \in \mathbb{R}^n, \alpha \in \mathbb{R}$ y el problema de optimización no restringida

$$(P_A) \qquad \begin{array}{ll} \min & f(x) = \frac{1}{2}x'Ax + b'x + \alpha \\ & x \in \mathbb{R}^n \end{array} \qquad (2)$$

Denotemos por
$F^* := argmin\{f(x) : x \in \mathbb{R}^n\} = \{x^* \in \mathbb{R}^n : f(x^*) \leq f(x),\ \forall x \in \mathbb{R}^n\}$ y sea $v(P_A)$ el valor óptimo del problema.

Proposición

Respecto al problema (P_A):

(i) Si A no es positiva semidefinida, entonces $v(P_A) = -\infty$

(ii) Si A es positiva semidefinida, $F^* \neq \emptyset \Leftrightarrow -b \in Im(A)$ y $argmin\{f\} = \{x \in \mathbb{R}^n : Ax + b = 0\}$

(iii) A es positiva definida $\Leftrightarrow F^*$ es un conjunto unitario y $F^* = \{-A^{-1}(b).\}$

Respecto al problema (P_A):

(i) Si A no es positiva semidefinida, entonces $v(P_A) = -\infty$

(ii) Si A es positiva semidefinida, $F^* \neq \emptyset \Leftrightarrow -b \in Im(A)$ y $argmin\{f\} = \{x \in \mathbb{R}^n : Ax + b = 0\}$

(iii) A es positiva definida $\Leftrightarrow F^*$ es un conjunto unitario y $F^* = \{-A^{-1}(b).\}$

NOTA:
Sea g una función de valorf real y creciente(estricta) en el rango de la función f definida en $C \subset \mathbb{R}^n$, entonces $x^* \in C$ resuelve el problema $\begin{array}{cc} \text{mín} & f(x) \\ & x \in C \end{array}$ si y solo si x^* resuelve el problema $\begin{array}{cc} \text{mín} & g(f(x)) \\ & x \in C \end{array}$

Regresión de mínimos cuadrados:

Regresión de mínimos cuadrados:

Problema básico:

Dada una muestra de observaciones $\{(t_i, s_i) \in \mathbb{R} \times \mathbb{R} : i = 1, \cdots, m\}$ y se asume que s depende linealmente de t. Se desea encontrar dos números reales x_1 y x_2 tales que al construir los valores residuales $r_i = x_1 + x_2 t_i - s_i$, aún más, el vector residual $r = (r_1, \cdots, r_m)$ tenga norma mínima. Se genera el problema de optimización

$$\begin{array}{ll} \min & \|r\| \\ & x = (x_1, x_2) \in \mathbb{R}^2 \end{array}$$

Regresión de mínimos cuadrados:
Problema básico:
Dada una muestra de observaciones $\{(t_i, s_i) \in \mathbb{R} \times \mathbb{R} : i = 1, \cdots, m\}$ y se asume que s depende linealmente de t. Se desea encontrar dos números reales x_1 y x_2 tales que al construir los valores residuales $r_i = x_1 + x_2 t_i - s_i$, aún más, el vector residual $r = (r_1, \cdots, r_m)$ tenga norma mínima. Se genera el problema de optimización

$$\begin{array}{ll} \text{mín} & \|r\| \\ & x = (x_1, x_2) \in \mathbb{R}^2 \end{array}$$

por la Nota anterior, este problema equivale a resolver

$$\begin{array}{ll} \text{mín} & \|r\|^2 \\ & x = (x_1, x_2) \in \mathbb{R}^2 \end{array}$$

¿Cuál es la ventaja?

Problema general:

Una variable escalar y se va a representar como una función lineal de la forma $y = \sum_{i=1}^{n} x_i u_i$ para ciertos coeficientes $x_1, \cdots, x_n$ desconocidos.

Se tiene N observaciones input-output $(u^j, y^j) \in \mathbb{R}^n \times \mathbb{R}$ para $j = 1, \cdots, N$. Para $x = (x_1, \cdots, x_n)'$, dados los valores residuales $r_j := x'u^j - y^j$, se propone minimizar la norma de $r = (r_1, \cdots, r_N)'$. Empleando la norma euclidiana, se genera el problema

$$\begin{array}{ll} \min & \sqrt{\sum_{j=1}^{N} (y^j - \sum_{i=1}^{n} x_i u_i^j)^2} \\ & x = (x_1, \cdots, x_n)' \in \mathbb{R}^n \end{array} \tag{3}$$

Sea U la matriz de orden $N \times n$ cuyas filas son los vectores u^j, el problema equivale a

$$\begin{array}{ll} \min & f(x) = \|y - Ux\|^2 = x'U'Ux - 2y'Ux + \|y\|^2 \\ & x = (x_1, \cdots, x_n)' \in \mathbb{R}^n \end{array} \tag{4}$$

Por ser f convexa y diferenciable, aplicamos las condiciones necesarias y suficientes de optimlaidad, por lo que se establece $\nabla f(x) = 2U'Ux - 2U'y = 0$.

Problema general:

Una variable escalar y se va a representar como una función lineal de la forma $y = \sum_{i=1}^{n} x_i u_i$ para ciertos coeficientes $x_1, \cdots, x_n$ desconocidos.

Se tiene N observaciones input-output $(u^j, y^j) \in \mathbb{R}^n \times \mathbb{R}$ para $j = 1, \cdots, N$. Para $x = (x_1, \cdots, x_n)'$, dados los valores residuales $r_j := x'u^j - y^j$, se propone minimizar la norma de $r = (r_1, \cdots, r_N)'$. Empleando la norma euclidiana, se genera el problema

$$\begin{array}{ll} \min & \sqrt{\sum_{j=1}^{N}(y^j - \sum_{i=1}^{n} x_i u_i^j)^2} \\ x = (x_1, \cdots, x_n)' \in \mathbb{R}^n & \end{array} \tag{3}$$

Sea U la matriz de orden $N \times n$ cuyas filas son los vectores u^j, el problema equivale a

$$\begin{array}{ll} \min & f(x) = \|y - Ux\|^2 = x'U'Ux - 2y'Ux + \|y\|^2 \\ x = (x_1, \cdots, x_n)' \in \mathbb{R}^n & \end{array} \tag{4}$$

Por ser f convexa y diferenciable, aplicamos las condiciones necesarias y suficientes de optimlaidad, por lo que se establece $\nabla f(x) = 2U'Ux - 2U'y = 0$. Si U tiene rango n, entonces $x^* = (U'U)^{-1}U'y$ resuelve el problema.

Problema general:
Una variable escalar y se va a representar como una función lineal de la forma $y = \sum_{i=1}^{n} x_i u_i$ para ciertos coeficientes $x_1, \cdots, x_n$ desconocidos.
Se tiene N observaciones input-output $(u^j, y^j) \in \mathbb{R}^n \times \mathbb{R}$ para $j = 1, \cdots, N$.
Para $x = (x_1, \cdots, x_n)'$, dados los valores residuales $r_j := x'u^j - y^j$, se propone minimizar la norma de $r = (r_1, \cdots, r_N)'$. Empleando la norma euclidiana, se genera el problema

$$\begin{array}{ll} \text{mín} & \sqrt{\sum_{j=1}^{N} (y^j - \sum_{i=1}^{n} x_i u_i^j)^2} \\ & x = (x_1, \cdots, x_n)' \in \mathbb{R}^n \end{array} \tag{3}$$

Sea U la matriz de orden $N \times n$ cuyas filas son los vectores u^j, el problema equivale a

$$\begin{array}{ll} \text{mín} & f(x) = \|y - Ux\|^2 = x'U'Ux - 2y'Ux + \|y\|^2 \\ & x = (x_1, \cdots, x_n)' \in \mathbb{R}^n \end{array} \tag{4}$$

Por ser f convexa y diferenciable, aplicamos las condiciones necesarias y suficientes de optimlaidad, por lo que se establece $\nabla f(x) = 2U'Ux - 2U'y = 0$.
Si U tiene rango n, entonces $x^* = (U'U)^{-1}U'y$ resuelve el problema.
¿Qué ocurre si U tiene rango menor que n ?

Considere el sistema lineal $Ax = b$ dende A es una matriz de orden $m \times n$ y $b \in \mathbb{R}^m$ y se asume que $m > n$. Se trata de encontrar $x \in \mathbb{R}^n$ tal que minimice $\|Ax - b\|$. Esto da lugar al problema

$$\begin{array}{ll} \text{mín} & \|Ax - b\|^2 \\ & x = (x_1, \cdots, x_n)' \in \mathbb{R}^n \end{array} \tag{5}$$

Siguiendo los procedmientos análogos, la solución x^* debe satisfacer $A'Ax^* = A'b$. Si A tiene rango n, entonces la solución es única.

Considere el sistema lineal $Ax = b$ dende A es una matriz de orden $m \times n$ y $b \in \mathbb{R}^m$ y se asume que $m > n$. Se trata de encontrar $x \in \mathbb{R}^n$ tal que minimice $\|Ax - b\|$. Esto da lugar al problema

$$\begin{array}{ll} \text{mín} & \|Ax - b\|^2 \\ & x = (x_1, \cdots, x_n)' \in \mathbb{R}^n \end{array} \tag{5}$$

Siguiendo los procedmientos análogos, la solución x^* debe satisfacer $A'Ax^* = A'b$. Si A tiene rango n, entonces la solución es única. Considerando la factorización QR de A de rango n. Existen matrices Q de orden m ortogonal y R triangular superior e invertible de orden n tal que $A = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$, se requiere minimizar

$\|Ax - b\| = \|Q(\begin{array}{c} R \\ 0 \end{array})x - b\| = \|Q'b - (\begin{array}{c} R \\ 0 \end{array})x\|$ Escribiendo $Q'b = (\begin{array}{c} c \\ d \end{array})$ con c de n componentes, entonces

$$\|r\|^2 = \|c - Rx\|^2 + \|d\|^2$$

Minimizar $\|r\|^2$ se reduce a resolver $Rx = c$.

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{6}$$

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{6}$$

El conjunto de las direcciones tangente a X en el punto x se denota por $T_X(x)$.

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{6}$$

El conjunto de las direcciones tangente a X en el punto x se denota por $T_X(x)$.
Equivalentemente, $d \in T_X(x)$ si y solo si:
existe una sucesión d^k que converge a d y una sucesión de números reales positivos t_k con $t_k \downarrow 0$ y $x + t_k d^k \in X$.

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{6}$$

El conjunto de las direcciones tangente a X en el punto x se denota por $T_X(x)$.
Equivalentemente, $d \in T_X(x)$ si y solo si:
existe una sucesión d^k que converge a d y una sucesión de números reales positivos t_k con $t_k \downarrow 0$ y $x + t_k d^k \in X$.

Proposición

$T_X(x)$ es un cono cerrado.

Definición

El cono de direcciones factibles en $x \in X$ está dado por

$$K_X(x) := \{d \in \mathbb{R}^n : d = \alpha(y - x), y \in X, \alpha \geq 0\}$$

Definición

El cono de direcciones factibles en $x \in X$ está dado por

$$K_X(x) := \{d \in \mathbb{R}^n : d = \alpha(y - x), y \in X, \alpha \geq 0\}$$

Proposición

Sea X un subconjunto convexo no vacío de $\mathbb{R}^n$ y $x \in X$, entonces

$$T_X(x) = \overline{K_X(x)} \tag{7}$$

Definición

El cono de direcciones factibles en $x \in X$ está dado por

$$K_X(x) := \{d \in \mathbb{R}^n : d = \alpha(y - x), y \in X, \alpha \geq 0\}$$

Proposición

Sea X un subconjunto convexo no vacío de $\mathbb{R}^n$ y $x \in X$, entonces

$$T_X(x) = \overline{K_X(x)} \tag{7}$$

Dem: Sea $d \in K_X(x)$, si $d = 0$ obviamente $d \in T_X(x)$, si $d \neq 0$, entonces existen $y \in X, \alpha > 0$ tales que $d = \alpha(y - x)$. Sean $x^k := x + \frac{1}{k} d \in X$ y cumple $\dfrac{x^k - x}{\frac{1}{k}} = d$, entonces $d \in T_X(x)$. Lo que significa que $K_X(x) \subset T_X(x)$ y así $\overline{K_X(x)} \subset T_X(x)$.
La otra inclusión (ejercicio).

Sea X un subconjunto no vacío de $\mathbb{R}^n$ y f una función de valor real definida en un conjunto del cual X es un subconjunto.
Se generan los problemas de optimización restringida:

$$\begin{array}{ll} \text{mín} & f(x) \\ s.a. & x \in X \end{array} \qquad\qquad \begin{array}{ll} \text{máx} & f(x) \\ s.a. & x \in X \end{array} \tag{8}$$

Sea X un subconjunto no vacío de $\mathbb{R}^n$ y f una función de valor real definida en un conjunto del cual X es un subconjunto.
Se generan los problemas de optimización restringida:

$$\begin{array}{ll} \min & f(x) \\ s.a. & x \in X \end{array} \qquad\qquad \begin{array}{ll} \max & f(x) \\ s.a. & x \in X \end{array} \tag{8}$$

Enseguida nos enfocamos al problema (P) : $\begin{array}{ll} \min & f(x) \\ s.a. & x \in X \end{array}$

¿Cómo se describe X ?
X puede tener una descripción funcional como for ejemplo es el conjunto solución de un sistema de ecuaciones o inecuaciones que involucren a las variables de decisión.

Teorema: Condición necesaria de optimalidad

Sea $X \subset \mathbb{R}^n$ no vacío y f continuamente diferenciable en $\overline{x} \in X$. Considere el problema

$$\begin{array}{ll} \text{mín} & f(x) \\ s.a. & x \in X \end{array} \tag{9}$$

Si $\overline{x}$ resuelve (9) entonces

$$-\nabla f(\overline{x}) \in (T_X(\overline{x}))^o \tag{10}$$

Recíprocamente, si X es convexo y la función f es convexa en X, y $\overline{x}$ satisface (10), entonces $\overline{x}$ resuelve (9).

Sea $X \subset \mathbb{R}^n$ no vacío y f continuamente diferenciable en $\overline{x} \in X$. Considere el problema

$$\begin{array}{ll} \text{mín} & f(x) \\ s.a. & x \in X \end{array} \tag{9}$$

Si $\overline{x}$ resuelve (9) entonces

$$-\nabla f(\overline{x}) \in (T_X(\overline{x}))^o \tag{10}$$

Recíprocamente, si X es convexo y la función f es convexa en X, y $\overline{x}$ satisface (10), entonces $\overline{x}$ resuelve (9).

Dem:(Por el absurdo) Suponga que $-\nabla f(\overline{x}) \notin (T_X(\overline{x}))^o$, entonces existe $d \in T_X(\bar{(}x))$ tal que $\langle \nabla f(\overline{x}), d\rangle < 0$. Para d, existe una sucesión x^k en X (convergente a $\overline{x}$) y una sucesión $t_k \downarrow 0$ tales que $\lim\limits_{k\to\infty} \dfrac{x^k - \overline{x}}{t_k} = d$. Para cada x^k, existe $a_k \in \mathbb{R}$ tal que

$$f(x^k) - f(\overline{x}) = \langle \nabla f(\overline{x}), x^k - \overline{x}\rangle + a_k$$

donde $\dfrac{a_k}{\|x^k - \overline{x}\|} \to 0$ y además $\lim\limits_{k\to\infty} \dfrac{a_k}{t_k} = \lim\limits_{k\to\infty} \dfrac{a_k\|d\|}{\|x^k - \overline{x}\|} = 0.$

Sea $X \subset \mathbb{R}^n$ no vacío y f continuamente diferenciable en $\overline{x} \in X$. Considere el problema

$$\begin{array}{ll} \text{mín} & f(x) \\ s.a. & x \in X \end{array} \tag{9}$$

Si $\overline{x}$ resuelve (9) entonces

$$-\nabla f(\overline{x}) \in (T_X(\overline{x}))^o \tag{10}$$

Recíprocamente, si X es convexo y la función f es convexa en X, y $\overline{x}$ satisface (10), entonces $\overline{x}$ resuelve (9).

Dem:(Por el absurdo) Suponga que $-\nabla f(\overline{x}) \notin (T_X(\overline{x}))^o$, entonces existe $d \in T_X(\bar{(x)})$ tal que $\langle \nabla f(\overline{x}), d\rangle < 0$. Para d, existe una sucesión x^k en X (convergente a $\overline{x}$) y una sucesión $t_k \downarrow 0$ tales que $\lim_{k\to\infty} \dfrac{x^k - \overline{x}}{t_k} = d$. Para cada x^k, existe $a_k \in \mathbb{R}$ tal que

$$f(x^k) - f(\overline{x}) = \langle \nabla f(\overline{x}), x^k - \overline{x}\rangle + a_k$$

donde $\dfrac{a_k}{\|x^k - \overline{x}\|} \to 0$ y además $\lim_{k\to\infty} \dfrac{a_k}{t_k} = \lim_{k\to\infty} \dfrac{a_k\|d\|}{\|x^k - \overline{x}\|} = 0$. Siendo

$$\frac{f(x^k) - f(\overline{x})}{t_k} = \langle \nabla f(\overline{x}), d\rangle + \langle \nabla f(\overline{x}), \frac{x^k - \overline{x}}{t_k} - d\rangle + \frac{a_k}{t_k}$$

se sigue que $\lim_{k\to\infty} \dfrac{f(x^k) - f(\overline{x})}{t_k} < 0$ y esto genera una contradicción.

Recíprocamente, para cada $y \in X$, $d := y - \overline{x} \in T_X(\overline{x})$
entonces $\langle \nabla f(\overline{x}), y - \overline{x} \rangle \geq 0$, y por la convexidad de f es sabido que

$$f(y) \geq f(\overline{x}) + \langle \nabla f(\overline{x}), y - \overline{x} \rangle$$

en consecuencia $f(y) \geq f(\overline{x})$, $\forall y \in X$.

Recíprocamente, para cada $y \in X$, $d := y - \overline{x} \in T_X(\overline{x})$ entonces $\langle \nabla f(\overline{x}), y - \overline{x} \rangle \geq 0$, y por la convexidad de f es sabido que

$$f(y) \geq f(\overline{x}) + \langle \nabla f(\overline{x}), y - \overline{x} \rangle$$

en consecuencia $f(y) \geq f(\overline{x})$, $\forall y \in X$.

¿Qué ocurre con (10) si $\overline{x}$ es un punto interior de X ?

X como el conjunto solución de un sistema de inecuaciones/ecuaciones algebraicas en las variables de decisión.

Sean $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \cdots, p$ funciones convexas y $C \subset \mathbb{R}^n$ un subconjunto convexo. Considere los siguientes sistemas :

(I) $f_1(x) < 0, \cdots, f_p(x) < 0,\ x \in C$;

(II) $\exists 0 \neq \lambda = (\lambda_1, \cdots, \lambda_p) \in \mathbb{R}^p_+$ tal que $\displaystyle\sum_{i=1}^{p} \lambda_i f_i(x) \geq 0, \forall x \in C.$

Entonces, uno y solo uno de los sistemas tiene solución y no ambos.

Sean $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \cdots, p$ funciones convexas y $C \subset \mathbb{R}^n$ un subconjunto convexo. Considere los siguientes sistemas :

(I) $f_1(x) < 0, \cdots, f_p(x) < 0,\ x \in C$;

(II) $\exists 0 \neq \lambda = (\lambda_1, \cdots, \lambda_p) \in \mathbb{R}^p_+$ tal que $\displaystyle\sum_{i=1}^{p} \lambda_i f_i(x) \geq 0, \forall x \in C.$

Entonces, uno y solo uno de los sistemas tiene solución y no ambos.

Esto significa: Si (I) no tiene solución, entonces (II) tiene solución, y si (II) no tiene solución entonces (I) tiene solución.

Considere el problema

$$(Pmin) \qquad \begin{array}{ll} \text{mín} & f(x) \\ s.a. & g_i(x) \leq 0, \; i = 1, \cdots, m. \end{array} \tag{11}$$

donde $f, g_i, i = 1, \cdots, m$ son funciones diferenciables en $\mathbb{R}^n$. Si x^* resuelve (Pmin), entonces existen escalares $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos cero, tales que

$$\begin{array}{l} (i) \; \lambda_0 \nabla f(x^*) + \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(x*) = 0 \\ (ii) \; \lambda_i g_i(x^*) = 0, \; i = 1, \cdots, m. \end{array} \tag{12}$$

Considere el problema

$$(Pmin) \qquad \begin{array}{ll} \text{mín} & f(x) \\ s.a. & g_i(x) \leq 0, \; i = 1, \cdots, m. \end{array} \tag{11}$$

donde $f, g_i, i = 1, \cdots, m$ son funciones diferenciables en $\mathbb{R}^n$. Si x^* resuelve (Pmin), entonces existen escalares $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos cero, tales que

$$\begin{array}{l} (i) \; \lambda_0 \nabla f(x^*) + \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(x*) = 0 \\ (ii) \; \lambda_i g_i(x^*) = 0, \; i = 1, \cdots, m. \end{array} \tag{12}$$

Dem: Considere el sistema de inecuaciones en la variable $v \in \mathbb{R}^n$:

$$(I) : \langle \nabla f(x^*), v \rangle < 0 \;\; ; \;\; g_i(x^*) + \langle \nabla g_i(x^*), v \rangle < 0, \; i = 1, \cdots, m$$

se prueba que este sistema no tiene solución en $\mathbb{R}^n$. Pues, caso contrario existe $v \in \mathbb{R}^n$ tal que satisface (I), se prueba que para $t > 0$ suficientemente pequeño : $g_i(x^* + tv) < 0$ lo que significa que $x^* + tv$ es factible para (Pmin) y de $\langle \nabla f(x^*), v \rangle < 0$ se sigue que $f(x^* + tv) < f(x^*)$ lo que significa una contradicción.

Por el teorema de la alternativa, existen $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos nulos tales que

$$\lambda_0 \langle \nabla f(x^*), v \rangle + \sum_{i=1}^{m} \lambda_i (g_i(x^*) + \langle \nabla g_i(x^*), v \rangle) \geq 0, \ \forall v \in \mathbb{R}^n \qquad (13)$$

En particular para $v = 0$ se obtiene $\sum_{i=1}^{m} \lambda_i (g_i(x^*) \geq 0$ y así se concluye con $\lambda_i g_i(x^*) = 0$ para cada $i = 1, \cdots, m$.

Por el teorema de la alternativa, existen $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos nulos tales que

$$\lambda_0 \langle \nabla f(x^*), v \rangle + \sum_{i=1}^{m} \lambda_i (g_i(x^*) + \langle \nabla g_i(x^*), v \rangle) \geq 0, \ \forall v \in \mathbb{R}^n \tag{13}$$

En particular para $v = 0$ se obtiene $\sum_{i=1}^{m} \lambda_i (g_i(x^*) \geq 0$ y así se concluye con $\lambda_i g_i(x^*) = 0$ para cada $i = 1, \cdots, m$. De este modo (13) se reduce a

$$\lambda_0 \langle \nabla f(x^*), v \rangle + \sum_{i=1}^{m} \lambda_i \langle \nabla g_i(x^*), v \rangle \geq 0, \ \forall v \in \mathbb{R}^n$$

lo que da lugar a $\quad \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0$

Por el teorema de la alternativa, existen $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos nulos tales que

$$\lambda_0 \langle \nabla f(x^*), v \rangle + \sum_{i=1}^{m} \lambda_i (g_i(x^*) + \langle \nabla g_i(x^*), v \rangle) \geq 0, \ \forall v \in \mathbb{R}^n \qquad (13)$$

En particular para $v = 0$ se obtiene $\sum_{i=1}^{m} \lambda_i (g_i(x^*) \geq 0$ y así se concluye con $\lambda_i g_i(x^*) = 0$ para cada $i = 1, \cdots, m$. De este modo (13) se reduce a

$$\lambda_0 \langle \nabla f(x^*), v \rangle + \sum_{i=1}^{m} \lambda_i \langle \nabla g_i(x^*), v \rangle \geq 0, \ \forall v \in \mathbb{R}^n$$

lo que da lugar a $\quad \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0$

Note que si $\lambda_0 \neq 0$, entonces podemos sostener que $\cdots$

Considere el problema

$$(P) \qquad \begin{array}{ll} \text{mín} & f(x) \\ s.a. & g_i(x) \leq 0,\ i = 1, \cdots, m. \\ & x \in C \end{array} \tag{14}$$

donde $f, g_i, i = 1, \cdots, m$ son funciones diferenciables en $\mathbb{R}^n$ y C es un subconjunto convexo no vacío de $\mathbb{R}^n$. Si x^* resuelve (P), entonces existen escalares $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos cero, tales que

$$\begin{array}{l} (i) 0 \in \ \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x*) + N_C(x^*) \\ (ii)\ \lambda_i g_i(x^*) = 0,\ i = 1, \cdots, m. \end{array} \tag{15}$$

Considere el problema

$$
(P) \qquad \begin{array}{ll} \text{mín} & f(x) \\ s.a. & g_i(x) \leq 0, \ i = 1, \cdots, m. \\ & x \in C \end{array} \tag{14}
$$

donde $f, g_i, i = 1, \cdots, m$ son funciones diferenciables en $\mathbb{R}^n$ y C es un subconjunto convexo no vacío de $\mathbb{R}^n$. Si x^* resuelve (P), entonces existen escalares $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos cero, tales que

$$
\begin{array}{l} (i) 0 \in \ \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x*) + N_C(x^*) \\ (ii) \ \lambda_i g_i(x^*) = 0, \ i = 1, \cdots, m. \end{array} \tag{15}
$$

Dem: Análogo a la demostración de la proposición anterior, probando que el sistema

$$
(I) : \langle \nabla f(x^*, v) \rangle < 0 \ \ ; \ \ g_i(x^*) + \langle \nabla g_i(x^*), v \rangle < 0, \ i = 1, \cdots, m \tag{16}
$$

no tiene solución para $v \in T_C(x^*)$.

Considere el problema

$$
(P) \qquad \begin{array}{ll} \text{mín} & f(x) \\ s.a. & g_i(x) \leq 0, \ i = 1, \cdots, m. \\ & x \in C \end{array} \tag{14}
$$

donde $f, g_i, i = 1, \cdots, m$ son funciones diferenciables en $\mathbb{R}^n$ y C es un subconjunto convexo no vacío de $\mathbb{R}^n$. Si x^* resuelve (P), entonces existen escalares $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos cero, tales que

$$
\begin{array}{l} (i) 0 \in \ \lambda_0 \nabla f(x^*) + \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(x*) + N_C(x^*) \\ (ii) \ \lambda_i g_i(x^*) = 0, \ i = 1, \cdots, m. \end{array} \tag{15}
$$

Dem: Análogo a la demostración de la proposición anterior, probando que el sistema

$$
(I) : \langle \nabla f(x^*, v) \rangle < 0 \ \ ; \ \ g_i(x^*) + \langle \nabla g_i(x^*), v \rangle < 0, \ i = 1, \cdots, m \tag{16}
$$

no tiene solución para $v \in T_C(x^*)$.Aplicando el teorema de la alternativa, existen $\lambda_0, \lambda_1, \cdots, \lambda_m$ no negativos y no todos cero, tales que

$$\lambda_0\langle\nabla f(x^*),v\rangle+\sum_{i=1}^{m}\lambda_i(g_i(x^*)+\langle\nabla g_i(x^*),v\rangle)\geq 0,\ \forall v\in T_C(x^*) \tag{17}$$

lo que indica que

$$-(\lambda_0\nabla f(x^*)+\sum_{i=1}^{m}\lambda_i\nabla g_i(x^*))\in(T_C(x^*))^\circ$$

Para el problema (P), denotemos por $I(x^*)$ al conjunto $\{i \in \{1, \cdots, m\} : g_i(x^*) = 0\}$ llamado el conjunto de los índices de las restricciones activas en x^*.

Corolario

Si en el problema (P), $x^* \in int(C)$ y los vectores $\{\nabla g_i(x^*)\}_{i \in I(x^*)}$ son l.i. Entonces existen $\lambda_1, \cdots, \lambda_m \geq 0$ tales que

$$\nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0$$
$$\lambda_i g_i(x^*) = 0, \; i = 1, \cdots, m.$$