Concavidad y Optimización
Continuación ejemplos y Lagrangiana

Abelardo Jordán Liza

Maestría en Matemáticas Aplicadas
PUCP

Setiembre 28, 2023

Supongamos $k+m$ pacientes que recibieron un tratamiento especial, de modo que se les ha evaluado en n características de desempeño(post tratamiento). Sean

$A = \{a^1, \cdots, a^k\} \subset \mathbb{R}^n$	en k pacientes el tratamiento fue efectivo
$B = \{b^1, \cdots, b^m\} \subset \mathbb{R}^n$	en m pacientes no respondió el tratamiento

Objetivo: Encontrar un clasificador escalar de las clases A y B. Podemos pensar en una función $\varphi : \mathbb{R}^n \to \mathbb{R}$ tal que

$$\varphi(\mathbf{a^i}) > \mathbf{0} \text{ para } i = 1, \cdots, k \qquad \text{y} \qquad \varphi(\mathbf{b^j}) < \mathbf{0} \text{ para } j = 1, \cdots, m.$$

Un problema básico de clasificación

Supongamos $k+m$ pacientes que recibieron un tratamiento especial, de modo que se les ha evaluado en n características de desempeño(post tratamiento). Sean

$$A = \{a^1, \cdots, a^k\} \subset \mathbb{R}^n \quad \text{en } k \text{ pacientes el tratamiento fue efectivo}$$
$$B = \{b^1, \cdots, b^m\} \subset \mathbb{R}^n \quad \text{en } m \text{ pacientes no respondió el tratamiento}$$

Objetivo: Encontrar un clasificador escalar de las clases A y B. Podemos pensar en una función $\varphi : \mathbb{R}^n \to \mathbb{R}$ tal que

$$\varphi(\mathbf{a^i}) > \mathbf{0} \text{ para } i = 1, \cdots, k \qquad \text{y} \qquad \varphi(\mathbf{b^j}) < \mathbf{0} \text{ para } j = 1, \cdots, m.$$

$\underline{Classificador\ lineal\ afín}$

$$\varphi(x) = v'x - \alpha$$

Se requiere conocer v y α.

Supongamos $k + m$ pacientes que recibieron un tratamiento especial, de modo que se les ha evaluado en n características de desempeño(post tratamiento). Sean

$$A = \{a^1, \cdots, a^k\} \subset \mathbb{R}^n \quad \text{en } k \text{ pacientes el tratamiento fue efectivo}$$
$$B = \{b^1, \cdots, b^m\} \subset \mathbb{R}^n \quad \text{en } m \text{ pacientes no respondió el tratamiento}$$

Objetivo: Encontrar un clasificador escalar de las clases A y B. Podemos pensar en una función $\varphi : \mathbb{R}^n \to \mathbb{R}$ tal que

$$\varphi(\mathbf{a^i}) > \mathbf{0} \text{ para } i = 1, \cdots, k \qquad \text{y} \qquad \varphi(\mathbf{b^j}) < \mathbf{0} \text{ para } j = 1, \cdots, m.$$

$\underline{Classificador\ lineal\ afín}$

$$\varphi(x) = v'x - \alpha$$

Se requiere conocer v y α.
¿ Cuál es el criterio para un "mejor " clasificador ?

Supongamos $k+m$ pacientes que recibieron un tratamiento especial, de modo que se les ha evaluado en n características de desempeño(post tratamiento). Sean

$$\begin{array}{ll} A = \{a^1, \cdots, a^k\} \subset \mathbb{R}^n & \text{en } k \text{ pacientes el tratamiento fue efectivo} \\ B = \{b^1, \cdots, b^m\} \subset \mathbb{R}^n & \text{en } m \text{ pacientes no respondió el tratamiento} \end{array}$$

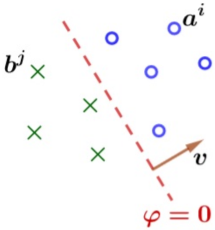
Objetivo: Encontrar un clasificador escalar de las clases A y B. Podemos pensar en una función $\varphi : \mathbb{R}^n \to \mathbb{R}$ tal que

$$\varphi(\mathbf{a^i}) > \mathbf{0} \text{ para } i = 1, \cdots, k \qquad \text{y} \qquad \varphi(\mathbf{b^j}) < \mathbf{0} \text{ para } j = 1, \cdots, m.$$

<u>*Classificador lineal afín*</u>

$$\varphi(x) = v'x - \alpha$$

Se requiere conocer v y α.
¿ Cuál es el criterio para un “mejor ” clasificador ?
¿Esto implica condiciones que involucran a v y α?



- Considerando las calificaciones que otorga φ tanto a los a^i como a los b^j. Se definen los correspondientes errores:

$$e_-(a^i) := \begin{cases} 0 & \text{si } \varphi(a^i) \geq 0 \\ \alpha - v'a^i & \text{si } \varphi(a^i) < 0 \end{cases} \quad \text{y} \quad e_+(b^j) := \begin{cases} 0 & \text{si } \varphi(b^j) \leq 0 \\ v'b^j - \alpha & \text{si } \varphi(b^j) > 0 \end{cases}$$

Podemos formular el problema de optimización

$$\min_{v,\alpha} \quad \sum_{i=1}^{k} e_-(a^i) + \sum_{j=1}^{m} e_+(b^j) \qquad \Big(e_-(a^i) = \max\{0, -\varphi(a^i)\}\Big) \qquad (1)$$

Note que la función objetivo es no-negativa y lineal por partes. Además, para $v = 0$ y $\alpha = 0$, ésta es identicamente nula para los a^i y los b^j y así la función objetivo resulta nula y φ no sería útil. Podemos asumir que $v \neq 0$ y agregar la restricción $\|v\| = 1$ al problema no-lineal (1)

- Considerando las calificaciones que otorga φ tanto a los a^i como a los b^j. Se definen los correspondientes errores:

$$e_-(a^i) := \begin{cases} 0 & \text{si } \varphi(a^i) \geq 0 \\ \alpha - v'a^i & \text{si } \varphi(a^i) < 0 \end{cases} \quad \text{y} \quad e_+(b^j) := \begin{cases} 0 & \text{si } \varphi(b^j) \leq 0 \\ v'b^j - \alpha & \text{si } \varphi(b^j) > 0 \end{cases}$$

Podemos formular el problema de optimización

$$\min_{v,\alpha} \quad \sum_{i=1}^{k} e_-(a^i) + \sum_{j=1}^{m} e_+(b^j) \qquad \left(e_-(a^i) = \max\{0, -\varphi(a^i)\}\right) \qquad (1)$$

Note que la función objetivo es no-negativa y lineal por partes. Además, para $v = 0$ y $\alpha = 0$, ésta es identicamente nula para los a^i y los b^j y así la función objetivo resulta nula y φ no sería útil. Podemos asumir que $v \neq 0$ y agregar la restricción $\|v\| = 1$ al problema no-lineal (1)
Hacemos las asignaciones $s_i = e_-(a^i)$ y $z_j = e_+(b^j)$, reformulamos el problema (1) por el problema

$$\begin{array}{lll} \displaystyle\min_{v,\alpha,s,z} & \displaystyle\sum_{i=1}^{k} s_i + \sum_{j=1}^{m} z_j & \\ s.a. & v'a^i - \alpha + s_i \geq 0 & i = 1, \cdots, k. \\ & v'b^j - \alpha + z_j \leq 0 & j = 1, \cdots, m. \\ & s \geq 0, z \geq 0 & \\ & v'v = 1 & \end{array} \qquad (2)$$

Otra manera de empoderar al discriminador(eliminando la restricción sobre $\|v\|$), es crear una zona de amortiguamiento entre los conjuntos A y B, esto se concretiza modificando las restricciones lineales de desigualdad:

$$\begin{array}{ll} v'a^i - \alpha + s_i \geq 1\,, & i = 1, \cdots, k. \\ v'b^j - \alpha + z_j \leq -1\,, & j = 1, \cdots, m. \end{array}$$

y para preveer que $\|v\|$ no sea bastante grande (caso contrario la zona de amortiguamiento resultaría muy pequeña), agregamos un término en la función objetivo que permita controlar a $\|v\|$.

Otra manera de empoderar al discriminador(eliminando la restricción sobre $\|v\|$), es crear una zona de amortiguamiento entre los conjuntos A y B, esto se concretiza modificando las restricciones lineales de desigualdad:

$$
\begin{array}{ll}
v'a^i - \alpha + s_i \geq 1\,, & i = 1, \cdots, k. \\
v'b^j - \alpha + z_j \leq -1\,, & j = 1, \cdots, m.
\end{array}
$$

y para preveer que $\|v\|$ no sea bastante grande (caso contrario la zona de amortiguamiento resultaría muy pequeña), agregamos un término en la función objetivo que permita controlar a $\|v\|$.
Generamos el problema, con $\delta > 0$ dado:

$$
\begin{array}{lll}
\min\limits_{v,\alpha,s,z} & \displaystyle\sum_{i=1}^{k} s_i + \sum_{j=1}^{m} z_j + \delta\|v\|^2 & \\
s.a. & v'a^i - \alpha + s_i \geq 1\,, & i = 1, \cdots, k. \\
 & v'b^j - \alpha + z_j \leq -1\,, & j = 1, \cdots, m. \\
 & s \geq 0\,, z \geq 0 &
\end{array}
\tag{3}
$$

Sean f, g_i, h_j funciones definidas en $\mathbb{R}^n$ y C un conjunto cerrado de $\mathbb{R}^n$. Para el problema

$$(ProMix) \qquad \begin{array}{ll} \text{mín} & f(x) \\ s.a. & g_i(x) \leq 0, \; i = 1, \cdots, m. \\ & h_j(x) = 0, \; j = 1, \cdots, p. \\ & x \in C \end{array} \tag{4}$$

definimos la función Lagrangiana

$$L(x, \lambda, \mu) := f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x) \tag{5}$$

con $x \in C, (\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p =: \Lambda_0$.

¿Cómo se usa la Función Lagrangiana?

Condición necesaria de optimalidad: Recuerde que si $\hat{x} \in C$ resuelve el problema(asumiendo diferenciabilidad de las funciones involucradas), entonces la condición necesaria para el (ProMix) puede expresarse como

$$-\nabla_x L(\hat{x}, \lambda, \mu) \in N_C(\hat{x}).$$

Condición necesaria de optimalidad: Recuerde que si $\hat{x} \in C$ resuelve el problema(asumiendo diferenciabilidad de las funciones involucradas), entonces la condición necesaria para el (ProMix) puede expresarse como

$$-\nabla_x L(\hat{x}, \lambda, \mu) \in N_C(\hat{x}).$$

Condición suficiente de optimalidad: Si C es convexo cerrado, f, g_i convexas y h_j afines lineales, entonces si existen $\hat{x} \in C, (\hat{\lambda}, \hat{\mu}) \in \Lambda_0$, que satisfacen las condiciones KKT, se sigue que $\hat{x}$ resuelve (Promix).

Condición necesaria de optimalidad: Recuerde que si $\hat{x} \in C$ resuelve el problema(asumiendo diferenciabilidad de las funciones involucradas), entonces la condición necesaria para el (ProMix) puede expresarse como

$$-\nabla_x L(\hat{x}, \lambda, \mu) \in N_C(\hat{x}).$$

Condición suficiente de optimalidad: Si C es convexo cerrado, f, g_i convexas y h_j afines lineales, entonces si existen $\hat{x} \in C, (\hat{\lambda}, \hat{\mu}) \in \Lambda_0$, que satisfacen las condiciones KKT, se sigue que $\hat{x}$ resuelve (Promix).
Dem de la condición de suficiencia:
$L(x, (\hat{\lambda}, \hat{\mu}))$ es convexa con respecto a $x \in C$, por lo que

$$L(x, (\hat{\lambda}, \hat{\mu})) \geq L(\hat{x}, (\hat{\lambda}, \hat{\mu})) + (L_x(\hat{x}, (\hat{\lambda}, \hat{\mu})))'(x - \hat{x}), \ \forall x \in C$$

y como $-\nabla_x L(\hat{x}, \hat{\lambda}, \hat{\mu}) \in N_C(\hat{x})$ es decir $(L_x(\hat{x}, (\hat{\lambda}, \hat{\mu})))'(x - \hat{x}) \geq 0$, entonces

$$L(x, (\hat{\lambda}, \hat{\mu})) \geq L(\hat{x}, (\hat{\lambda}, \hat{\mu}))$$

y si $x \in X$(el conjunto factible) es decir $g_i(x) \leq 0, \ h_j(x) = 0$, entonces
$L(\hat{x}, (\hat{\lambda}, \hat{\mu})) \leq f(x)$ con igualdad en $\hat{x}$.

Condición necesaria de optimalidad: Recuerde que si $\hat{x} \in C$ resuelve el problema(asumiendo diferenciabilidad de las funciones involucradas), entonces la condición necesaria para el (ProMix) puede expresarse como

$$-\nabla_x L(\hat{x}, \lambda, \mu) \in N_C(\hat{x}).$$

Condición suficiente de optimalidad: Si C es convexo cerrado, f, g_i convexas y h_j afines lineales, entonces si existen $\hat{x} \in C, (\hat{\lambda}, \hat{\mu}) \in \Lambda_0$, que satisfacen las condiciones KKT, se sigue que $\hat{x}$ resuelve (Promix).
Dem de la condición de suficiencia:
$L(x, (\hat{\lambda}, \hat{\mu}))$ es convexa con respecto a $x \in C$, por lo que

$$L(x, (\hat{\lambda}, \hat{\mu})) \geq L(\hat{x}, (\hat{\lambda}, \hat{\mu})) + (L_x(\hat{x}, (\hat{\lambda}, \hat{\mu})))'(x - \hat{x}), \ \forall x \in C$$

y como $-\nabla_x L(\hat{x}, \hat{\lambda}, \hat{\mu}) \in N_C(\hat{x})$ es decir $(L_x(\hat{x}, (\hat{\lambda}, \hat{\mu})))'(x - \hat{x}) \geq 0$, entonces

$$L(x, (\hat{\lambda}, \hat{\mu})) \geq L(\hat{x}, (\hat{\lambda}, \hat{\mu}))$$

y si $x \in X$(el conjunto factible) es decir $g_i(x) \leq 0, \ h_j(x) = 0$, entonces $L(\hat{x}, (\hat{\lambda}, \hat{\mu})) \leq f(x)$ con igualdad en $\hat{x}$.
En consecuencia $f(\hat{x}) \leq f(x), \ \forall x \in X$.

Ejemplo

Sean $f_1, \cdots, f_m$ funciones continuamente diferenciables en $\mathbb{R}^n$, se define la función $f(x) := \max\limits_{\overline{1,m}} \{f_i(x)\}$ y se formula el problema

$$\begin{array}{c} \min f(x) \\ x \in \mathbb{R}^n \end{array} \tag{6}$$

Sean $f_1, \cdots, f_m$ funciones continuamente diferenciables en $\mathbb{R}^n$, se define la función $f(x) := \underset{\overline{1,m}}{\text{máx}}\{f_i(x)\}$ y se formula el problema

$$\begin{array}{c} \text{mín } f(x) \\ x \in \mathbb{R}^n \end{array} \tag{6}$$

Haciendo la siguiente asignación $v := \underset{\overline{1,m}}{\text{máx}}\{f_i(x)\}$, entonces para cada $i = 1, \cdots, m$ se cumple $f_i(x) \leq v$ y (6) se puede formular alternativamente como

$$\begin{array}{ll} \underset{x,v}{\text{mín}} & v \\ s.a. & f_i(x) - v \leq 0, \; i = 1, \cdots, m. \end{array} \tag{7}$$

Sean $f_1, \cdots, f_m$ funciones continuamente diferenciables en $\mathbb{R}^n$, se define la función $f(x) := \underset{\overline{1,m}}{\text{máx}}\{f_i(x)\}$ y se formula el problema

$$\begin{array}{c} \text{mín } f(x) \\ x \in \mathbb{R}^n \end{array} \tag{6}$$

Haciendo la siguiente asignación $v := \underset{\overline{1,m}}{\text{máx}}\{f_i(x)\}$, entonces para cada $i = 1, \cdots, m$ se cumple $f_i(x) \leq v$ y (6) se puede formular alternativamente como

$$\begin{array}{ll} \underset{x,v}{\text{mín}} & v \\ s.a. & f_i(x) - v \leq 0, \; i = 1, \cdots, m. \end{array} \tag{7}$$

Note que si $(\hat{x}, \hat{v})$ resuelve (7), entonces $\hat{x}$ resuelve (6). Recíprocamente, si $\hat{x}$ resuelve (6), entonces $(\hat{x}, f(\hat{x}))$ resuelve (7).
Si cada f_i es una función convexa, entonces las condiciones necesarias resultan ser condiciones suficientes.

Sean $f_1, \cdots, f_m$ funciones continuamente diferenciables en $\mathbb{R}^n$, se define la función $f(x) := \max\limits_{\underline{1,m}}\{f_i(x)\}$ y se formula el problema

$$\begin{array}{c} \min f(x) \\ x \in \mathbb{R}^n \end{array} \tag{6}$$

Haciendo la siguiente asignación $v := \max\limits_{\underline{1,m}}\{f_i(x)\}$, entonces para cada $i = 1, \cdots, m$ se cumple $f_i(x) \leq v$ y (6) se puede formular alternativamente como

$$\begin{array}{ll} \min\limits_{x,v} & v \\ s.a. & f_i(x) - v \leq 0, \; i = 1, \cdots, m. \end{array} \tag{7}$$

Note que si $(\hat{x}, \hat{v})$ resuelve (7), entonces $\hat{x}$ resuelve (6). Recíprocamente, si $\hat{x}$ resuelve (6), entonces $(\hat{x}, f(\hat{x}))$ resuelve (7).
Si cada f_i es una función convexa, entonces las condiciones necesarias resultan ser condiciones suficientes.
Tarea: Asumiendo que el problema tiene solución, verificar la condición de Mangasarian y así aplicar las condiciones KKT.

- Sean $X = \{2, 3\}$ y $Y = \{-1, 0, 1\}$. Se define la función L en $X \times Y$ mediante:

x \ y	-1	0	1
2	1	2	-1
3	4	-1	0

- Sean $X = \{2, 3\}$ y $Y = \{-1, 0, 1\}$. Se define la función L en $X \times Y$ mediante:

x \ y	-1	0	1
2	1	2	-1
3	4	-1	0

$\min\limits_{x} \max\limits_{y} L(x, y) = 2$ y este valor se alcanza en $(x, y) = (2, 0)$.

- Sean $X = \{2,3\}$ y $Y = \{-1,0,1\}$. Se define la función L en $X \times Y$ mediante:

x \ y	-1	0	1
2	1	2	-1
3	4	-1	0

  $\min\limits_{x} \max\limits_{y} L(x,y) = 2$ y este valor se alcanza en $(x,y) = (2,0)$. Mientras que $\max\limits_{y} \min\limits_{x} L(x,y) = 1$ y este valor se alcanza en $\cdots$.

- Sean $X = \{2,3\}$ y $Y = \{-1,0,1\}$. Se define la función L en $X \times Y$ mediante:

x \ y	-1	0	1
2	1	2	-1
3	4	-1	0

  $\min_x \max_y L(x,y) = 2$ y este valor se alcanza en $(x,y) = (2,0)$. Mientras que $\max_y \min_x L(x,y) = 1$ y este valor se alcanza en $\cdots$.

- Sean $X = Y = \mathbb{R}$ y la función $L(x,y) = x^2 - (y-1)^2$ para todo $(x,y) \in \mathbb{R} \times \mathbb{R}$. Se cumple

$$\min_x \max_y L(x,y) = \max_y \min_x L(x,y) = 0.$$

- Sean $X = \{2,3\}$ y $Y = \{-1,0,1\}$. Se define la función L en $X \times Y$ mediante:

x \ y	-1	0	1
2	1	2	-1
3	4	-1	0

  $\min_x \max_y L(x,y) = 2$ y este valor se alcanza en $(x,y) = (2,0)$. Mientras que $\max_y \min_x L(x,y) = 1$ y este valor se alcanza en $\cdots$.

- Sean $X = Y = \mathbb{R}$ y la función $L(x,y) = x^2 - (y-1)^2$ para todo $(x,y) \in \mathbb{R} \times \mathbb{R}$. Se cumple

$$\min_x \max_y L(x,y) = \max_y \min_x L(x,y) = 0.$$

En general, para una función L definida en $X \times Y$ se cumple:

$$\min_x \max_y L(x,y) \geq \max_y \min_x L(x,y)$$

Dada una función L definida en $X \times Y$, se dice que $(\hat{x}, \hat{y}) \in X \times Y$ es un punto silla de L en $X \times Y$ si se cumple

$$\min_{x} \max_{y} L(x, y) = \max_{y} \min_{x} L(x, y) = L(\hat{x}, \hat{y})$$

Dada una función L definida en $X \times Y$, se dice que $(\hat{x}, \hat{y}) \in X \times Y$ es un punto silla de L en $X \times Y$ si se cumple

$$\min_x \max_y L(x, y) = \max_y \min_x L(x, y) = L(\hat{x}, \hat{y})$$

En un ejemplo anterior con $L(x, y) = x^2 - (y-1)^2$ en $\mathbb{R} \times \mathbb{R}$, se tiene que $(0, 1)$ es un punto silla para L.

Dada una función L definida en $X \times Y$, se dice que $(\hat{x}, \hat{y}) \in X \times Y$ es un punto silla de L en $X \times Y$ si se cumple

$$\min_x \max_y L(x,y) = \max_y \min_x L(x,y) = L(\hat{x}, \hat{y})$$

En un ejemplo anterior con $L(x,y) = x^2 - (y-1)^2$ en $\mathbb{R} \times \mathbb{R}$, se tiene que $(0,1)$ es un punto silla para L.
Ejemplo:
Para $X = \mathbb{R}$, $Y = \mathbb{R}_+$, analizar la existencia de punto silla de la función $L(x,y) = x^2 - x + yx$.

Sean f, g_i, h_j funciones definidas en $\mathbb{R}^n$ y C un conjunto cerrado de $\mathbb{R}^n$. Para el problema

$$(ProMix) \qquad \begin{array}{ll} \min & f(x) \\ s.a. & g_i(x) \leq 0,\ i = 1, \cdots, m. \\ & h_j(x) = 0,\ j = 1, \cdots, p. \\ & x \in C \end{array} \tag{8}$$

definimos la función Lagrangiana

$$L(x, \lambda, \mu) := f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x) \tag{9}$$

con $x \in C, (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^p =: \Lambda_0$.

Definimos la función primal asociada al problema (ProMix) como:

$$L_P(x) := \sup_{(\lambda,\mu)\in\Lambda_0} L(x,(\lambda,\mu)) \ , \ x \in C \tag{10}$$

Mientras que la función dual asociada al problema (ProMix) se defini como

$$L_D((\lambda,\mu)) := \inf_{x\in C} L(x,(\lambda,\mu)) \ , \ (\lambda,\mu) \in \Lambda_0 \tag{11}$$

Definimos la función primal asociada al problema (ProMix) como:

$$L_P(x) := \sup_{(\lambda,\mu)\in\Lambda_0} L(x,(\lambda,\mu)) \ , \ x \in C \tag{10}$$

Mientras que la función dual asociada al problema (ProMix) se defini como

$$L_D((\lambda,\mu)) := \inf_{x\in C} L(x,(\lambda,\mu)) \ , \ (\lambda,\mu) \in \Lambda_0 \tag{11}$$

Si en (10) el supremo resulta $+\infty$, entonces se asigna $L_P(x) = +\infty$. Análogamente, si en (11) el ínfimo resulta $-\infty$, entonces se asigna $L_D((\lambda,\mu)) = -\infty$.

- Considere el problema

$$\begin{array}{ll} \text{mín} & x^2+x \\ s.a. & x \leq 1 \end{array} \quad \Rightarrow \quad \begin{array}{ll} \text{mín} & x^2+x \\ s.a. & x-1 \leq 0 \end{array}$$

Aquí $C = \mathbb{R}$ y $\Lambda_0 = \mathbb{R}_+$; la función lagrangiana asociada es $L(x,\lambda) = x^2 + x + \lambda x - \lambda$ y

$$L_P(x) = \sup_{\lambda \geq 0}[f(x) + \lambda(x-1)] = \begin{cases} f(x), & \text{si } x \leq 1 \\ +\infty, & \text{si } x > 1 \end{cases}$$

y

$$L_D(\lambda) = \inf_{x \in \mathbb{R}}(f(x) + \lambda x - \lambda) = -\frac{1}{4}(\lambda+1)^2 - \lambda,\ \lambda \geq 0.$$

- Considere el problema

$$\begin{array}{ll} \text{mín} & x_1^2 + 2x_2^2 - x_1 + x_2 \\ s.a. & x_1 x_2 = 1 \\ & x_1 \geq 0, x_2 \geq 0 \end{array} \quad \overset{(una\, manera)}{\Rightarrow} \quad \begin{array}{ll} \text{mín} & x_1^2 + 2x_2^2 - x_1 + x_2 \\ s.a. & x_1 x_2 - 1 = 0 \\ & (x_1, x_2) \in \mathbb{R}^2_+ \end{array}$$

Aquí $C = \mathbb{R}^2_+$ y $\Lambda_0 = \mathbb{R}$; la función lagrangiana asociada es $L((x_1, x_2), \mu) = f(x_1, x_2) + \mu(x_1 x_2 - 1))$ y

$$L_P(x_1, x_2) = \sup_{\mu \in \mathbb{R}} [f(x_1, x_2) + \mu(x_1 x_2 - 1)] = \begin{cases} f(x), & \text{si } x_1 x_2 = 1, x \in \mathbb{R}^2_+ \\ +\infty, & \text{si } x_1 x_2 \neq 1, x \in \mathbb{R}^2_+ \end{cases}$$

y

$$L_D(\mu) = \inf_{x_1 \geq 0, x_2 \geq 0} (x_1^2 + 2x_2^2 - x_1 + x_2 + \mu x_1 x_2 - \mu), \ \mu \in \mathbb{R}.$$

- Sean A una matriz de orden $m \times n$, $b \in \mathbb{R}^m$ y c un punto de $\mathbb{R}^n$. Considere el problema

$$\begin{array}{ll} \text{mín} & \frac{1}{2}\|x-c\|^2 \\ s.a. & Ax = b \end{array} \quad \Rightarrow \quad \begin{array}{ll} \text{mín} & \frac{1}{2}\|x-c\|^2 \\ s.a. & Ax - b = 0 \end{array}$$

Aquí $C = \mathbb{R}^n$ y $\Lambda_0 = \mathbb{R}^m$; la función lagrangiana asociada es $L(x,\mu) = \frac{1}{2}\|x-c\|^2 + \mu'(Ax-b)$ y

$$L_P(x) = \sup_{\mu}[\frac{1}{2}\|x-c\|^2 + \mu'(Ax-b)] = \begin{cases} \frac{1}{2}\|x-c\|^2, & \text{si } Ax-b=0 \\ +\infty, & \text{si } Ax-b \neq 1 \end{cases}$$

y

$$\begin{aligned} L_D(\mu) &= \inf_{x\in\mathbb{R}^n} (\underbrace{\frac{1}{2}\|x-c\|^2 + \mu' Ax}_{G(x) \text{ convex}} - \mu' b) \\ &= -\frac{1}{2}\|A'\mu\|^2 + \mu'(Ac-b). \end{aligned}$$

(a)

$$L_P(x) = \begin{cases} f(x), & \text{si } x \text{ es factible} \\ +\infty, & \text{en otro caso.} \end{cases}$$

(a)

$$L_P(x) = \begin{cases} f(x), & \text{si } x \text{ es factible} \\ +\infty, & \text{en otro caso.} \end{cases}$$

(b) $L_D(\lambda, \mu)$ es una función cóncava en Λ_0.

(a)

$$L_P(x) = \begin{cases} f(x), & \text{si } x \text{ es factible} \\ +\infty, & \text{en otro caso.} \end{cases}$$

(b) $L_D(\lambda, \mu)$ es una función cóncava en Λ_0.

(c) Asuma que para $(\lambda^*, \mu^*) \in \Lambda_0$, podemos encontrar $x^* \in C$ tal que $L_D(\lambda^*, \mu^*) = L(\lambda^*, \mu^*, x^*)$, entonces

$$L_D(\lambda, \mu) \leq L_D(\leq \lambda^*, \mu^*) + (g(x^*))'(\lambda - \lambda^*) + (h(x^*))'(\mu - \mu^*), \quad \forall (\lambda, \mu) \in \Lambda_0$$

(a)

$$L_P(x) = \begin{cases} f(x), & \text{si } x \text{ es factible} \\ +\infty, & \text{en otro caso.} \end{cases}$$

(b) $L_D(\lambda, \mu)$ es una función cóncava en Λ_0.

(c) Asuma que para $(\lambda^*, \mu^*) \in \Lambda_0$, podemos encontrar $x^* \in C$ tal que $L_D(\lambda^*, \mu^*) = L(\lambda^*, \mu^*, x^*)$, entonces

$$L_D(\lambda, \mu) \leq L_D(\leq \lambda^*, \mu^*) + (g(x^*))'(\lambda - \lambda^*) + (h(x^*))'(\mu - \mu^*), \quad \forall (\lambda, \mu) \in \Lambda_0$$

¿Qué papel desempeña $(g(x^*); h(x^*))$ para L_D en (λ^*, μ^*) ?

La función lagrangiana $L(x,(\lambda,\mu))$ está definida en $\Lambda = C \times \Lambda_0($ en su respectiva versión), se sabe que siempre se cumple

$$\sup_{(\lambda,\mu)} \inf_{x\in C} L(x,(\lambda,\mu)) \leq \inf_{x\in C} \sup_{(\lambda,\mu)} L(x,(\lambda,\mu))$$

es decir

$$\sup_{(\lambda,\mu)} L_D(\lambda,\mu) \leq \inf_{x\in C} L_P(x)$$

y en el caso que estos últimos $\sup$ e $\inf$ se alcancen, resulta que el máximo de la función dual es una cota inferior del mínimo de la función primal.

La función lagrangiana $L(x,(\lambda,\mu))$ está definida en $\Lambda = C \times \Lambda_0$(en su respectiva versión), se sabe que siempre se cumple

$$\sup_{(\lambda,\mu)} \inf_{x\in C} L(x,(\lambda,\mu)) \leq \inf_{x\in C} \sup_{(\lambda,\mu)} L(x,(\lambda,\mu))$$

es decir

$$\sup_{(\lambda,\mu)} L_D(\lambda,\mu) \leq \inf_{x\in C} L_P(x)$$

y en el caso que estos últimos $\sup$ e $\inf$ se alcancen, resulta que el máximo de la función dual es una cota inferior del mínimo de la función primal.
Tenga presente que en el conjunto factible X del problema (ProMix), se cumple que $L_P(x) = f(x)$ para cada $x \in X$. En consecuencia

$$\max_{(\lambda,\mu)} L_D(\lambda,\mu) \leq \text{ valor óptimo del problema (ProMix).}$$

Esta desigualdad se conoce como la condición de "Dualidad débil"

- **Problema Primal**:

$$(PP) \qquad \min_{x \in C} L_P(x)$$

y supongamos que su valor óptimo es α.

- **Problema Primal**:

$$(PP) \qquad \min_{x \in C} L_P(x)$$

y supongamos que su valor óptimo es α.

- **Problema Dual**:

$$(PD) \qquad \max_{(\lambda,\mu)\in\Lambda_0} L_D(\lambda,\mu)$$

con valor óptimo β.

- **Problema Primal**:

$$(PP) \qquad \min_{x\in C} L_P(x)$$

y supongamos que su valor óptimo es α.

- **Problema Dual**:

$$(PD) \qquad \max_{(\lambda,\mu)\in\Lambda_0} L_D(\lambda,\mu)$$

con valor óptimo β.

La condición de dualidad débil, sostiene que:

$$\beta \leq \alpha = \text{Valor óptimo de (ProMix)}$$