# Modelamiento y Optimización

## Clase 22

Gonzalo Muñoz

7 de Junio 2024

# Dualidad Lagrangeana

# Recapitulemos

Estamos estudiando problemas del tipo

$$
\begin{aligned}
\text{mín } & f(x) \\
\text{s.a. } & g_i(x) \leq 0 \quad i = 1, \ldots, m \\
& h_j(x) = 0 \quad j = 1, \ldots, p \\
& x \in \mathbb{R}^n
\end{aligned}
$$

$h_j(x) = a_j^T x + b_j$

Donde las funciones $h_j$ son lineales afín. Hasta ahora hemos visto que:

- Si la región factible $S$ es compacta y la función $f$ es continua, entonces existe un óptimo global.
- Si todas las funciones $g_i$ son convexas la región factible es un conjunto convexo. Si adicionalmente $f$ es convexa, el problema se llama problema convexo y los óptimos locales son globales.

Ahora veremos cómo encontrar mínimos (globales y locales).

# Simplificando el problema

Nos gustaría deshacernos de las restricciones para poder resolver un optimización sin restricciones; esto lo hacemos penalizando restricciones:

$$S\left\{\begin{array}{ll} \text{mín } f(x) & \\ \text{s.a. } g_i(x) \leq 0 & i = 1, \dots, m \\ h_j(x) = 0 & j = 1, \dots, p \\ x \in \mathbb{R}^n & \end{array}\right. \rightsquigarrow \min_{x \in \mathbb{R}^n} \overbrace{f(x) + \sum_{i=1}^{m} \underset{\geq 0}{\lambda_i} g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x)}^{L(x,\lambda,\mu)}$$

La función a optimizar se llama Lagrangeano, y siempre se tiene que para todo $x \in S$, $\lambda \in \mathbb{R}^m_+$ y $\mu \in \mathbb{R}^p$:

$$f(x) \geq L(x, \lambda, \mu) \qquad (*)$$

$$L(x,\lambda,\mu) = f(x) + \sum_{i=1}^{m} \overbrace{\lambda_i}^{\geq 0} \underbrace{g_i(x)}_{\leq 0} + \sum_{j=1}^{p} \mu_j \underbrace{h_j(x)}_{=0}$$

$$\leq f(x)$$

pues $x \in S$   pues $x \in S$

# Función dual

¿Qué pasa si minimizamos el Lagrangeano? Esto se conoce como la función dual:

$$d(\lambda, \mu) = \underbrace{\min_{x \in \mathbb{R}^n} L(x, \lambda, \mu).}_{\text{Sin restricciones}}$$

## Ejemplo

Calculemos el Lagrangeano y la función dual de

$$\min\ x^2 + 2x$$
$$\text{s.a. } x + 5 \leq 0$$
$$x \in \mathbb{R}$$

$$L(x, \lambda) = x^2 + 2x + \lambda(x + 5), \qquad \lambda \geq 0$$

$$d(\lambda) = \min_{x \in \mathbb{R}} \underbrace{x^2 + 2x + \lambda(x + 5)}_{\text{Convexa para todo } \lambda}$$

haremos $\frac{d}{dx} L(x, \lambda) = 0$

# Ejemplo

$$\frac{dL}{dx} = 2x + 2 + \lambda = 0$$

$$\Rightarrow \quad x = \frac{-(2+\lambda)}{2} = -1 - \frac{\lambda}{2}$$

$$\Rightarrow \quad d(\lambda) = L\left(-1 - \frac{\lambda}{2}, \lambda\right)$$

$$= \left(-1 - \frac{\lambda}{2}\right)^2 + 2\left(-1 - \frac{\lambda}{2}\right) + \lambda\left(-1 - \frac{\lambda}{2} + 5\right)$$

$$\vdots$$

$$= -\frac{1}{4}(2+\lambda)^2 + 5\lambda$$

# Relación de la función dual y el óptimo

La función dual simplifica las cosas, y siempre podemos garantizar que:

## Lema

*Sea* $(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^p$ *y* $p^*$ *el valor óptimo del problema original. Luego,*

$$\underbrace{d(\lambda, \mu) \leq p^*}.$$

sale principalmente de (*)

¿Qué tipo de información podemos deducir en nuestro ejemplo anterior?

$$d(\lambda) = -\frac{1}{4}(2+\lambda)^2 + 5\lambda$$

$$d(2) = -\frac{1}{4}(2+2)^2 + 5 \cdot 2 = 6 \Rightarrow p^* \geq 6$$

$$d(8) = -\frac{1}{4}(2+8)^2 + 5 \cdot 8 = 15 \Rightarrow p^* \geq 15$$

óptimo es $\geq 15$.

# La mejor cota inferior

Como sabemos que la función dual $d(\lambda, \mu)$ nos da una cota inferior, la mejor cota dada por el Lagrangeano es

$$d^* = \max\{d(\lambda, \mu) : \lambda \geq 0\}. \qquad (**)$$

Este es el problema dual, y por lo que vimos anteriormente

$$d^* \leq p^*$$

Lo que se conoce como dualidad débil. Calculemos $d^*$ en el ejemplo anterior.

$$d(\lambda) = \underbrace{-\frac{1}{4}(2+\lambda)^2 + 5\lambda}$$

cóncava $\Rightarrow$ haremos $\dfrac{d\, d(\lambda)}{d\lambda} = 0$

$$\frac{d}{d\lambda} d(\lambda) = -\frac{1}{2}(\lambda+2) + 5 = 0 \Rightarrow \boxed{\lambda = 8}$$

Como $\lambda \geq 0$, es óptimo para $(**)$

$$\Rightarrow \quad d^* = 15$$

(calculado anteriormente)

# La mejor cota inferior

En programación lineal, (casi) siempre se tenía que $d^* = p^*$. Sin embargo, en el caso no-lineal no siempre es el caso.

A $p^* - d^* \geq 0$ lo llamamos *gap de dualidad*, y cuando es cero diremos que el par primal-dual satisface *dualidad fuerte*.

Ejercicio: verificar "a mano" que el ejemplo anterior satisface dualidad fuerte.

Por ahora dejaremos pendiente algunas condiciones para garantizar que se satisfaga dualidad fuerte. Partiremos por estudiar algunas de sus consecuencias.

# Consecuencias de dualidad fuerte

Sea $x^*$ óptimo para el primal y $(\lambda^*, \mu^*)$ óptimo para el dual. Si se cumple dualidad fuerte:

$$f(x^*) = p^* = d^* = d(\lambda^*, \mu^*) = \min_{x \in \mathbb{R}^n} L(x, \lambda^*, \mu^*)$$

$$\leq L(x^*, \lambda^*, \mu^*)$$ / pues a la izq. hay un min

sandwich

$$= f(x^*) + \sum_{i=1}^{m} \lambda_i^* \underbrace{g_i(x^*)}_{\leq 0} + \sum_{j=1}^{p} \mu_j^* \underbrace{h_j(x^*)}_{=0}$$

$x^*$ es factible        $x^*$ es factible

$$\leq f(x^*)$$

$$\Rightarrow \sum_{i=1}^{m} \lambda_i^* \underbrace{g_i(x^*)}_{\text{todos son } \leq 0} = 0 \quad \Bigg\} \Rightarrow \underbrace{\lambda_i^* g_i(x^*) = 0}_{\text{Holgura complementaria}} \quad \forall i$$

# Otra consecuencia

De acuerdo a lo que acabamos de ver

$$\min_{x\in\mathbb{R}^n} L(x,\lambda^*,\mu^*) = L(x^*,\lambda^*,\mu^*)$$

Es decir, $x^*$ es un mínimo *global* de la función $L(x,\lambda^*,\mu^*)$. Por lo tanto

$$0 = \underbrace{\nabla_x L(x^*,\lambda^*,\mu^*)}_{\text{gradiente y luego evaluar}} = \nabla_x \Big( f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{i=1}^{p} \mu_j h_j(x^*) \Big)$$

$$= \nabla_x f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla h_j(x^*)$$

# Condiciones de KKT

juntando todo :

factibilidad primal
$$\begin{cases} g_i(x^*) \leq 0 & \forall i = 1, \ldots, m \\ h_j(x^*) = 0 & \forall j = 1, \ldots, p \end{cases}$$

factibilidad dual
$$\lambda_i^* \geq 0 \qquad \forall i = 1, \ldots, m$$

$$\lambda_i^* g_i(x^*) = 0 \qquad \forall i = 1, \ldots, m$$

Holgura complementaria

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla g_j(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla h_j(x^*) = 0$$

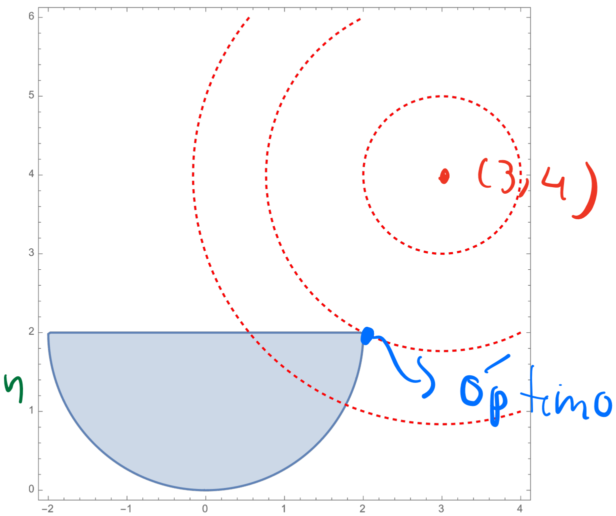
Cond. 1er orden en el Lagrangeano

Estas son las condiciones de Karush-Kuhn-Tucker, o KKT. Siempre que haya dualidad fuerte, un par primal-dual óptimo las debe cumplir.

( ojo , si algo cumple KKT $\not\Rightarrow$ es óptimo necesariamente )

# Ejemplo

Veamos que el óptimo del ejemplo de la clase pasada cumple KKT




óptimo (2,2)

$$
\begin{aligned}
\text{mín} \quad & (x_1-3)^2+(x_2-4)^2 \\
\text{sujeto a:} \quad & x_2 \leq 2 \quad \checkmark \\
& x_1^2+(x_2-2)^2 \leq 4 \quad \checkmark
\end{aligned}
$$

se cumplen

H-C:

$$\lambda_1 (\overbrace{x_2 - 2}^{=0 \text{ en } (2,2)}) = 0$$

$$\lambda_2 (\underbrace{x_1^2 + (x_2-2)^2 - 4}_{=0 \text{ en } (2,2)}) = 0$$

$\nabla_x L = 0$:

$$\frac{\partial L}{\partial x_1} = 2(x_1-3) + 2\lambda_2 \overset{\text{en } (2,2)}{=} -2 + 2\lambda_2$$

$$\frac{\partial L}{\partial x_2} = 2(x_2-4) + \lambda_1 + 2\lambda_2(x_2-2)$$

$= -4 + \lambda_1$
en (2,2)

Factibilidad dual

$\Rightarrow \quad \left.\begin{array}{l} -2 + 2\lambda_2 = 0 \\ -4 + \lambda_1 = 0 \end{array}\right\} \quad \begin{array}{l} \lambda_2 = 1 \geq 0 \\ \lambda_1 = 4 \geq 0 \end{array}$

Como $(\lambda_1, \lambda_2) \geq 0 \Rightarrow x^* = (2,2)$, $\lambda^* = (4,1)$ cumplen KKT.