Optimización convexa

Fernando Lozano

Universidad de los Andes

26 de septiembre de 2022

El Langrangiano

$$\begin{gathered}\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})\\ \text{sujeto a}\\ f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m\\ h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p\end{gathered}$$

El Langrangiano

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

sujeto a

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p$$

- El Lagrangiano $L : (\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p) \to \mathbb{R}$,

El Langrangiano

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

sujeto a

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p$$

- El Lagrangiano $L : (\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p) \to \mathbb{R}$,

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x})$$

El Langrangiano

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

$$\text{sujeto a}$$

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p$$

- El Lagrangiano $L : (\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p) \to \mathbb{R}$,

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x})$$

El Langrangiano

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

sujeto a

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$
$$h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p$$

- El Lagrangiano $L : (\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p) \to \mathbb{R}$,

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x})$$

El Langrangiano

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

sujeto a

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$
$$h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p$$

- El Lagrangiano $L : (\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p) \to \mathbb{R}$,

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x})$$

- $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_m]$ y $\boldsymbol{\nu} = [\nu_1, \nu_2, \ldots, \nu_p]$ son las variables duales o multiplicadores de Lagrange.

La función Dual

- Definimos la función dual como el valor mínimo del Lagrangiano sobre $\mathbf{x}$:

La función Dual

- Definimos la función dual como el valor mínimo del Lagrangiano sobre $\mathbf{x}$:

$$\begin{aligned} g(\boldsymbol{\lambda}, \boldsymbol{\nu}) &= \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \\ &= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \right) \end{aligned}$$

La función Dual

- Definimos la función dual como el valor mínimo del Lagrangiano sobre $\mathbf{x}$:

$$
\begin{aligned}
g(\boldsymbol{\lambda}, \boldsymbol{\nu}) &= \min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \\
&= \min_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \right)
\end{aligned}
$$

La función Dual

- Definimos la función dual como el valor mínimo del Lagrangiano sobre $\mathbf{x}$:

$$
\begin{aligned}
g(\boldsymbol{\lambda}, \boldsymbol{\nu}) &= \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \right)
\end{aligned}
$$

La función Dual

- Definimos la función dual como el valor mínimo del Lagrangiano sobre $\mathbf{x}$:

$$
\begin{aligned}
g(\boldsymbol{\lambda}, \boldsymbol{\nu}) &= \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \right)
\end{aligned}
$$

- $g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ es una función cóncava.

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$.

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$. Para cualquier $\boldsymbol{\lambda} \geq \mathbf{0}$ y cualquier $\boldsymbol{\nu}$ tenemos:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*$$

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$. Para cualquier $\boldsymbol{\lambda} \geq \mathbf{0}$ y cualquier $\boldsymbol{\nu}$ tenemos:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*$$

- Suponga que $\mathbf{x}^\circ$ es un punto factible y $\boldsymbol{\lambda} \geq \mathbf{0}$, entonces

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$. Para cualquier $\boldsymbol{\lambda} \geq \mathbf{0}$ y cualquier $\boldsymbol{\nu}$ tenemos:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*$$

- Suponga que $\mathbf{x}^\circ$ es un punto factible y $\boldsymbol{\lambda} \geq \mathbf{0}$, entonces

$$\sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) \leq 0$$

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$. Para cualquier $\boldsymbol{\lambda} \geq \mathbf{0}$ y cualquier $\boldsymbol{\nu}$ tenemos:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*$$

- Suponga que $\mathbf{x}^\circ$ es un punto factible y $\boldsymbol{\lambda} \geq \mathbf{0}$, entonces

$$\sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) \leq 0$$

$$f_0(\mathbf{x}^\circ) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) \leq f_0(\mathbf{x}^\circ)$$

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$. Para cualquier $\boldsymbol{\lambda} \geq \mathbf{0}$ y cualquier $\boldsymbol{\nu}$ tenemos:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*$$

- Suponga que $\mathbf{x}^\circ$ es un punto factible y $\boldsymbol{\lambda} \geq \mathbf{0}$, entonces

$$\begin{aligned} \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) &\leq 0 \\ f_0(\mathbf{x}^\circ) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) &\leq f_0(\mathbf{x}^\circ) \\ g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq L(\mathbf{x}^\circ, \boldsymbol{\lambda}, \boldsymbol{\nu}) &\leq f_0(\mathbf{x}^\circ) \end{aligned}$$

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$. Para cualquier $\boldsymbol{\lambda} \geq \mathbf{0}$ y cualquier $\boldsymbol{\nu}$ tenemos:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*$$

- Suponga que $\mathbf{x}^\circ$ es un punto factible y $\boldsymbol{\lambda} \geq \mathbf{0}$, entonces

$$\begin{aligned}
\sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) &\leq 0 \\
f_0(\mathbf{x}^\circ) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) &\leq f_0(\mathbf{x}^\circ) \\
g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq L(\mathbf{x}^\circ, \boldsymbol{\lambda}, \boldsymbol{\nu}) &\leq f_0(\mathbf{x}^\circ)
\end{aligned}$$

- Podemos reemplazar $\mathbf{x}^\circ$ por $\mathbf{x}^*$.

La función dual es una cota inferior de $f(\mathbf{x}^*)$

- Sea $f(\mathbf{x}^*) = p^*$. Para cualquier $\boldsymbol{\lambda} \geq \mathbf{0}$ y cualquier $\boldsymbol{\nu}$ tenemos:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*$$

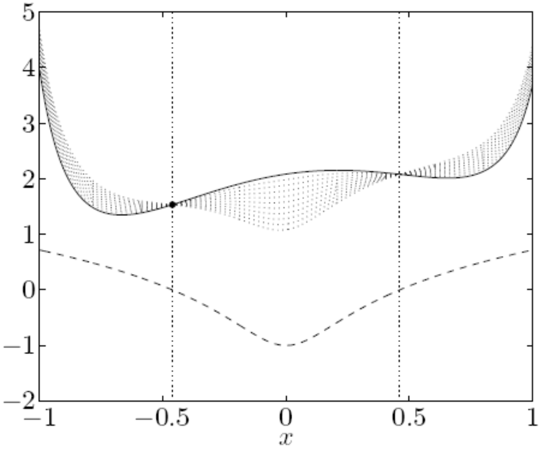
- Suponga que $\mathbf{x}^\circ$ es un punto factible y $\boldsymbol{\lambda} \geq \mathbf{0}$, entonces

$$\sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) \leq 0$$

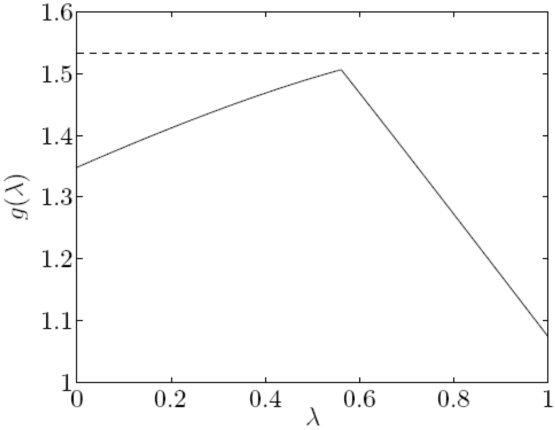
$$f_0(\mathbf{x}^\circ) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}^\circ) \leq f_0(\mathbf{x}^\circ)$$

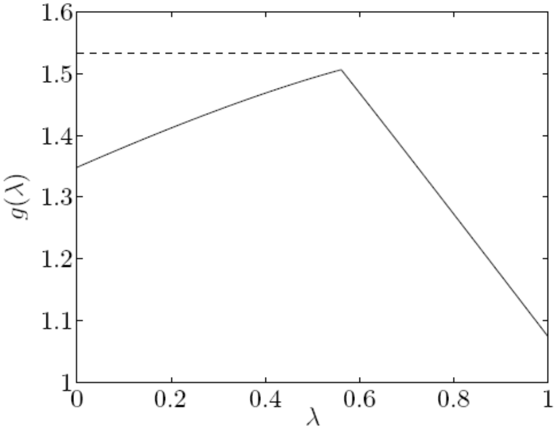
$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \operatorname*{ínf}_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq L(\mathbf{x}^\circ, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq f_0(\mathbf{x}^\circ)$$

- Podemos reemplazar $\mathbf{x}^\circ$ por $\mathbf{x}^*$.
- Cota no trivial sólo si $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) > -\infty$



5
4
3
2
1
0
−1
−2
−1
−0.5
0
0.5
1
x


1.6
1.5
1.4
1.3
1.2
1.1
1
$g(\lambda)$
0
0.2
0.4
0.6
0.8
1
λ



- Note que aunque f_0 y f_i no son convexas, g es cóncava.

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu})$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} +$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) =$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) = 2\mathbf{x} + \mathbf{A}^T\boldsymbol{\nu}$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) = 2\mathbf{x} + \mathbf{A}^T\boldsymbol{\nu} = 0$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) = 2\mathbf{x} + \mathbf{A}^T\boldsymbol{\nu} = 0 \Rightarrow \mathbf{x} = -\frac{1}{2}\mathbf{A}^T\boldsymbol{\nu}$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) = 2\mathbf{x} + \mathbf{A}^T\boldsymbol{\nu} = 0 \Rightarrow \mathbf{x} = -\frac{1}{2}\mathbf{A}^T\boldsymbol{\nu}$$

Reemplazando:

$$g(\boldsymbol{\nu}) = L(-\frac{1}{2}\mathbf{A}^T\boldsymbol{\nu}, \boldsymbol{\nu})$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) = 2\mathbf{x} + \mathbf{A}^T\boldsymbol{\nu} = 0 \Rightarrow \mathbf{x} = -\frac{1}{2}\mathbf{A}^T\boldsymbol{\nu}$$

Reemplazando:

$$g(\boldsymbol{\nu}) = L(-\frac{1}{2}\mathbf{A}^T\boldsymbol{\nu}, \boldsymbol{\nu}) = -\frac{1}{4}\boldsymbol{\nu}^T(\mathbf{A}\mathbf{A}^T)\boldsymbol{\nu} - \mathbf{b}^T\nu$$

Ejemplo

$$\min \mathbf{x}^T\mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p\times n}$$

- El Lagrangiano:

$$L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

- Podemos hallar la función dual minimizando con respecto a $\mathbf{x}$:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) = 2\mathbf{x} + \mathbf{A}^T\boldsymbol{\nu} = 0 \Rightarrow \mathbf{x} = -\frac{1}{2}\mathbf{A}^T\boldsymbol{\nu}$$

  Reemplazando:

$$g(\boldsymbol{\nu}) = L(-\frac{1}{2}\mathbf{A}^T\boldsymbol{\nu}, \boldsymbol{\nu}) = -\frac{1}{4}\boldsymbol{\nu}^T(\mathbf{A}\mathbf{A}^T)\boldsymbol{\nu} - \mathbf{b}^T\nu$$

- Función cuadrática cóncava.

El Problema Dual de Lagrange

El Problema Dual de Lagrange

- Para cada $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ con $\boldsymbol{\lambda} > \mathbf{0}$, la función dual de Lagrange nos da una cota inferior de p^*.

El Problema Dual de Lagrange

- Para cada $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ con $\boldsymbol{\lambda} > \mathbf{0}$, la función dual de Lagrange nos da una cota inferior de p^*.
- Cuál es la mejor cota inferior que podemos obtener?

El Problema Dual de Lagrange

- Para cada $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ con $\boldsymbol{\lambda} > \mathbf{0}$, la función dual de Lagrange nos da una cota inferior de p^*.
- Cuál es la mejor cota inferior que podemos obtener?
- Problema de optimización:

$$\begin{aligned} \text{máx} \quad & g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \\ \text{sujeto a} \quad & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned}$$

El Problema Dual de Lagrange

- Para cada $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ con $\boldsymbol{\lambda} > \mathbf{0}$, la función dual de Lagrange nos da una cota inferior de p^*.
- Cuál es la mejor cota inferior que podemos obtener?
- Problema de optimización:

$$\begin{aligned} \text{máx} \quad & g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \\ \text{sujeto a} \quad & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned}$$

- Este es el problema dual de Lagrange asociado al problema primal original.

El Problema Dual de Lagrange

- Para cada $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ con $\boldsymbol{\lambda} > \mathbf{0}$, la función dual de Lagrange nos da una cota inferior de p^*.
- Cuál es la mejor cota inferior que podemos obtener?
- Problema de optimización:

$$\begin{aligned} \text{máx} \quad & g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \\ \text{sujeto a} \quad & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned}$$

- Este es el problema dual de Lagrange asociado al problema primal original.
- $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ es factible en el dual si $\boldsymbol{\lambda} \geq \mathbf{0}$ y $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) > -\infty$.

El Problema Dual de Lagrange

- Para cada $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ con $\boldsymbol{\lambda} > \mathbf{0}$, la función dual de Lagrange nos da una cota inferior de p^*.
- Cuál es la mejor cota inferior que podemos obtener?
- Problema de optimización:

$$\begin{aligned} \text{máx} \quad & g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \\ \text{sujeto a} \quad & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned}$$

- Este es el problema dual de Lagrange asociado al problema primal original.
- $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ es factible en el dual si $\boldsymbol{\lambda} \geq \mathbf{0}$ y $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) > -\infty$.
- Solución óptima dual $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ (multiplicadores de Lagrange óptimos).

Ejemplo

- Primal:

$$\min_{\mathbf{x}} \mathbf{x}^T \mathbf{x} \quad \text{sujeto a} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p \times n}$$

- Dual

$$\max_{\boldsymbol{\nu}} -\frac{1}{4}\boldsymbol{\nu}^T(\mathbf{A}\mathbf{A}^T)\boldsymbol{\nu} - \mathbf{b}^T \nu$$

Dualidad Débil y Dualidad Fuerte

Dualidad Débil y Dualidad Fuerte

- Sea $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, entonces,

Dualidad Débil y Dualidad Fuerte

- Sea $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, entonces,

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

Dualidad Débil y Dualidad Fuerte

- Sea $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, entonces,

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

- $p^* - d^*$ es la brecha de dualidad óptima.

Dualidad Débil y Dualidad Fuerte

- Sea $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, entonces,

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

- $p^* - d^*$ es la brecha de dualidad óptima.
- Si la brecha de dualidad óptima es cero $(d^* = p^*)$ tenemos dualidad fuerte.

Dualidad Débil y Dualidad Fuerte

- Sea $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, entonces,

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

- $p^* - d^*$ es la brecha de dualidad óptima.
- Si la brecha de dualidad óptima es cero $(d^* = p^*)$ tenemos dualidad fuerte.
- Dualidad fuerte: condiciones de Karush-Kuhn-Tucker.

Dualidad Débil y Dualidad Fuerte

- Sea $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, entonces,

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

- $p^* - d^*$ es la brecha de dualidad óptima.
- Si la brecha de dualidad óptima es cero $(d^* = p^*)$ tenemos dualidad fuerte.
- Dualidad fuerte: condiciones de Karush-Kuhn-Tucker.
- Métodos de solución.

Dualidad Débil y Dualidad Fuerte

- Sea $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, entonces,

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

- $p^* - d^*$ es la brecha de dualidad óptima.
- Si la brecha de dualidad óptima es cero $(d^* = p^*)$ tenemos dualidad fuerte.
- Dualidad fuerte: condiciones de Karush-Kuhn-Tucker.
- Métodos de solución.
- Tenemos dualidad fuerte, por ejemplo cuando se minimiza una función convexa en un poliedro convexo.

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$f_0(\mathbf{x}^*) = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$$

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$
\begin{aligned}
f_0(\mathbf{x}^*) &= g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}) \right)
\end{aligned}
$$

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$
\begin{aligned}
f_0(\mathbf{x}^*) &= g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}) \right) \\
&\leq f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}^*)
\end{aligned}
$$

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$\begin{aligned} f_0(\mathbf{x}^*) &= g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \\ &= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}) \right) \\ &\leq f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}^*) \\ &\leq f_0(\mathbf{x}^*) \end{aligned}$$

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$
\begin{aligned}
f_0(\mathbf{x}^*) &= g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}) \right) \\
&\leq f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}^*) \\
&\leq f_0(\mathbf{x}^*)
\end{aligned}
$$

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$\begin{aligned}
f_0(\mathbf{x}^*) &= g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}) \right) \\
&\color{red}{=} f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}^*) \\
&\color{red}{=} f_0(\mathbf{x}^*)
\end{aligned}$$

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$
\begin{aligned}
f_0(\mathbf{x}^*) &= g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}) \right) \\
&= f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}^*) \\
&= f_0(\mathbf{x}^*)
\end{aligned}
$$

- $\mathbf{x}^*$ minimiza $L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ sobre $\mathbf{x}$.

Holgura Complementaria

- Suponga que tenemos $p^* = d^*$.
- Tenemos:

$$
\begin{aligned}
f_0(\mathbf{x}^*) &= g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \\
&= \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}) \right) \\
&= f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}^*) \\
&= f_0(\mathbf{x}^*)
\end{aligned}
$$

- $\mathbf{x}^*$ minimiza $L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ sobre $\mathbf{x}$.
- $\sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) = 0$

- Podemos expresar la condición de holgura complementaria:

$$\sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) = 0$$

como

- Podemos expresar la condición de holgura complementaria:

$$\sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) = 0$$

como

$$\lambda_i^* > 0 \Rightarrow f_i(\mathbf{x}^*) = 0$$

- Podemos expresar la condición de holgura complementaria:

$$\sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) = 0$$

como

$$\lambda_i^* > 0 \Rightarrow f_i(\mathbf{x}^*) = 0$$

o

$$f_i(\mathbf{x}^*) < 0 \Rightarrow \lambda_i^* = 0$$

- Podemos expresar la condición de holgura complementaria:

$$\sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) = 0$$

como

$$\lambda_i^* > 0 \Rightarrow f_i(\mathbf{x}^*) = 0$$

o

$$f_i(\mathbf{x}^*) < 0 \Rightarrow \lambda_i^* = 0$$

- Es decir, el i-ésimo multiplicador de Lagrange es cero, a no ser que la restricción correspondiente sea activa

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferenciables.

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \dots, f_m, h_1, \dots, h_p$ diferenciables.
- En un par de puntos óptimos para el primal $(\mathbf{x}^*)$ y el dual $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, se cumplen las siguientes condiciones:

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferenciables.
- En un par de puntos óptimos para el primal $(\mathbf{x}^*)$ y el dual $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, se cumplen las siguientes condiciones:

$$f_i(\mathbf{x}^*) \leq 0, \quad i = 1, \ldots, m.$$

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferenciables.
- En un par de puntos óptimos para el primal $(\mathbf{x}^*)$ y el dual $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, se cumplen las siguientes condiciones:

$$\begin{aligned} f_i(\mathbf{x}^*) &\leq 0, \quad i = 1, \ldots, m. \\ h_i(\mathbf{x}^*) &= 0, \quad i = 1, \ldots, p. \end{aligned}$$

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferenciables.
- En un par de puntos óptimos para el primal $(\mathbf{x}^*)$ y el dual $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, se cumplen las siguientes condiciones:

$$\begin{aligned} f_i(\mathbf{x}^*) &\leq 0, \quad i = 1, \ldots, m. \\ h_i(\mathbf{x}^*) &= 0, \quad i = 1, \ldots, p. \\ \lambda_i^* &\geq 0, \quad i = 1, \ldots, m. \end{aligned}$$

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferenciables.
- En un par de puntos óptimos para el primal ($\mathbf{x}^*$) y el dual ($\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*$), se cumplen las siguientes condiciones:

$$\begin{aligned} f_i(\mathbf{x}^*) \leq 0, &\quad i = 1, \ldots, m. \\ h_i(\mathbf{x}^*) = 0, &\quad i = 1, \ldots, p. \\ \lambda_i^* \geq 0, &\quad i = 1, \ldots, m. \\ \lambda_i^* f_i(\mathbf{x}^*) = 0, &\quad i = 1, \ldots, m. \end{aligned}$$

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferenciables.
- En un par de puntos óptimos para el primal $(\mathbf{x}^*)$ y el dual $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, se cumplen las siguientes condiciones:

$$\begin{aligned} f_i(\mathbf{x}^*) &\leq 0, \quad i = 1, \ldots, m. \\ h_i(\mathbf{x}^*) &= 0, \quad i = 1, \ldots, p. \\ \lambda_i^* &\geq 0, \quad i = 1, \ldots, m. \\ \lambda_i^* f_i(\mathbf{x}^*) &= 0, \quad i = 1, \ldots, m. \end{aligned}$$

$$\nabla f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(\mathbf{x}^*) = 0$$

Condiciones de Karush-Khun-Tucker (KKT)

- Suponga $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferenciables.
- En un par de puntos óptimos para el primal ($\mathbf{x}^*$) y el dual ($\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*$), se cumplen las siguientes condiciones:

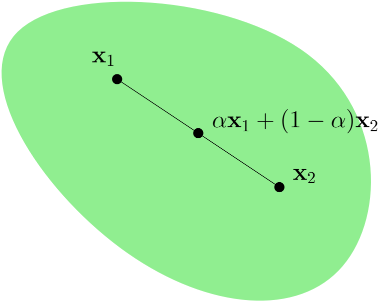
$$
\begin{aligned}
f_i(\mathbf{x}^*) &\leq 0, \quad i = 1, \ldots, m. \\
h_i(\mathbf{x}^*) &= 0, \quad i = 1, \ldots, p. \\
\lambda_i^* &\geq 0, \quad i = 1, \ldots, m. \\
\lambda_i^* f_i(\mathbf{x}^*) &= 0, \quad i = 1, \ldots, m.
\end{aligned}
$$

$$
\nabla f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(\mathbf{x}^*) = 0
$$

- En cualquier problema de optimización con funciones objetivo y de restricciones diferenciables, para el cual se tenga dualidad fuerte, cualquier par de puntos óptimos del primal y el dual deben satisfacer las condiciones de KKT.

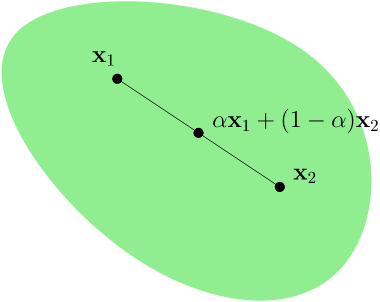
Conjuntos convexos

$$\mathbf{x}_1, \mathbf{x}_2 \in C, \quad 0 \leq \alpha \leq 1 \Rightarrow \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in C$$



Conjuntos convexos

$$\mathbf{x}_1, \mathbf{x}_2 \in C, \quad 0 \leq \alpha \leq 1 \Rightarrow \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in C$$

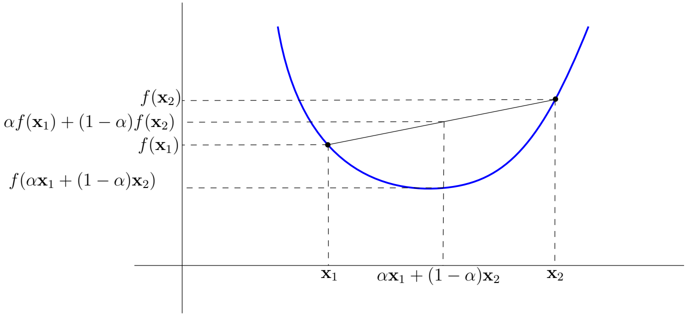


- Intersección de conjuntos convexos es un conjunto convexo.

Funciones convexas

Una función f definida sobre un conjuto convexo Ω es convexa si para todo $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ y para todo $\alpha \in [0,1]$,

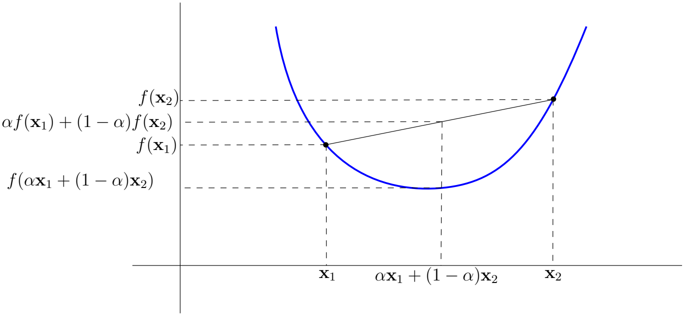
$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$$



Funciones convexas

Una función f definida sobre un conjuto convexo Ω es convexa si para todo $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ y para todo $\alpha \in [0,1]$,

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$$



- f convexa $\Leftrightarrow f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y})$

- f convexa $\Leftrightarrow f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y})$
- f convexa $\Leftrightarrow \nabla^2 f(\mathbf{x})$ es positiva semidefinida.

- f convexa $\Leftrightarrow f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y})$
- f convexa $\Leftrightarrow \nabla^2 f(\mathbf{x})$ es positiva semidefinida.
- f definida sobre un conjuto convexo Ω:

- f convexa $\Leftrightarrow f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y})$
- f convexa $\Leftrightarrow \nabla^2 f(\mathbf{x})$ es positiva semidefinida.
- f definida sobre un conjuto convexo Ω:
  - $\mathbf{x}^*$ es mínimo local $\Leftrightarrow \mathbf{x}^*$ es mínimo global.

- f convexa $\Leftrightarrow f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y})$
- f convexa $\Leftrightarrow \nabla^2 f(\mathbf{x})$ es positiva semidefinida.
- f definida sobre un conjuto convexo Ω:
  - $\mathbf{x}^*$ es mínimo local $\Leftrightarrow \mathbf{x}^*$ es mínimo global.
  - Conjunto de minimizadores es convexo.

- f convexa $\Leftrightarrow f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y})$
- f convexa $\Leftrightarrow \nabla^2 f(\mathbf{x})$ es positiva semidefinida.
- f definida sobre un conjuto convexo Ω:
    - $\mathbf{x}^*$ es mínimo local $\Leftrightarrow \mathbf{x}^*$ es mínimo global.
    - Conjunto de minimizadores es convexo.
    - $\nabla f(\mathbf{x}) = 0 \Rightarrow \mathbf{x}^*$ es mínimo global

- f convexa $\Leftrightarrow f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y})$
- f convexa $\Leftrightarrow \nabla^2 f(\mathbf{x})$ es positiva semidefinida.
- f definida sobre un conjuto convexo Ω:
  - $\mathbf{x}^*$ es mínimo local $\Leftrightarrow \mathbf{x}^*$ es mínimo global.
  - Conjunto de minimizadores es convexo.
  - $\nabla f(\mathbf{x}) = 0 \Rightarrow \mathbf{x}^*$ es mínimo global
- Combinación lineal de funciones convexas con coeficientes positivos es convexa.

Problema de Optimización Convexo

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

$$\text{sujeto a}$$

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$\mathbf{a}_i^T\mathbf{x} = b_i, \quad i = 1, \ldots, p$$

Problema de Optimización Convexo

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

$$\text{sujeto a}$$

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$\mathbf{a}_i^T\mathbf{x} = b_i, \quad i = 1, \ldots, p$$

- $f_0(\mathbf{x})$ es convexa.

Problema de Optimización Convexo

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

$$\text{sujeto a}$$

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$\mathbf{a}_i^T\mathbf{x} = b_i, \quad i = 1, \ldots, p$$

- $f_0(\mathbf{x})$ es convexa.
- $f_i(\mathbf{x})$ son convexas.

Problema de Optimización Convexo

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

$$\text{sujeto a}$$

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$\mathbf{a}_i^T \mathbf{x} = b_i, \quad i = 1, \ldots, p$$

- $f_0(\mathbf{x})$ es convexa.
- $f_i(\mathbf{x})$ son convexas.
- Restricciones de igualdad son funciones afines.

Problema de Optimización Convexo

$$\min_{\mathbf{x}\in\Omega} \quad f_0(\mathbf{x})$$

$$\text{sujeto a}$$

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m$$

$$\mathbf{a}_i^T \mathbf{x} = b_i, \quad i = 1, \ldots, p$$

- $f_0(\mathbf{x})$ es convexa.
- $f_i(\mathbf{x})$ son convexas.
- Restricciones de igualdad son funciones afines.
- El conjunto factible es convexo.

Condiciones de Slater

- En un problema de optimización se tiene dualidad fuerte si

Condiciones de Slater

- En un problema de optimización se tiene dualidad fuerte si
  1. Es convexo.

Condiciones de Slater

- En un problema de optimización se tiene dualidad fuerte si
  1. Es convexo.
  2. Existe un punto estrictamente factible en el interior relativo del dominio de f_0.

Condiciones de Slater

- En un problema de optimización se tiene dualidad fuerte si
  1. Es convexo.
  2. Existe un punto estrictamente factible en el interior relativo del dominio de f_0.

$$f_i(\mathbf{x}) < 0, \quad i = 1, \ldots, m. \quad \mathbf{A}\mathbf{x} = \mathbf{b}.$$

Condiciones de Slater

- En un problema de optimización se tiene dualidad fuerte si
  1. Es convexo.
  2. Existe un punto estrictamente factible en el interior relativo del dominio de f_0.

$$f_i(\mathbf{x}) < 0, \quad i = 1, \ldots, m. \quad \mathbf{Ax} = \mathbf{b}.$$

- Esta última condición se cumple por ejemplo si las restricciones $f_1, \ldots, f_m$ son de la forma $\mathbf{c}_i^T \mathbf{x} \leq d_i$.

- Suponga ahora que el problema primal es convexo.

- Suponga ahora que el problema primal es convexo.
- Suponga que $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ satisfacen KKT:

$$\begin{aligned} f_i(\tilde{\mathbf{x}}) &\leq 0, \quad i = 1, \ldots, m. \\ h_i(\tilde{\mathbf{x}}) &= 0, \quad i = 1, \ldots, p. \\ \tilde{\lambda}_i &\geq 0, \quad i = 1, \ldots, m. \\ \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) &= 0, \quad i = 1, \ldots, m. \end{aligned}$$

$$\nabla f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i \nabla h_i(\tilde{\mathbf{x}}) = 0$$

- Suponga ahora que el problema primal es convexo.
- Suponga que $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ satisfacen KKT:

$$
\begin{aligned}
f_i(\tilde{\mathbf{x}}) \leq 0, &\quad i = 1, \ldots, m. \\
h_i(\tilde{\mathbf{x}}) = 0, &\quad i = 1, \ldots, p. \\
\tilde{\lambda}_i \geq 0, &\quad i = 1, \ldots, m. \\
\tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) = 0, &\quad i = 1, \ldots, m.
\end{aligned}
$$

$$
\nabla f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i \nabla h_i(\tilde{\mathbf{x}}) = 0
$$

- $\tilde{\mathbf{x}}$ es factible,

- Suponga ahora que el problema primal es convexo.
- Suponga que $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ satisfacen KKT:

$$\begin{aligned} f_i(\tilde{\mathbf{x}}) \le 0, &\quad i = 1, \ldots, m. \\ h_i(\tilde{\mathbf{x}}) = 0, &\quad i = 1, \ldots, p. \\ \tilde{\lambda}_i \ge 0, &\quad i = 1, \ldots, m. \\ \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) = 0, &\quad i = 1, \ldots, m. \\ \nabla f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i \nabla h_i(\tilde{\mathbf{x}}) = 0 & \end{aligned}$$

- $\tilde{\mathbf{x}}$ es factible, $L(\mathbf{x}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ es convexo en $\mathbf{x}$,

- Suponga ahora que el problema primal es convexo.
- Suponga que $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ satisfacen KKT:

$$\begin{aligned} f_i(\tilde{\mathbf{x}}) &\leq 0, \quad i = 1, \ldots, m. \\ h_i(\tilde{\mathbf{x}}) &= 0, \quad i = 1, \ldots, p. \\ \tilde{\lambda}_i &\geq 0, \quad i = 1, \ldots, m. \\ \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) &= 0, \quad i = 1, \ldots, m. \end{aligned}$$

$$\nabla f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i \nabla h_i(\tilde{\mathbf{x}}) = 0$$

- $\tilde{\mathbf{x}}$ es factible, $L(\mathbf{x}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ es convexo en $\mathbf{x}$, y $\tilde{\mathbf{x}}$ minimiza $L(\mathbf{x}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$.

- Tenemos:

- Tenemos:

$$g(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) = L(\tilde{\mathrm{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$$

- Tenemos:

$$
\begin{aligned}
g(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) &= L(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \\
&= f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i h_i(\tilde{\mathbf{x}})
\end{aligned}
$$

- Tenemos:

$$
\begin{aligned}
g(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) &= L(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \\
&= f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i h_i(\tilde{\mathbf{x}}) \\
&= f_0(\tilde{\mathbf{x}})
\end{aligned}
$$

- Tenemos:

$$\begin{aligned} g(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) &= L(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \\ &= f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i h_i(\tilde{\mathbf{x}}) \\ &= f_0(\tilde{\mathbf{x}}) \end{aligned}$$

- Es decir $\tilde{\mathbf{x}}$ y $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ tienen brecha de dualidad cero.

- Tenemos:

$$\begin{aligned} g(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) &= L(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \\ &= f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i h_i(\tilde{\mathbf{x}}) \\ &= f_0(\tilde{\mathbf{x}}) \end{aligned}$$

- Es decir $\tilde{\mathbf{x}}$ y $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ tienen brecha de dualidad cero.
- $\tilde{\mathbf{x}}$ y $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ son óptimos para el primal y el dual.

- Tenemos:

$$\begin{aligned} g(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) &= L(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \\ &= f_0(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{p} \tilde{\nu}_i h_i(\tilde{\mathbf{x}}) \\ &= f_0(\tilde{\mathbf{x}}) \end{aligned}$$

- Es decir $\tilde{\mathbf{x}}$ y $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ tienen brecha de dualidad cero.
- $\tilde{\mathbf{x}}$ y $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ son óptimos para el primal y el dual.
- En cualquier problema de optimización convexo con funciones objetivo y de restricciones diferenciables, cualquier par de puntos que satisfagan KKT son óptimos para el primal y el dual.