Two-player zero-sum game

A Game between two players $P$ and $D$

- **Game setting**
  - $\mathcal{X}$: A set of strategies for $P$
  - $\mathcal{Y}$: A set of strategies for $D$
  - Payoff function, $\psi(x, y)$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$

- **Example**
  - Let $\mathcal{X} = \{1, 2\}$, $\mathcal{Y} = \{1, 2\}$
  - Payoff $\psi(x, y) = a_{x,y}$ where $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$

- **Game Rules**:
  - $P$ chooses a strategy $x \in \mathcal{X}$ and $D$ chooses a strategy $y \in \mathcal{Y}$ independently
  - The referee reveals both the strategies simultaneously
  - Game Outcome: Depends on $\psi(x, y)$
Two-player zero-sum game

A Game between two players $P$ and $D$

- Game Outcome:
  \[
  \psi(x, y) > 0 \implies P \text{ pays an amount } \psi(x, y) \text{ to } D \\
  \psi(x, y) < 0 \implies D \text{ pays an amount } -\psi(x, y) \text{ to } P
  \]

- $P$ wishes to minimize payoff to $D$, while $D$ wishes to receive maximum payoff from $P$

- Assume that minimum and maximum exist
Example: Game 1

\[ \mathcal{X} = \{1, 2\}, \quad \mathcal{Y} = \{1, 2\}, \quad \psi(x, y) = a_{x,y}, \quad \text{where} \]

\[ A = \begin{pmatrix} -2 & 1 \\ 2 & -3 \end{pmatrix} \]

**Player P’s strategy**

\[ \min \{ \max_y a_{1,y}, \max_y a_{2,y} \} \]

\[ = \min \{ 1, 2 \} \]

\[ = 1 \]

Choose \( x = 1 \)

**Player D’s strategy**

\[ \max \{ \min_x a_{x,1}, \min_x a_{x,2} \} \]

\[ = \max \{ -2, -3 \} \]

\[ = -2 \]

Choose \( y = 1 \)

\[ \text{min-max} \geq \text{max-min} \]
Example: Game 2

\[ \mathcal{X} = \{1, 2\}, \quad \mathcal{Y} = \{1, 2\}, \quad \psi(x, y) = a_{x,y}, \quad \text{where} \]

\[ A = \begin{pmatrix} -2 & 1 \\ 2 & 3 \end{pmatrix} \]

Player P’s strategy

\[
\min \{ \max_y a_{1,y}, \max_y a_{2,y} \} \\
= \min \{ 1, 3 \} \\
= 1
\]

Choose \( x = 1 \)

Player D’s strategy

\[
\max \{ \min_x a_{x,1}, \min_x a_{x,2} \} \\
= \max \{ -2, 1 \} \\
= 1
\]

Choose \( y = 2 \)

min-max = max-min
Primal Problem

\[
\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)
\]

Solution

Dual Problem

\[
\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y)
\]

The two problems are dual to each other.

For any \( x \in \mathcal{X} \) and \( y \in \mathcal{Y} \)

\[
\min_{x \in \mathcal{X}} \psi(x, y) \leq \psi(x, y) \leq \max_{y \in \mathcal{Y}} \psi(x, y)
\]

\[
\therefore \min_{x \in \mathcal{X}} \psi(x, y) \leq \max_{y \in \mathcal{Y}} \psi(x, y)
\]

\[
\therefore \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) \leq \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)
\]

Weak Duality

\[
\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) \leq \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)
\]
Weak Duality

$$\max_{y \in Y} \min_{x \in X} \psi(x, y) \leq \min_{x \in X} \max_{y \in Y} \psi(x, y)$$

- When does the equality hold?

Definition

Let $x^* \in X$ and $y^* \in Y$. A point $(x^*, y^*)$ is a saddle point for $\psi(x, y)$ if

$$\psi(x^*, y) \leq \psi(x^*, y^*) \leq \psi(x, y^*) \quad \forall x \in X, y \in Y$$

- $x^* = \arg\min_{x \in X} \psi(x, y^*)$
- $y^* = \arg\max_{y \in Y} \psi(x^*, y)$
Theorem

The following equality holds

\[
\max_{y \in Y} \min_{x \in X} \psi(x, y) = \min_{x \in X} \max_{y \in Y} \psi(x, y)
\]

if and only if there exists a saddle point, \((x^*, y^*)\), for \(\psi(x, y)\).

Proof.

(a) Let \((x^*, y^*)\) be a saddle point for \(\psi(x, y)\).

\[
\therefore \psi(x^*, y) \leq \psi(x^*, y^*) \leq \psi(x, y^*) \quad \forall x \in X, y \in Y
\]

\[
\therefore \max_{y \in Y} \psi(x^*, y) \leq \psi(x^*, y^*) \leq \min_{x \in X} \psi(x, y^*)
\]

Note that

\[
\min_{x \in X} \max_{y \in Y} \psi(x, y) \leq \max_{x \in X} \psi(x^*, y)
\]

\[
\min_{x \in X} \psi(x, y^*) \leq \max_{y \in Y} \min_{x \in X} \psi(x, y^*)
\]
Proof. (continued)

Therefore,

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y) \leq \psi(x^*, y^*) \leq \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y)$$

But, we know that

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) \leq \psi(x^*, y^*) \leq \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

Therefore,

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y) = \psi(x^*, y^*) = \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y)$$
Proof. (continued)

(b) Suppose the following equality holds for some $x^* \in \mathcal{X}$, $y^* \in \mathcal{Y}$,

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) = \psi(x^*, y^*) = \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

Now,

$$\max_{y \in \mathcal{Y}} \psi(x^*, y) = \psi(x^*, y^*) = \min_{x \in \mathcal{X}} \psi(x, y^*)$$

$\therefore$ $\psi(x^*, y) \leq \max_{y \in \mathcal{Y}} \psi(x^*, y) = \psi(x^*, y^*) = \min_{x \in \mathcal{X}} \psi(x, y^*) \leq \psi(x, y^*)$

Therefore, $(x^*, y^*)$ is a saddle point for $\psi(x, y)$. □
Strong Duality

\[
\max_{y \in Y} \min_{x \in X} \psi(x, y) = \min_{x \in X} \max_{y \in Y} \psi(x, y)
\]

Consider the problem (NLP):

\[
\begin{align*}
\min & \quad f(x) \\
\text{s.t.} & \quad h_j(x) \leq 0, \ j = 1, \ldots, l \\
& \quad e_i(x) = 0, \ i = 1, \ldots, m
\end{align*}
\]

- Can we define a game with a payoff function \( \psi(\cdot) \) so that the solution to NLP is a solution to the **primal** problem, \( \min_x \max_y \psi(x, y) \)?
- What is the saddle point condition in terms of \( f \), \( h_j \)'s and \( e_i \)'s?
Consider the problem (P):

\[
\begin{align*}
\min & \quad f(x) \\
\text{s.t.} & \quad h_j(x) \leq 0, \ j = 1, \ldots, l \\
& \quad x \in X
\end{align*}
\]

Define a payoff function as the Lagrangian,

\[
\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^{l} \lambda_j h_j(x)
\]

where \( x \in X \) and \( \lambda_j \geq 0, \ j = 1, \ldots, l \)

- \( x \): Primal Variables, \( \lambda \): Dual Variables
- \( X = \mathcal{X}, \mathcal{Y} = \{ \lambda \in \mathbb{R}^l : \lambda_j \geq 0, \ j = 1, \ldots, l \} \)

Duality: Define a \textbf{min max} problem equivalent to the primal problem \( P \). Then, the corresponding dual \textbf{max min} problem is the dual problem \( D \).
Assumption: Minimum and Maximum exist for the problems defined here (Use infimum or supremum appropriately).

**Primal Function**

\[
\text{Primal Function} = \max_{\lambda \geq 0} \mathcal{L}(x, \lambda)
\]

\[
= \max_{\lambda \geq 0} f(x) + \sum_{j=1}^{l} \lambda_j h_j(x)
\]

\[
= \begin{cases} 
  f(x) & \text{if } h_j(x) \leq 0 \ \forall \ j \\
  +\infty & \text{Otherwise.}
\end{cases}
\]

**Primal Problem:**

\[
\min_{x \in X} \max_{\lambda \geq 0} \mathcal{L}(x, \lambda)
\]

That is, (ignoring the possibility of \( h_j(x) > 0 \ \forall \ j \)),

\[
\min f(x) \\
\text{s.t.} \quad h_j(x) \leq 0, \ j = 1, \ldots, l \\
x \in X
\]
For $\lambda \geq 0$, define

**Dual Function**

\[
\begin{align*}
\theta(\lambda) &= \min_{x \in X} \mathcal{L}(x, \lambda) \\
&= \min_{x \in X} f(x) + \sum_{j=1}^{l} \lambda_j h_j(x)
\end{align*}
\]

**Dual Problem:**

\[
\max_{\lambda \geq 0} \min_{x \in X} f(x) + \sum_{j=1}^{l} \lambda_j h_j(x)
\]
Consider the problem:

\[
\min \quad x^2 \\
\text{s.t.} \quad x \geq 1
\]

- Primal solution: \( x^* = 1, \ f(x^*) = 1 \).
- Dual function: \( \theta(\lambda) = \min_x x^2 + \lambda(1 - x) \).
  At the minimum, \( x^* = \frac{\lambda}{2} \).
  For \( \lambda \geq 0 \), \( \theta(\lambda) = -\frac{1}{4}\lambda^2 + \lambda \).
  Therefore, the dual problem is

\[
\max_{\lambda \geq 0} -\frac{1}{4}\lambda^2 + \lambda
\]

- \( \lambda^* = 2, \ \theta(\lambda^*) = 1 \)
- \( f(x^*) = 1 = \theta(\lambda^*) \)
Geometric Interpretation

Consider the problem (P1):

\[
\min_{x \in X} \begin{cases} f(x) \\ h(x) \leq 0 \end{cases} \equiv \min_{x \in X} \max_{\lambda \geq 0} f(x) + \lambda h(x)
\]

Define \( G = \{(y, z) : y = h(x), z = f(x), x \in X\} \).
$G = \{ (y, z) : y = h(x), z = f(x), x \in X \}$
Geometric Interpretation

Consider the problem (P1):

\[
\min_{x \in X} \begin{cases} f(x) \\ h(x) \leq 0 \end{cases} \equiv \min_{x \in X} \max_{\lambda \geq 0} f(x) + \lambda h(x)
\]

Define \( G = \{(y, z) : y = h(x), z = f(x), x \in X\} \).

A solution to the primal problem \textbf{P1} is a point in \( G \) with \( y \leq 0 \) and has minimum ordinate \( z \).
Geometric Interpretation

Consider the problem \((\textbf{P1})\):

\[
\min_{x \in X} \begin{cases}
  f(x) \\
  h(x) \leq 0
\end{cases} 
\equiv \min_{x \in X} \max_{\lambda \geq 0} f(x) + \lambda h(x)
\]

Define \( G = \{(y, z) : y = h(x), z = f(x), x \in X\} \).

A solution to the primal problem \(\textbf{P1}\) is a point in \( G \) with \( y \leq 0 \) and has minimum ordinate \( z \).

Let \((y^*, z^*)\) be this point in \( y - z \) space.

For a given \( \lambda \geq 0 \),

- Define \( \theta(\lambda) = \min_{x \in X} f(x) + \lambda h(x) \).
- \( \theta(\lambda) \) is a minimum \( z + \lambda y \) over feasible \( G \) in \( y - z \) space.
Shirish Shevade
Numerical Optimization
Geometric Interpretation

Consider the problem \((P1)\):

\[
\min_{x \in X} \quad f(x)
\]

s.t.

\[
h(x) \leq 0
\]

\[
\equiv \min_{x \in X} \max_{\lambda \geq 0} f(x) + \lambda h(x)
\]

Define \( G = \{(y, z) : y = h(x), z = f(x), x \in X\} \).

A solution to the primal problem \(P1\) is a point in \(G\) with \(y \leq 0\) and has minimum ordinate \(z\).

Let \((y^*, z^*)\) be this point in \(y - z\) space.

For a given \(\lambda \geq 0\),

- Define \(\theta(\lambda) = \min_{x \in X} f(x) + \lambda h(x)\).
- \(\theta(\lambda)\) is a minimum \(z + \lambda y\) over feasible \(G\) in \(y - z\) space.

Lagrangian Dual Problem \((D1)\):

\[
\max_{\lambda \geq 0} \quad \theta(\lambda) \equiv \max_{\lambda \geq 0} \min_{x \in X} f(x) + \lambda h(x).
\]
\[ \theta(\lambda^*) = (y^*, z^*) \]

\[ z + \lambda y = \beta \]

Slope = $-\lambda^*$
Slope = -\( \lambda_1^\pm \)
Slope = $-\lambda_2^*$
Slope $= -\lambda_3^* = 0$

Slope $= -\lambda_2^*$

Slope $= -\lambda_1^*$
Feasible Set

Optimal Primal Objective
Duality Gap

Optimal Dual Objective
### Primal Problem

\[
\begin{align*}
\text{min} & \quad f(x) \\
\text{s.t.} & \quad h_j(x) \leq 0, j = 1, \ldots, l \\
& \quad e_i(x) = 0, i = 1, \ldots, m \\
& \quad x \in X
\end{align*}
\]

### Dual Problem

\[
\begin{align*}
\text{max} & \quad \theta(\lambda, \mu) \\
\text{s.t.} & \quad \lambda \geq 0
\end{align*}
\]

where \( \theta(\lambda, \mu) = \min_{x \in X} \mathcal{L}(x, \lambda, \mu) \).

### Theorem

Let \( x \) be primal feasible and \( (\lambda, \mu) \) be dual feasible. Then

\[
f(x) \geq \theta(\lambda, \mu).
\]
Primal Problem

\[
\begin{align*}
\text{min} & \quad f(x) \\
\text{s.t.} & \quad h_j(x) \leq 0, j = 1, \ldots, l \\
& \quad e_i(x) = 0, i = 1, \ldots, m \\
& \quad x \in X
\end{align*}
\]

Dual Problem

\[
\begin{align*}
\text{max} & \quad \theta(\lambda, \mu) \\
\text{s.t.} & \quad \lambda \geq 0
\end{align*}
\]

where \( \theta(\lambda, \mu) = \min_{x \in X} \mathcal{L}(x, \lambda, \mu) \).

Proof.

Let \( x \) and \( (\lambda, \mu) \) be primal and dual feasible respectively.

\[
\begin{align*}
\theta(\lambda, \mu) &= \min_{x \in X} \mathcal{L}(x, \lambda, \mu) \\
&= \min_{x \in X} f(x) + \sum_{j=1}^{l} \lambda_j h_j(x) + \sum_{i=1}^{m} \mu_i e_i(x) \\
&\leq f(x)
\end{align*}
\]
Primal Problem

$$\min \quad f(x)$$

s.t.  
$$h_j(x) \leq 0, \quad j = 1, \ldots, l$$
$$e_i(x) = 0, \quad i = 1, \ldots, m$$
$$x \in X$$

Dual Problem

$$\max \quad \theta(\lambda, \mu)$$

s.t.  
$$\lambda \geq 0$$

where
$$\theta(\lambda, \mu) = \min_{x \in X} \mathcal{L}(x, \lambda, \mu).$$

Weak Duality Theorem

Let $p^*$ and $d^*$ be optimal primal and dual objective function values respectively. Let $x$ be primal feasible and $(\lambda, \mu)$ be dual feasible. Then
$$f(x) \geq \theta(\lambda, \mu).$$

$$\therefore \quad \min \{f(x) : h_j(x) \leq 0 \forall j, e_i(x) = 0 \forall i, x \in X\} \geq \max \{\theta(\lambda, \mu) : \lambda \geq 0\}$$

$$\therefore p^* \geq d^*.$$
Example:
Consider the problem:

\[
\begin{align*}
\min & \quad x^3 \\
\text{s.t.} & \quad x = 1 \\
& \quad x \in \mathbb{R}
\end{align*}
\]

- \(x^* = 1, \ f(x^*) = 1\).
- Dual function:

\[
\begin{align*}
\theta(\mu) &= \min_{x \in \mathbb{R}} x^3 + \mu(x - 1) \\
&= \min_{x \in \mathbb{R}} x^3 + \mu x - \mu \\
&= -\infty \quad \forall \ \mu \in \mathbb{R}
\end{align*}
\]

\[
\therefore \theta(\mu^*) = -\infty < f(x^*) \Rightarrow d^* < p^*
\]

\[
\Rightarrow \text{There exists a duality gap.}
\]
Recall the example of two-player zero-sum game.

Example: Game 2

\[ X = \{1, 2\}, \ Y = \{1, 2\}, \ \psi(x, y) = a_{x,y}, \text{ where} \]

\[
A = \begin{pmatrix}
-2 & 1 \\
2 & 3
\end{pmatrix}
\]

Player \( P \)'s strategy

\[
\min\{\max_y a_{1,y}, \max_y a_{2,y}\} = \min\{1, 3\} = 1
\]

Choose \( x = 1 \)

Player \( D \)'s strategy

\[
\max\{\min_x a_{x,1}, \min_x a_{x,2}\} = \max\{-2, 1\} = 1
\]

Choose \( y = 2 \)

\[\min\text{-}\max = \max\text{-}\min\]
Primal Problem

\[
\begin{align*}
\min & \quad f(x) \\
\text{s.t.} & \quad h_j(x) \leq 0, j = 1, \ldots, l \\
& \quad e_i(x) = 0, i = 1, \ldots, m \\
& \quad x \in X
\end{align*}
\]

Dual Problem

\[
\begin{align*}
\max & \quad \theta(\lambda, \mu) \\
\text{s.t.} & \quad \lambda \geq 0
\end{align*}
\]

where \( \theta(\lambda, \mu) = \min_{x \in X} \mathcal{L}(x, \lambda, \mu) \).

Let \( x^* \) and \( (\lambda^*, \mu^*) \) be optimal solutions to the primal and dual problems respectively. Let \( p^* \) and \( d^* \) be optimal primal and dual objective function values respectively.

\[ p^* = d^* \Rightarrow \text{There is no duality gap.} \]

Under what conditions is \( p^* = d^* \)?

Optimal primal and dual objective function values are same (\( p^* = d^* \)) if and only if \( (x^*, \lambda^*, \mu^*) \) is a Lagrangian saddle point, that is, for \( x, x^* \in X \) and \( \lambda, \lambda^* \geq 0 \),

\[
\mathcal{L}(x^*, \lambda, \mu) \leq \mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*).
\]
Proof.

(a)

Let \((x^*, \lambda^*, \mu^*)\) be a Lagrangian saddle point where \(x^* \in X\) and \(\lambda^* \geq 0\). Let \(\lambda \geq 0\).

\[
\mathcal{L}(x^*, \lambda, \mu) \leq \mathcal{L}(x^*, \lambda^*, \mu^*)
\]

\[
\therefore f(x^*) + \sum_{j=1}^{l} \lambda_j h_j(x^*) + \sum_{i=1}^{m} \mu_i e_i(x^*)
\]

\[
\leq f(x^*) + \sum_{j=1}^{l} \lambda_j^* h_j(x^*) + \sum_{i=1}^{m} \mu_i^* e_i(x^*)
\]

\[
\therefore h_j(x^*) \leq 0 \quad \forall j \\
\quad e_i(x^*) = 0 \quad \forall i \quad \text{and } x^* \in X \Rightarrow x^* \text{ is primal feasible}
\]
\[ \mathcal{L}(x^*, \lambda, \mu) \leq \mathcal{L}(x^*, \lambda^*, \mu^*) \]

\[
\therefore \sum_{j=1}^{l} \lambda_j h_j(x^*) + \sum_{i=1}^{m} \mu_i e_i(x^*) \leq \sum_{j=1}^{l} \lambda_j^* h_j(x^*) + \sum_{i=1}^{m} \mu_i^* e_i(x^*)
\]

\[
\therefore \sum_{j=1}^{l} \lambda_j h_j(x^*) \leq \sum_{j=1}^{l} \lambda_j^* h_j(x^*) \quad (\because e_i(x^*) = 0 \ \forall \ i)
\]

\[
\therefore 0 \leq \sum_{j=1}^{l} \lambda_j^* h_j(x^*) \quad (\text{Letting } \lambda_j = 0 \ \forall \ j)
\]

Also, \[ 0 \geq \sum_{j=1}^{l} \lambda_j^* h_j(x^*) \quad (\because \lambda_j^* \geq 0, \ h_j(x^*) \leq 0 \ \forall \ j) \]

\[
\therefore \sum_{j=1}^{l} \lambda_j^* h_j(x^*) = 0 \Rightarrow \lambda_j^* h_j(x^*) = 0 \ \forall \ j
\]
(x*, λ*, µ*) is a saddle point. \( L(x*, λ*, µ*) \leq L(x, λ*, µ*) \).

Therefore, the dual function at (λ*, µ*),

\[
\theta(λ*, µ*) = \min_{x \in X} f(x) + \sum_{j=1}^{l} λ_j^* h_j(x) + \sum_{i=1}^{m} µ_i^* e_i(x)
\]

\[
= \min_{x \in X} L(x, λ*, µ*)
\]

\[
= L(x*, λ*, µ*)
\]

\[
= f(x*) + \sum_{j=1}^{l} λ_j^* h_j(x*) + \sum_{i=1}^{m} µ_i^* e_i(x*)
\]

\[
= f(x*)
\]

∴ \( d^* = p^* \).
Proof. (continued)

(b)

Let \( f(x^*) = \theta(\lambda^*, \mu^*) \). Note that \( x^* \) is primal feasible and \( (\lambda^*, \mu^*) \) is dual feasible. Let \( x \) be primal feasible and \( \lambda_j \geq 0 \ \forall \ j \).

\[
\therefore \theta(\lambda^*, \mu^*) = \min_{x \in X} f(x) + \sum_{j=1}^{l} \lambda_j^* h_j(x) + \sum_{i=1}^{m} \mu_i^* e_i(x)
\]

\[
\leq f(x^*) + \sum_{j=1}^{l} \lambda_j^* h_j(x^*) + \sum_{i=1}^{m} \mu_i^* e_i(x^*)
\]

\[
= f(x^*) + \sum_{j=1}^{l} \lambda_j^* h_j(x^*)
\]

\[
\leq f(x^*) \quad (\therefore \lambda_j^* \geq 0, h_j(x^*) \leq 0)
\]

But, \( \theta(\lambda^*, \mu^*) = f(x^*) \). Therefore, \( \lambda_i^* h_i(x^*) = 0 \ \forall \ j \).
Proof.

\[ \mathcal{L}(x^*, \lambda^*, \mu^*) = f(x^*) + \sum_{j=1}^{l} \lambda_j^* h_j(x^*) + \sum_{i=1}^{m} \mu_i^* e_i(x^*) \]

\[ = f(x^*) \]

\[ = \theta(\lambda^*, \mu^*) \]

\[ = \min_{x \in X} \mathcal{L}(x, \lambda^*, \mu^*) \]

\[ \therefore \mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*) \ldots (1) \]

Also,

\[ \mathcal{L}(x^*, \lambda^*, \mu^*) = f(x^*) \]

\[ \geq f(x^*) + \sum_{j=1}^{l} \lambda_j^* h_j(x^*) + \sum_{i=1}^{m} \mu_i^* e_i(x^*) \]

\[ \therefore \mathcal{L}(x^*, \lambda^*, \mu^*) \geq \mathcal{L}(x^*, \lambda, \mu) \ldots (2) \]

From (1) and (2), \((x^*, \lambda^*, \mu^*)\) is a Lagrangian saddle point. \(\square\)
How to find a saddle point if it exists?
Consider the problem (NLP):

\[
\begin{align*}
\min & \quad f(x) \\
\text{s.t.} & \quad h_j(x) \leq 0, \quad j = 1, \ldots, l \\
& \quad e_i(x) = 0, \quad i = 1, \ldots, m \\
& \quad x \in X
\end{align*}
\]

**Theorem**

Let \( f \) and \( h_j \)'s be continuously differentiable convex functions, \( e_i(x) = a_i^T x - b_i \forall i \) and \( X \) be a convex set. Assume that Slater’s condition holds. Then, \( (x^*, \lambda^*, \mu^*) \) is a KKT point \( \Rightarrow \) \( (x^*, \lambda^*, \mu^*) \) is a Lagrangian saddle point.

If \( x^* \) is primal feasible, \( x^* \in \text{int}(X) \), \( \lambda^* \) is dual feasible and \( (x^*, \lambda^*, \mu^*) \) is a Lagrangian saddle point, then \( (x^*, \lambda^*, \mu^*) \) is a KKT point.
Proof.

\( \mathbf{x}^\ast \) is primal feasible. \( \therefore h_j(\mathbf{x}^\ast) \leq 0 \ \forall \ j \) and \( e_i(\mathbf{x}^\ast) = 0 \ \forall \ i \).

(\( \mathbf{x}^\ast, \lambda^\ast, \mu^\ast \)) is a KKT point. Therefore,

\[
\nabla f(\mathbf{x}^\ast) + \sum_{j=1}^{l} \lambda_j^\ast \nabla h_j(\mathbf{x}^\ast) + \sum_{i=1}^{m} \mu_i^\ast \nabla e_i(\mathbf{x}^\ast) = 0
\]

\( \lambda_j^\ast h_j(\mathbf{x}^\ast) = 0 \ \forall \ j \)

\( \lambda_j^\ast \geq 0 \ \forall \ j \)

\( f \) is convex. Therefore, for all \( \mathbf{x} \in X \),

\[
f(\mathbf{x}) \geq f(\mathbf{x}^\ast) + \nabla f(\mathbf{x}^\ast)^T (\mathbf{x} - \mathbf{x}^\ast). \quad \ldots (3)
\]

Similarly, since every \( h_j \) is convex,

\[
h_j(\mathbf{x}) \geq h_j(\mathbf{x}^\ast) + \nabla h_j(\mathbf{x}^\ast)^T (\mathbf{x} - \mathbf{x}^\ast). \quad \ldots (4)
\]

Every \( e_i \) is an affine function. Therefore,

\[
e_i(\mathbf{x}) = e_i(\mathbf{x}^\ast) + \nabla e_i(\mathbf{x}^\ast)^T (\mathbf{x} - \mathbf{x}^\ast). \quad \ldots (5)
\]
Proof. (continued)

Multiplying (4) by $\lambda_j^*$ and (5) by $\mu_i^*$, adding and using KKT conditions,

$$f(x) + \sum_{j=1}^{l} \lambda_j^* h_j(x) + \sum_{i=1}^{m} \mu_i^* e_i(x)$$

$$\geq f(x^*) + \sum_{j=1}^{l} \lambda_j^* h_j(x^*) + \sum_{i=1}^{m} \mu_i^* e_i(x^*)$$

$$\therefore \mathcal{L}(x, \lambda^*, \mu^*) \geq \mathcal{L}(x^*, \lambda^*, \mu^*) \quad \ldots (6)$$

Also,

$$f(x^*) = f(x^*) + \sum_{j=1}^{l} \lambda_j^* h_j(x^*) + \sum_{i=1}^{m} \mu_i^* e_i(x^*)$$

$$\geq f(x^*) + \sum_{j=1}^{l} \lambda_j h_j(x^*) + \sum_{i=1}^{m} \mu_i e_i(x^*)$$

$$\therefore \mathcal{L}(x^*, \lambda^*, \mu^*) \geq \mathcal{L}(x^*, \lambda, \mu) \quad \ldots (7)$$

Therefore, $(x^*, \lambda^*, \mu^*)$ is a Lagrangian saddle point.
(b) \((x^*, \lambda^*, \mu^*)\) is a Lagrangian saddle point, where \(x^*\) is primal feasible, \(x^* \in \text{int}(X)\) and \(\lambda^*\) is dual feasible. Therefore,

\[
h_j(x^*) \leq 0 \quad \forall \ j \quad \text{and} \quad \lambda^*_j \geq 0 \quad \forall \ j \quad \ldots (8)
\]

and

\[
\mathcal{L}(x^*, \lambda, \mu) \leq \mathcal{L}(x^*, \lambda^*, \mu^*).
\]

\[
\therefore \sum_{j=1}^{l} \lambda_j h_j(x^*) + \sum_{i=1}^{m} \mu_i e_i(x^*) \leq \sum_{j=1}^{l} \lambda^*_j h_j(x^*) + \sum_{i=1}^{m} \mu^*_i e_i(x^*)
\]

\[
\therefore \lambda^*_j h_j(x^*) = 0 \quad \forall \ j \quad \ldots (9)
\]

Also,

\[
\mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*).
\]

\[
\therefore x^* = \arg\min_{x \in X} \mathcal{L}(x, \lambda^*, \mu^*)
\]
Proof.(continued)

\[ \mathbf{x}^* = \arg \min_{x \in X} \mathcal{L}(x, \lambda^*, \mu^*) \]

Note that,

\[ \mathcal{L}(x, \lambda^*, \mu^*) = f(x) + \sum_{j=1}^{l} \lambda_j^* h_j(x) + \sum_{i=1}^{m} \mu_i^* e_i(x). \]

\( \mathcal{L}(x, \lambda^*, \mu^*) \) is a convex function of \( x \) (since \( f \) and \( h_j \)'s are convex functions, \( e_i \)'s are affine functions and \( \lambda_j^* \geq 0 \)). Further, \( x^* \in \text{int}(X) \).

\[ \therefore \nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0 \quad \ldots (10) \]

Therefore, from (8), (9) and (10), we see that \( (x^*, \lambda^*, \mu^*) \) is a KKT point. \[\square\]
Consider the convex programming problem (CP):

\[
\begin{align*}
\min & \quad f(x) \\
\text{s.t.} & \quad h_j(x) \leq 0, \; j = 1, \ldots, l \\
& \quad e_i(x) = 0, \; e_i(x) = a_i^T x - b_i, \; i = 1, \ldots, m \\
& \quad x \in \mathbb{R}^n
\end{align*}
\]

where \(f\) and \(h_j\)'s are continuously differentiable convex functions. Assume that Slater's condition holds.

The Lagrangian function is given by:

\[
\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{j=1}^{l} \lambda_j h_j(x) + \sum_{i=1}^{m} \mu_i e_i(x)
\]

The Dual Problem is:

\[
\max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu)
\]

which is the \textbf{Wolfe Dual} of CP:

\[
\begin{align*}
\max_{x, \lambda, \mu} & \quad \mathcal{L}(x, \lambda, \mu) \\
\text{s.t.} & \quad \nabla_x \mathcal{L}(x, \lambda, \mu) = 0 \\
& \quad \lambda > 0
\end{align*}
\]
Example:

\[
\begin{align*}
\min & \quad (x - 2)^2 \\
\text{s.t.} & \quad 2x + 1 \leq 0 \\
& \quad x \in [-1, 1]
\end{align*}
\]

- Convex Programming Problem
- Slater’s condition holds
- \( x^* = -\frac{1}{2}, \quad p^* = f(x^*) = \frac{25}{4} \)
- Dual function: \( \theta(\lambda) = \min_{x \in [-1, 1]} (x - 2)^2 + \lambda(2x + 1) \)

The Wolfe dual problem is:

\[
\begin{align*}
\max & \quad -\lambda^2 + 5\lambda \\
\text{s.t.} & \quad \lambda \in [1, 3]
\end{align*}
\]

Solution: \( \lambda^* = \frac{5}{2} \)

Optimal Dual Objective Value, \( d^* = \frac{25}{4} = p^* \)
Example:
Consider the problem:
\[
\min \quad x_1^2 + x_2^2 + \ldots + x_n^2 \\
\text{s.t.} \quad x_1 + x_2 + \ldots + x_n = 1
\]
- Convex programming problem
- Slater’s condition holds
- \( x^* = (\frac{1}{n}, \ldots, \frac{1}{n})^T, \ f(x^*) = \frac{1}{n} \)
- \( \mathcal{L}(x, \mu) = x_1^2 + \ldots + x_n^2 + \mu(x_1 + \ldots + x_n - 1) \)
- \( \nabla_x \mathcal{L}(x, \mu) = 0 \implies x_i = -\frac{\mu}{2} \forall i \)

Wolfe dual problem:
\[
\max \quad \mathcal{L}(x, \mu) \\
\text{s.t.} \quad \nabla_x \mathcal{L}(x, \mu) = 0
\]
\[
\equiv \max_{\mu \in \mathbb{R}} -\frac{n}{4}\mu^2 - \mu
\]
Solution to the dual problem: \( \mu^* = -\frac{2}{n} \implies x_i^* = \frac{1}{n} \forall i \)
Example: Consider the *Linear Program* (LP),

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad Ax = b \\
& \quad x \geq 0
\end{align*}
\]

where \( A \in \mathbb{R}^{m \times n} \) and \( \text{rank}(A) = m < n \).

- Convex programming problem
- Slater’s condition holds
- \( \mathcal{L}(x, \lambda, \mu) = c^T x + \mu^T (b - Ax) - \lambda^T x \)
- \( \nabla_x \mathcal{L}(x, \lambda, \mu) = 0 \implies c - A^T \mu - \lambda = 0 \)

**Wolfe dual problem (Dual-LP):**

\[
\begin{align*}
\max & \quad \mathcal{L}(x, \lambda, \mu) \\
\text{s.t.} & \quad \nabla_x \mathcal{L}(x, \lambda, \mu) = 0 \\
& \quad \lambda \geq 0
\end{align*}
\]

\[
\left\{ \begin{array}{c}
\max b^T \mu \\
\text{s.t.} \quad A^T \mu \leq c
\end{array} \right. \equiv
\]

The dual of **Dual-LP** is LP!
Example: Consider the *Quadratic Program*,

\[
\min \quad \frac{1}{2} x^T H x + c^T x \\
\text{s.t.} \quad A x \geq b
\]

where \( H \in \mathbb{R}^{n \times n} \) is a symmetric positive semi-definite matrix and \( A \in \mathbb{R}^{m \times n}, \ \text{rank}(A) = m. \)

\[
\mathcal{L}(x, \lambda) = \frac{1}{2} x^T H x + c^T x + \lambda^T (b - A x)
\]

\[
\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow H x + c - A^T \lambda = 0
\]

Therefore, the **Wolfe dual problem** is,

\[
\max \quad \frac{1}{2} x^T H x + c^T x + \lambda^T (b - A x) \\
\text{s.t.} \quad H x - A^T \lambda = -c \\
\quad \lambda \geq 0.
\]

The dual problem cannot be given explicitly in terms of dual variables.
Example: Consider the Quadratic Program,

\[
\begin{align*}
\min & \quad \frac{1}{2}x^T H x + c^T x \\
\text{s.t.} & \quad Ax \geq b
\end{align*}
\]

where \( H \in \mathbb{R}^{n \times n} \) is a symmetric positive definite matrix.

\[
\mathcal{L}(x, \lambda) = \frac{1}{2}x^T H x + c^T x + \lambda^T (b - Ax)
\]

\[
\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow Hx + c - A^T \lambda = 0
\]

Therefore, the Wolfe dual problem is,

\[
\begin{align*}
\max & \quad \frac{1}{2}x^T H x + c^T x + \lambda^T (b - Ax) \\
\text{s.t.} & \quad Hx + c - A^T \lambda = 0 \\
& \quad \lambda \geq 0.
\end{align*}
\]

Using \( x = H^{-1}(A^T \lambda - c) \), the dual problem is,

\[
\begin{align*}
\max_{\lambda \geq 0} & \quad -\frac{1}{2} \lambda^T A H^{-1} A^T \lambda + (A H^{-1} c + b)^T \lambda
\end{align*}
\]
Example:

\[
\begin{align*}
\min & \sum_{i=1}^{n} x_i \log \left( \frac{x_i}{c_i} \right) \\
\text{s.t.} & \quad Ax = b \\
& \quad x \geq 0
\end{align*}
\]

where \( c_i > 0 \ \forall \ i, \ A \in \mathbb{R}^{m \times n} \) and \( m \ll n \).

- Convex programming problem
- Slater’s condition holds

The Wolfe dual problem is:

\[
\max_{\mu \in \mathbb{R}^m} - \sum_i c_i \exp\{(A^T \mu)_i - 1\} + b^T \mu
\]
Consider the problem (NLP):

\[
\min_{x} f(x) \\
\text{s.t. } h_j(x) \leq 0, \ j = 1, \ldots, l \\
e_i(x) = 0, \ i = 1, \ldots, m \\
\mathbf{x} \in X \ 	ext{where } X \text{ is a compact set.}
\]

- **Dual Function:**

\[
\theta(\lambda, \mu) = \min_{x \in X} f(x) + \sum_{j=1}^{l} \lambda_j h_j(x) + \sum_{i=1}^{m} \mu_i e_i(x)
\]

- Dual function is a pointwise minimum of a family of affine functions of \((\lambda, \mu)\).

\[
\theta(\lambda, \mu) \text{ is a concave function.}
\]

\[
\max_{\lambda} \theta(\lambda, \mu) \\
\text{s.t. } \lambda \geq 0
\]

Therefore, the dual problem is a convex programming problem even if the primal problem is not!