Interpretation: Saddle-point

Example: $f(w, z) = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}$

$z = 1, 2, 3$

$\min_{w \in W} f(w, 1) = 1$

$\min_{w \in W} f(w, 2) = 1$

$\min_{w \in W} f(w, 3) = 1$

$\max_{z \in Z} \min_{w \in W} f(w, z) = 1$

$\max_{z \in Z} f(1, z) = 3$

$\max_{z \in Z} f(2, z) = 3$

$\max_{z \in Z} f(3, z) = 3$

$\min_{w \in W} \max_{z \in Z} f(w, z) = 3$

$\max_{w \in W} \min_{z \in Z} f(w, z) = 3$

---

Interpretation: Saddle-point

Example: $f(w, z) = \begin{bmatrix} 4 & 6 & 3 \\ 2 & 3 & 1 \\ 3 & 5 & 2 \end{bmatrix}$

$z = 1, 2, 3$

$\omega = 2$

$\max_{z \in Z} \min_{w \in W} f(w, z) = 3$

$\min_{w \in W} \max_{z \in Z} f(w, z) = 3$

$\min_{w \in W} f(2, z) = 3$

$\max_{w \in W} \min_{z \in Z} f(w, z) = 3$

\text{W 7A midterm W 7B}
Interpretation: Saddle-point

Claim: Result of II $\geq$ Result of I

Given an arbitrary pair $(\tilde{w}, \tilde{z}) \in D$

\[
\min_{w \in W} f(w, \tilde{z}) \leq f(\tilde{w}, \tilde{z}) \leq \max_{z \in Z} f(\tilde{w}, z) \quad \forall \tilde{w}, \tilde{z} \in D
\]

\[
\min_{w \in W} f(w, \tilde{z}) \leq \max_{z \in Z} f(\tilde{w}, z)
\]

Thus \[
\max_{z \in Z} \min_{w \in W} f(w, z) \leq \min_{w \in W} \max_{z \in Z} f(w, z)
\]

Interpretation: Saddle-point

Example: $f(w, z)$ $w = 2$ $2$ $2$ $2$

$3$ $3$ $3$ $3$

$z = 1, 2, 3$

\[
\min_{w \in W} f(w, 1) = 1 \quad \max_{z \in Z} f(1, z) = 1
\]

\[
\min_{w \in W} f(w, 2) = 1 \quad \max_{z \in Z} f(2, z) = 2
\]

\[
\min_{w \in W} f(w, 3) = 1 \quad \max_{z \in Z} f(3, z) = 3
\]

\[
\min_{z \in Z} \min_{w \in W} f(w, z) = 1 \quad \min_{w \in W} \max_{z \in Z} f(w, z) = 1
\]
\[ g(\lambda) = \min_{x} f_0(x) + \lambda u \]

\[ f_0(x) = (-\lambda u) \]

\[ g(\lambda) = \max_{\lambda \leq \lambda_3} \min_{x} f_0(x) + \lambda u \]
Primal
\[ \min f_0(x) \]
subject to
\[ \begin{align*}
  f_i(x) &\leq 0 \\
  h_i(x) &= 0
\end{align*} \]
\[ \min \max L(x, \lambda, u) \]

Dual
\[ \max f(x) \]
\[ \min \max L(x, \lambda, u) \]
\[ \max \min L(x, \lambda, u) \]
\[ \max \min L(x, \lambda, u) \]
\[ \min L \text{ for } x \in D \]

The constraints are enforced.

\[ L(x, \lambda, u) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \frac{p}{2} \sum_{i=1}^{m} h_i(x) \]

Let
\[ f(a, b) = \min \max f(w, z), \quad f(c, d) = \max \min f(w, z) \]

Then
\[ f(a, b) \geq f(a, d) \geq f(c, d) \]

\[ f(w, z) \text{ is convex w.r.t. } w \]

\[ f(w, z) \text{ is concave w.r.t. } z \]

A.21.0
I. Saddle Point

Given function $f(\omega, z)$

$(\bar{\omega}, \bar{z})$ \text{ is a saddle point of } f

if

$$\max_{z} f(\bar{\omega}, z) = f(\bar{\omega}, \bar{z})$$

$$\min_{\omega} f(\omega, \bar{z}) = f(\bar{\omega}, \bar{z})$$

\[ \text{II. Theorem: } \max_{z} \min_{\omega} f(\omega, z) = \min_{\omega} \max_{z} f(\omega, z) \]

iff a saddle point of f exists.
Convexity of $f_0(x)$
Proof: Necessity

Assume that
\[ \min_w \max_z f(w, z) = \max_w \min_z f(w, z) \]

Let \( \bar{w} = \arg \min_w \max_z f(w, z) \)
\( \bar{z} = \arg \max_z \min_w f(w, z) \)

We have
\[ f(\bar{w}, \bar{z}) \leq \max_z f(\bar{w}, z) = \min_w f(w, \bar{z}) \leq f(\bar{w}, \bar{z}) \]

By definition, \((\bar{w}, \bar{z})\) is a saddle point.

Sufficiency

Assume that \((\bar{w}, \bar{z})\) is a saddle point

We have
\[ \max_z \min_w f(w, z) = \min_w f(w, \bar{z}) = f(\bar{w}, \bar{z}) \]
\[ \min_w \max_z f(w, z) \leq \max_z f(\bar{w}, z) = \bar{f}(\bar{w}, \bar{z}) \]
Thus,
\[ \max_z \min_w f(w, z) = \min_w \max_z f(w, z) \]
Formulation: The row & column selection is formulated as a bilinear optimization problem.

\[
\min_{\omega} \max_{z} f(\omega, z) = \sum_{i,j} a_{ij} w_i z_j \quad \text{maximize} \ \frac{f(\omega, z)}{z \omega}
\]

1. row & column selection constraints
   \(a_{ij} \in \{0, 1\}\) \(\sum w_i = 1\) \(\sum z_j = 1\)

2. relaxed constraints
   \(\sum w_i = 1\) \(\sum z_j = 1\) \(w_i \geq 0, z_j \geq 0, \forall i, j\)

A. The optimization problem with relaxed constraints can be solved with algorithms Dantzig

\[
\min_{\omega} \max_{z} f(\omega, z) = \max_{z} \min_{\omega} f(\omega, z)
\]

B. Since \(f(\omega, z)\) is convex w.r.t \(\omega\), concave w.r.t \(z\).
   The solution can reduce to constraint 1 (row & column selection)

\[f(\omega, z)\]

C. From B, \((\omega, z)\) is a saddle point.