21. Duality

- Consider all hyperplanes that lie below \( w(z) \).

\[
\mathbf{w}(z) = \max \{ \text{y-intercept of all hyperplanes that lie below } w(z) \} 
\]

- Thus, \( \mathbf{w}(z) \) (which is a min problem) can be converted to a max problem involving the y-intercepts of hyperplanes.

- The max y-intercept is achieved by the hyperplane determined
by the Lagrange multipliers $z_0$.

This is the essence of duality theory.

Consider

$$\inf_{x \in \mathcal{X}} f(x) = \mu_0$$

s.t.

$$g(x) \leq 0$$

Define the dual functional:

$$\varphi(g^*) = \min_{x \in \mathcal{X}} f(x) + \langle g(x), g^* \rangle$$

on the positive cone $\mathcal{P}^+ \ (we\ will\ say\ g^* \geq 0)$

Theorem:

1. $\varphi(g^*) \leq \mu_0$. Thus,

$$\inf_{g^* \geq 0} \varphi(g^*) \leq \mu_0$$

2. $\varphi(g^*)$ is concave.
Note: No condition are needed on $f, g, \Omega$.

Proof:

(1) \[ f(x) + \langle g(x), z^* \rangle \leq f(x) \quad \text{since} \quad z^* \in \Phi^+ \quad \text{and} \quad g(x) \leq 0 \quad \text{w.r.t.} \quad \Phi \]

Thus, $\Phi(z^*) = \min_{x \in \Omega} f(x) + \langle g(x), z^* \rangle$

\[ \leq \min_{x \in \Omega} f(x) + \langle g(x), z^* \rangle \quad \text{if} \quad g(x) \leq 0 \]

\[ = \min_{x \in \Omega} f(x) = \mu_0 \quad \text{if} \quad g(x) \leq 0 \]

(2) \[ \Phi(\alpha z_1^* + (1-\alpha) z_2^*) = \min_{x \in \Omega} f(x) + \langle g(x), \alpha z_1^* + (1-\alpha) z_2^* \rangle \]

\[ = \min_{x \in \Omega} \alpha (f(x) + \langle g(x), z_1^* \rangle) \]

\[ + (1-\alpha) (f(x) + \langle g(x), z_2^* \rangle) \]
\[ \alpha \varphi(g_1^+) + (1-\alpha) \varphi(g_2^+) \]

**Theorem:** Assume the conditions in the necessary part of the global theory.

(i) Then,

\[ \mu_0 := \inf_{x \in \mathcal{A}, \varrho(x) \leq 0} f(x) = \max_{g^+ \geq 0} \varphi(g^+) \]

and the max on the RHS is achieved at \( g_0^+ \geq 0 \).

(ii) If the inf on the LHS is achieved at some \( x_0 \), then

\[ \langle \varrho(x_0), g_0^+ \rangle = 0 \]

and

\[ \mu_0 = \inf_{x \in \mathcal{A}} f(x) + \langle \varrho(x), g_0^+ \rangle. \]

**Proof:** (i) We already showed that \( \varphi(g^+) \leq \mu_0 \). The global theory established the
existence of \( \xi^* \) achieving \( \mu_0 \).

Thus, (i) is proved.

(ii) Established in global theory.

The following theorem makes the connection between \( \varphi(\xi^*) \) and the figure we drew at the beginning of this lecture.

**Theorem:**

\[
\varphi(\xi^*) = \inf_{\xi \in \Gamma} \omega(\xi) + \langle \xi, \xi^* \rangle
\]

where \( \Gamma = \{ \xi : \exists x \in \mathbb{R} \text{ s.t. } C(x) \leq \xi \} \)

**Proof:**

\[
\varphi(\xi^*) = \inf_{x \in \mathbb{R}} f(x) + \langle C(x), \xi^* \rangle
\]

\[
= \inf_{x \in \mathbb{R}} f(x) + \langle C(x) - \xi, \xi^* \rangle + \langle \xi, \xi^* \rangle
\]

\[
\leq \inf_{\xi \in \Gamma} f(x) + \langle \xi, \xi^* \rangle
\]
\[ \Rightarrow \psi(\mathbf{z}^*) \leq \inf_{\mathbf{z} \in \Gamma} \omega(\mathbf{z}) + \langle \mathbf{z}, \mathbf{z}^* \rangle \]

Conversely, \( \forall x_i \in \mathcal{L} \)

\[ f(x_i) + \langle c(x_i), \mathbf{z}^* \rangle \]

\[ \geq \inf_{x \in \mathcal{L}} f(x) + \langle c(x), \mathbf{z}^* \rangle \]

\[ = \inf_{x \in \mathcal{L}} f(x) + \langle \mathbf{z}, \mathbf{z}^* \rangle \]

\[ c(x) = c(x_i) \]

\[ \geq \inf_{x \in \mathcal{L}} f(x) + \langle \mathbf{z}, \mathbf{z}^* \rangle \]

\[ = \inf_{x \in \mathcal{L}} f(x) + \langle \mathbf{z}, \mathbf{z}^* \rangle \]

\[ = \omega(c(x_i)) + \langle c(x_i), \mathbf{z}^* \rangle \]

\[ \geq \inf_{\mathbf{z} \in \Gamma} \omega(\mathbf{z}) + \langle \mathbf{z}, \mathbf{z}^* \rangle \]

\[ \Rightarrow \psi(\mathbf{z}^*) \geq \inf_{\mathbf{z} \in \Gamma} \omega(\mathbf{z}) + \langle \mathbf{z}, \mathbf{z}^* \rangle \]

D
Back to the figure in $\mathbb{R}^n$

1. $\phi(\lambda) = \inf_{z \in P} w(z) + \lambda^T z$, $\lambda \geq 0$

$\gamma + \lambda^T z = k$ is a hyperplane

1. States that this hyperplane supports the region above the curve $w$ when $k = \phi(\lambda)$.

When $z = 0$, $\gamma = \phi(\lambda)$ for this hyperplane, so $\phi(\lambda)$ is the $y$-intercept of this hyperplane.