\[
\inf_{x \in \mathbb{R}^n} f(x)
\]

\[ (\star) \]

Assumptions:

(i) \( x \in \mathbb{R}^n \)

(ii) \( f \) is a convex functional:

\[
f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)
\]

(iii) \( G : \mathbb{R}^n \to \mathbb{R}^m \) “\( \leq \)” means

\( \text{elementwise} \leq \).

(iv) \( \Omega \) is a convex set

(v) Let \( \mu_0 = \inf f(x) < \infty \)

(vi) \( \exists x, : \: G(x) \leq 0 \), where

“\( \leq \)” is again element wise.

(vii) Each element of \( G \) is a convex function.
**Theorem:** \( \exists \lambda > 0 \text{ s.t. } \mu_0 = \inf_{x \in \mathbb{R}} f(x) + \lambda g(x). \)

If \( x_0 \) solves \( \Omega \), then

\[ \nabla g(x_0) = 0 \]

**Proof:** Define

\[ w(z) = \inf_{x \in \mathbb{R}} f(x) \quad \text{for some } x \in \Omega \]

Under our assumptions, it is easy to show that \( w(z) \) is a convex function.

Let

\[ A = \{(x, z) : y \geq f(x), \quad \exists x \in \Omega \} \]

\[ B = \{(x, z) : y \leq \mu_0, \quad z \leq 0 \} \]

It is easy to show the following:

(i) \( A \) and \( B \) are convex sets.

(ii) \( A \cap B = \emptyset \)

(iii) \( B \) is non-empty.
Applying the separating hyperplane theorem, \( \exists (x_0, \lambda_0), c \) s.t.

\[
\gamma x_0 + \lambda_0 g \leq c \quad \forall (x, g) \in B
\]

\[
\gamma x_0 + \lambda_0 g \geq c \quad \forall (x, g) \in A.
\]

We now argue that \( x_0 > 0, \lambda_0 > 0 \)

Suppose \( x_0 < 0 \), \((-M, 0) \in B \neq \mathbb{R}^n - \mu_0 \)

\[ -M y \leq c \quad \forall -M < \mu_0 \]

Choose \( M \) very large, then \(-M y \)

is a large positive number violating the above inequality. Thus, \( x_0 > 0 \).

Next, suppose \( \lambda_0(i) < 0 \), where \( \lambda_0(i) \) is the \( i \)th component of \( \lambda_0 \).

Choose \( z_i \) to be a "large" negative
number and again we get a contradiction. Thus, \( \lambda > 0 \).

Next, we show that \( \gamma > 0 \).

Since \((\mu_0, 0) \in B\)

\[ \gamma_0 \mu_0 \leq \gamma \gamma_0 + \lambda^T \eta \not\in \{ \gamma, \beta \} \in A \]

If \( \gamma_0 = 0 \), then

\[ 0 \leq \lambda_0 \eta \]

Also \( \lambda \not= 0 \), since \((\gamma_0, \lambda_0) \not= 0 \)

for the separating hyperplane to exist. Next, note that \((f(x), C(x)) \in A\).

This implies

\[ 0 \leq \lambda_0^T C(x) \not\leq x \]

since \( \exists x_1 \) s.t. \( C(x_1) < 0 \) and \( \lambda_0 \not= 0 \).

\[ \lambda_0 \not= 0, \lambda_0^T C(x_1) < 0 \] which is a contradiction. \( \Rightarrow \gamma_0 > 0 \).

Dividing by \( \gamma_0 \) and redefining

\[ \lambda_0 \leq \frac{\lambda_0}{\gamma_0}, \quad C \leq \frac{C}{\gamma_0}, \]
\[ y + \lambda^T z \leq c \quad \forall \ (y, z) \in A \]
\[ y + \lambda^T z > c \quad \forall \ (y, z) \in B \]

\[ y_B + \lambda^T z_B \leq y_A + \lambda^T z_A \quad \forall \ (y_b, z_B) \in B \]
\[ (y_A, z_A) \in A \]

\[ \text{Since} \quad (\mu_0, 0) \in B \]
\[ \mu_0 \leq y_A + \lambda^T z_A \quad \forall \ (y_A, z_A) \in A \]
\[ \Rightarrow \mu_0 \leq \inf \ y - \lambda^T z \]
\[ (y, z) \in A \]

\[ \leq \inf \ f(x) + \lambda^T g(x) \]
\[ x \in \Omega \]
\[ g(x) \leq 0 \]

\[ \leq \inf \ f(x) = \mu_0 \]
\[ x \in \Omega \]
\[ g(x) \leq 0 \]

(a) follows from the fact that
\[ \{ f(x), g(x) \} : x \in \Omega, g(x) \leq 0 \quad \forall \ y \in A \]

This proves the first part of the
Theorem. If \( x_0 \) achieves the min in (1), then by the first part of the theorem

\[
\mu_0 = \inf_{x \in \Omega} f(x) + \lambda_0^T a(x)
\]

\[
\leq f(x_0) + \lambda_0^T a(x_0)
\]

\[
\leq f(x_0) \quad (\text{since } a(x_0) \leq 0)
\]

\[
= \mu_0
\]

\[
\Rightarrow \lambda_0^T a(x_0) = 0
\]

\[\square\]

Saddle-point theorem:

Define

\[
L(x, \lambda) = f(x) + \lambda^T a(x)
\]

Then, \( \exists \lambda_0 \) s.t.

\[
L(x_0, \lambda) \leq L(x_0, \lambda_0) \leq L(x, \lambda_0)
\]

\( \forall x_0 \in \Omega, \lambda \geq 0 \)

Proof: The RHS inequality is true...
from the previous theorem. To see the LHS, recall that $\lambda^T A(x_0) = 0$.

\[
L(x_0, \lambda) = f(x_0) + \lambda^T A(x_0)
\]

since $\lambda \geq 0$ and $A(x_0) \leq 0$.

$\square$

**Sufficiency:**

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

--- (2)

- s.t. $A(x) \leq 0$

**Theorem:** If $\int_0^\infty \rho x^T A(x) dx = 0$ and $f(x) + x^T A(x) \leq f(x) + x^T A(x)$

then $x_0$ solves (2).

(Note: no condition are needed on $f$, $A$).

**Proof:** Suppose $\int_0^1 \rho \leq 0$ s.t.

\[
f(x_1) \leq f(x_0) \quad \text{and} \quad A(x_1) \leq 0
\]
Then,
\[ f(x_i) + \lambda^T G(x_i) < f(x_0) + \lambda^T G(x_0) \]
\[ \leq f(x_0) \quad \text{since } \lambda^T G(x) \leq 0 \]
\[ = f(x_0) + \lambda^T G(x_0) \]
which contradicts the assumption of the theorem.

Saddle point sufficiency:
Suppose there exists a \( \lambda_0 \geq 0 \) s.t.
the SP property holds:
\[ L(x_0, \lambda) \leq L(x_0, \lambda_0) \leq L(x, \lambda_0), \quad \forall \lambda \geq \lambda_0, \quad x \in \mathbb{R}^n \]
Then \( x_0 \) solves (2).

Proof: By the LHS inequality,
\[ \lambda^T G(x_0) \leq \lambda_0^T G(x_0) \quad \forall \lambda \geq \lambda_0 \]
\[ \lambda^T G(x_0) = (\lambda + \lambda_0)^T G(x_0) - \lambda_0^T G(x_0) \]
\[ \leq \lambda_0^T G(x_0) - \lambda_0^T G(x_0) = 0 \]
Hence \( \lambda + \lambda_0 \geq 0 \) and apply above inequality.
Thus, $\lambda^T \mathcal{A}(x_0) \leq 0 \quad \forall \lambda \geq 0$

$\Rightarrow \quad \mathcal{A}(x_0) \leq 0$

$\Rightarrow \quad x_0$ is a feasible point of (2).

$f(x_0) \geq f(x_0) + \lambda^T \mathcal{A}(x_0) \geq f(x_0) + \lambda^T \mathcal{A}(x_0)$ (SP condition

LHS) $\forall \lambda \geq 0$

Letting $\lambda = 0$,

$f(x_0) \geq f(x_0) + 0 = f(x_0)$

$\Rightarrow \quad \lambda^T \mathcal{A}(x_0) = 0$

Now the result follows from the previous theorem. $\blacksquare$

Arbitrary normed spaces:

- $f : x \to \mathbb{R}$, convex.
- $\Omega$ : convex
- $G : x \to \mathbb{Z}$, what does

$G(x) \leq 0$ and $G$ is
Convex mean. Let $P \subseteq \mathbb{R}$ be a convex cone: $P$ is a convex set and $\cdot \in P$ is a cone, i.e., $x \in P \Rightarrow \alpha x \in P \quad \forall \alpha \geq 0$.

Two vectors $\mathbf{g}_1, \mathbf{g}_2 \in \mathbb{R}$ are said to satisfy $\mathbf{g}_1 \geq \mathbf{g}_2$ if $\mathbf{g}_1 - \mathbf{g}_2 \in P$. A cone defining $\geq$ relation as above is called a positive cone.

If $f(x) \leq 0$ and $f$ is a convex mapping and defined w.r.t. some $P$. It is easy to see that $x \leq y$, $y \leq z \Rightarrow x \leq z$ and $x \leq \mathbf{0}$ since $\mathbf{0} \in P$.

Note: In $\mathbb{R}^n$, $P$ is taken to be positive orthant.
- What does there exist $x$ s.t. $c(x) < 0$ mean?

$c(x)$ is an interior point of $P$.

- $\mathbf{x}^T \mathbf{z}$ is replaced by $\langle z, z^* \rangle$.

$z^* \geq 0$ is replaced by $z^* \in P^\circ$ where

$P^\circ = \{ z^* : \langle z, z^* \rangle \geq 0 \}
\quad \forall z \in P \}.$

**Note:**

(i) If $f$ is a convex functional over a convex set $C$, then any local min of $f$ over $C$ is also a global min.

Easy to prove. See book.
(2) Under the necessary conditions of the global theory, if
\[ L(x, z) = f(x) + \langle g(x), z^* \rangle \]
is differentiable w.r.t. \( x \), then the necessary condition for \( x_0 \) to be an optimum can be stated as follows: \( \exists z^* \) s.t. \( x_0 \) is a stationary point of \( L(x, z^*) \) and \( \langle g(x_0), z^* \rangle = 0 \). This is called the Karush-Kuhn-Tucker theorem.