Chapter 6

Convex Programming

- Recall that
  - ◇ A set K is said to be *convex* if
$$\alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in K, \quad \text{whenever } \mathbf{x}, \mathbf{y} \in K \text{ and } 0 \le \alpha \le 1.$$
  - ◇ A real-valued functional $\phi : X \to \mathbb{R}$ is said to be a *convex functional* on a convext subset K if
$$\phi(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \le \lambda\phi(\mathbf{x}) + (1-\lambda)\mathbf{y}$$
for each $\mathbf{x}, \mathbf{y} \in K$ and $0 \le \lambda \le 1$.
- Suppose f is a convex functional over a convex set C. Let $\mu = \inf_{\mathbf{x}\in K} f(\mathbf{x})$. Then
  - ◇ $\Omega := \{\mathbf{x} | f(\mathbf{x}) = \mu\}$ is a convex set.
  - ◇ If $\mathbf{x}_0$ is a local minimum, then $f(\mathbf{x}_0) = \mu$.
  - ◇ A local minimum is always a global minimum.

Partial Ordering by Cone

The convex theory can be interpreted geometrically.

- The minimization can be considered as contours of f.
- The Lagrange multiplier can be considered as an element in the dual space X^* and therefore is a hyperplane.
- By introducing a cone to define *positive* vectors in X, it is possible to consider *inequality problems* in an abstract space.

Positive Cones

- Given any convex cone $P \subset X$, define

$$\mathbf{x} \geq \mathbf{y} \quad \text{if and only if} \quad \mathbf{x} - \mathbf{y} \in P.$$

  - $\mathbf{x} \geq \mathbf{x}$ because $\mathbf{0} \in P$.
  - $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{y} \geq \mathbf{z}$ imply that $\mathbf{x} \geq \mathbf{z}$.

- The above definition generalizes the notion of inequality between vectors.
  - Let $P \subset Z$ specify a positive cone. A mapping $G : \Omega \subset X \to Z$ is said to be *convex* if and only if Ω is convex in X and

$$G(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha G(\mathbf{x}_1) + (1-\alpha)G(\mathbf{x}_2), \quad \text{for all } \mathbf{x}_1, \mathbf{x}_2 \in X, 0 < \alpha < 1.$$

  - Convexity is not an intrinsic property of a mapping but is dependent on the specified positive cone in the range space.
  - If G is convex, then for every $\mathbf{z} \in Z$ the set

$$\{\mathbf{x} \in \Omega | G(\mathbf{x}) \leq \mathbf{z}\}$$

    is convex.

Lagrange Multiplier in Abstract Space

- Consider the basic convex programming problem:

$$\begin{array}{ll} \text{Minimize} & f(\mathbf{x}), \\ \text{Subject to} & \mathbf{x} \in \Omega, \quad G(\mathbf{x}) \leq \mathbf{0}, \end{array}$$

  where

  - ◇ Ω is a convex subset in X.
  - ◇ f is a real-valued convex functional on Ω.
  - ◇ G is a convex mapping from Ω into Z with a positive cone P.

- For an arbitrary $\mathbf{z} \in Z$, consider the associate problem:

$$\begin{array}{ll} \text{Minimize} & f(\mathbf{x}), \\ \text{Subject to} & \mathbf{x} \in \Omega, \quad G(\mathbf{x}) \leq \mathbf{z}, \end{array}$$

  and the set $\Gamma \subset Z$ where the feasible set is not empty, i.e.,

$$\Gamma := \{\mathbf{z} \in Z | \text{There exists at least one } \mathbf{x} \in \Omega \text{ such that } G(\mathbf{x}) \leq \mathbf{z}\}.$$

  - ◇ Γ is convex.
  - ◇ $\mathbf{0} \in \Gamma$, if the original problem has feasible set.

- Define the *primal functional* $\omega : \Gamma \to \mathbb{R}$ by

$$\omega(\mathbf{z}) := \inf\{f(\mathbf{x}) | \mathbf{x} \in \Omega, G(\mathbf{x}) \leq \mathbf{z}\}.$$

  - ⋄ The original problem is equivalent to determining $\omega(\mathbf{0})$.
  - ⋄ What is the meaning of ω?

- The entire theory is based on the functional ω:

  - ⋄ ω is a convex functional on Γ, because

$$\begin{aligned}
\omega(\alpha\mathbf{z}_1 + (1-\alpha)\mathbf{z}_2) &= \inf\{f(\mathbf{x}) | \mathbf{x} \in \Omega, G(\mathbf{x}) \leq \alpha\mathbf{z}_1 + (1-\alpha)\mathbf{z}_2\} \\
&\leq \inf\{f(\mathbf{x}) | \mathbf{x} = \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_2 \in \Omega, G(\mathbf{x}_1) \leq \mathbf{z}_1, G(\mathbf{x}_2) \leq \mathbf{z}_2\} \\
&\leq \alpha \inf\{f(\mathbf{x}_1) | \mathbf{x}_1 \in \Omega, G(\mathbf{x}_1) \leq \mathbf{z}_1\} \\
&\quad (1-\alpha) \inf\{f(\mathbf{x}_2) | \mathbf{x}_2 \in \Omega, G(\mathbf{x}_2) \leq \mathbf{z}_2\} \\
&\leq \alpha\omega(\mathbf{z}_1) + (1-\alpha)\omega(\mathbf{z}_2).
\end{aligned}$$

  - ⋄ If $\mathbf{z}_1 \geq \mathbf{z}_2$, then $\omega(\mathbf{z}_1) \leq \omega(\mathbf{z}_2)$. That is, ω is *decreasing*.

General Result

- Assume in the basic convex programming problem that
  - ◇ The positive cone P contains an interior point.
  - ◇ $G : \Omega \to Z$ is convex with respect to P.
  - ◇ There exists a point $\mathbf{x}_1 \in \Omega$ such that $G(\mathbf{x}_1) < \mathbf{0}$.
  - ◇ $\mu_0 := \inf\{f(\mathbf{x}) | \mathbf{x} \in \Omega, G(\mathbf{x}) \leq \mathbf{0}\}$ is finite.
- Then
  - ◇ There is an element $\mathbf{z}_0^* \in Z^*$ such that
$$\mu_0 = \inf_{\mathbf{x}\in\Omega}\{f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}_0^*\rangle\}.$$
  - ◇ $\mathbf{z}_0^* \geq \mathbf{0}$ with respect to the positive cone
$$P^{\oplus} := \{\mathbf{x}^* | \langle \mathbf{x}, \mathbf{x}^*\rangle \geq 0, \text{ for all } \mathbf{x} \in P\}.$$
  - ◇ If the infimum of f is achieved by $\mathbf{x}_0 \in \Omega$, $G(\mathbf{x}_0) \leq \mathbf{0}$, then
$$\langle G(\mathbf{x}_0), \mathbf{z}_0^*\rangle = 0.$$
- Note that the Lagrange multiplier is played out by $\mathbf{z}_0^*$ in the dual space Z^* and that
$$\langle G(\mathbf{x}), \mathbf{z}_0^*\rangle \leftrightarrow \underbrace{\mathbf{z}_0^*}_{\lambda} G(\mathbf{x}).$$

Proof

- In the space $\mathbb{R} \times Z$, define

$$\begin{aligned} A &:= \{(r, \mathbf{z}) \in \mathbb{R} \times Z | f(\mathbf{x}) \leq r, \underbrace{G(\mathbf{x}) \leq \mathbf{z} \text{ for some } \mathbf{x} \in \Omega}_{\mathbf{z} \in \Gamma}\}, \\ B &:= \{(r, \mathbf{z}) \in \mathbb{R} \times Z | r \leq \mu_0, \mathbf{z} \leq \mathbf{0}\}. \end{aligned}$$

  - ◇ Both A and B are convex due to the convexity of f and G.
  - ◇ A contains no interior point of B due to the definition of μ_0.
  - ◇ B contains an interior point due to the assumption P has an interior point.

- By the Separation Hyperplane Theorem, there exists a nonzero element $(r_0, \mathbf{z}_0^*) \in \mathbb{R} \times Z^*$ such that

$$r_0 r_A + \langle \mathbf{z}_A, \mathbf{z}_0^* \rangle \geq r_0 r_B + \langle \mathbf{z}_B, \mathbf{z}_0^* \rangle$$

  for all $(r_A, \mathbf{z}_A) \in A$, $(r_B, \mathbf{z}_B) \in B$.

  - ◇ $r_0 \geq 0$ and $\mathbf{z}_0^* \geq \mathbf{0}$. (Why?)
    - ▷ Take $\mathbf{z}_A = \mathbf{z}_B = \mathbf{0}$. Since $r_A \geq r_B$, $r_0 \geq 0$.
    - ▷ Take $\mathbf{z}_A = \mathbf{0}$. The $\langle \mathbf{z}_B, \mathbf{z}_0^* \rangle \leq 0$ for all $\mathbf{z}_B \leq \mathbf{0}$.
  - ◇ $r_0 > 0$.
    - ▷ $(\mu_0, \mathbf{0}) \in B \Rightarrow r_0 r_A + \langle \mathbf{z}_A, \mathbf{z}_0^* \rangle \geq r_0 \mu_0$ for all $(r_A, \mathbf{z}_A) \in A$.
    - ▷ If $\mu_0 = 0$, then $\langle G(\mathbf{x}_1), \mathbf{z}_0^* \rangle \geq 0$ and $\mathbf{z}_0^* \neq \mathbf{0}$.
    - ▷ Since $G(\mathbf{x}_1) < \mathbf{0}$ and $\mathbf{z}_0^* \geq \mathbf{0}$, it must be $\langle G(\mathbf{x}_1), \mathbf{z}_0^* \rangle < 0$.
    - ▷ May take $r_0 = 1$.

- By definition of μ_0,

$$\mu_0 = \inf_{(r,\mathbf{z})\in A} (r + \langle \mathbf{z}, \mathbf{z}_0^* \rangle) \le \inf_{\mathbf{x}\in\Omega} (f(\mathbf{x}) + \langle G(\mathbf{x}, \mathbf{z}_0^* \rangle) \le \inf_{\mathbf{x}\in\Omega, G(\mathbf{x})\le \mathbf{0}} f(\mathbf{x}) \le \mu_0.$$

- If there exist an $\mathbf{x}_0 \in \Omega$ such that $G(\mathbf{x}_0) \le \mathbf{0}$ and $f(\mathbf{x}_0) = \mu_0$, then

$$\mu_0 \le f(\mathbf{x}_0) + \langle G(\mathbf{x}_0), \mathbf{z}_0^* \rangle \le f(\mathbf{x}_0) \le \mu_0.$$

  So it must be that $\langle G(\mathbf{x}_0), \mathbf{z}_0^* \rangle = 0$.

- Three important assumptions:
  - ◇ Both f and G are convex. This assumption guarantees that A and ω are convex.
  - ◇ The positive cone P contains an interior point. This assumption guarantees the existence of a separating hyperplane.
  - ◇ There exists $\mathbf{x}_1$ such that $G(\mathbf{x}_1) < \mathbf{0}$. This assumption guarantees the the separating hyperplane is non-vertical. (Where is this used?)

Saddle Point Theory

- Algebraic formulation of the Lagrange multiplier theorem.
- Sufficient condition for non-convex problems.
- Sensitivity

Algebraic Formulation for Convex Programming

- Suppose that
  - ◇ The same assumptions as in the basic *convex* programming problem hold, i.e.,
    - ▷ The positive cone P contains an interior point.
    - ▷ $G : \Omega \to Z$ is convex with respect to P.
    - ▷ There exists a point $\mathbf{x}_1 \in \Omega$ such that $G(\mathbf{x}_1) < \mathbf{0}$.
    - ▷ $\mu_0 := \inf\{f(\mathbf{x}) | \mathbf{x} \in \Omega, G(\mathbf{x}) \leq \mathbf{0}\}$ is finite.
  - ◇ At the point $\mathbf{x}_0$ the constrained minimum is achieved.
- Then
  - ◇ There exists a Lagrangian multiplier $\mathbf{z}_0^* \geq \mathbf{0}$ such that the Lagrangian

$$L(\mathbf{x}_0, \mathbf{z}^*) := f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}^* \rangle$$

    has a saddle point at $(\mathbf{x}_0, \mathbf{z}_0^*)$.
  - ◇ That is,

$$L(\mathbf{x}_0, \mathbf{z}^*) \leq L(\mathbf{x}_0, \mathbf{z}_0^*) \leq L(\mathbf{x}, \mathbf{z}_0^*)$$

    for all $\mathbf{x} \in \Omega$, $\mathbf{z}^* \geq \mathbf{0}$.
    - ▷ The right inequality is by the definition of μ_0.
    - ▷ The left inequality is due to

$$L(\mathbf{x}_0, \mathbf{z}^*) - L(\mathbf{x}_0, \mathbf{z}_0^*) = \langle G(\mathbf{x}_0), \mathbf{z}^* \rangle - \langle G(\mathbf{x}_0), \mathbf{z}_0^* \rangle = \langle G(\mathbf{x}_0), \mathbf{z}^* \rangle \leq \mathbf{0}.$$

Sufficient Condition in General

- Suppose that
  - ◇ $f : \Omega \subset X \to \mathbb{R}$.
  - ◇ $G : \Omega \to Z$.
  - ◇ Z has an nonempty positive cone P which is closed.
  - ◇ There exist a $\mathbf{z}_0^* \in Z^*$, $\mathbf{z}_0^* \geq \mathbf{0}$ and $\mathbf{x}_0 \in \Omega$ such that the Lagrangian

    $$L(\mathbf{x}_0, \mathbf{z}^*) := f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}^* \rangle$$

    has a saddle point at $(\mathbf{x}_0, \mathbf{z}_0^*)$, i.e.,

    $$L(\mathbf{x}_0, \mathbf{z}^*) \leq L(\mathbf{x}_0, \mathbf{z}_0^*) \leq L(\mathbf{x}, \mathbf{z}_0^*)$$

    for all $\mathbf{x} \in \Omega$, $\mathbf{z}^* \geq \mathbf{0}$.
- Then, $\mathbf{x}_0$ solves the constrained problem:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & G(\mathbf{x}) \leq \mathbf{0}, \quad \mathbf{x} \in \Omega. \end{array}$$

Proof

- By the left inequality in the saddle-point condition

$$\langle G(\mathbf{x}_0), \mathbf{z}^* \rangle \leq \langle G(\mathbf{x}_0), \mathbf{z}_0^* \rangle$$

  for all $\mathbf{z}^* \geq \mathbf{0}$.

  - ◇ Given any $\mathbf{z}_1^* \geq \mathbf{0}$, take $\mathbf{z}^* = \mathbf{z}_1^* + \mathbf{z}_0^*$. It follows that

$$\langle G(\mathbf{x}_0), \mathbf{z}_1^* \rangle \leq \mathbf{0}.$$

  - ◇ Claim that $G(\mathbf{x}_0) \leq \mathbf{0}$.
    - ▷ Assume otherwise, then by the Separating Hyperplane Theorem there exists $\mathbf{x}^* \in X^*$ such that

$$\langle G(\mathbf{x}_0), \mathbf{x}^* \rangle > \langle \mathbf{n}, \mathbf{x}^* \rangle \quad \text{for all } \mathbf{n} \in -P.$$

    - ▷ Since $-P$ is a cone, $\langle \mathbf{n}, \mathbf{x}^* \rangle$ can never be positive because then $\langle G(\mathbf{x}_0), \mathbf{x}^* \rangle < \langle \alpha\mathbf{n}, \mathbf{x}^* \rangle$ for some $\alpha > 0$.
    - ▷ It follows that $\mathbf{x}^* \geq \mathbf{0}$.
    - ▷ But $\sup_{\mathbf{n} \in -P} \langle \mathbf{n}, \mathbf{x}^* \rangle = 0$. We have $\langle G(\mathbf{x}_0), \mathbf{x}^* \rangle > 0$. This is a contradiction.

Sensitivity

- Assume that
  - f and G are convex on a convex set Ω.
  - $\mathbf{x}_1$ and $\mathbf{x}_2$ solve the problems

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}), \\ \text{subject to} & \mathbf{x} \in \Omega, \quad \text{and} \quad G(\mathbf{x}) \leq \mathbf{z}_1, \quad G(\mathbf{x}) \leq \mathbf{z}_2, \end{array}$$

    respectively, with Lagrange multipliers $\mathbf{z}_1^*$ and $\mathbf{z}_2^*$.
- Then

$$\langle \mathbf{z}_2 - \mathbf{z}_1, \mathbf{z}_2^* \rangle \leq f(\mathbf{x}_1) - f(\mathbf{x}_2) \leq \langle \mathbf{z}_2 - \mathbf{z}_1, \mathbf{z}_1^* \rangle.$$

  - Using the Lagrange multiplier $\mathbf{z}_1^*$,

$$f(\mathbf{x}_1) + \langle G(\mathbf{x}_1) - \mathbf{z}_1, \mathbf{z}_1^* \rangle \leq f(\mathbf{x}) + \langle G(\mathbf{x}) - \mathbf{z}_1, \mathbf{z}_1^* \rangle$$

    for all $\mathbf{x} \in \Omega$.
  - Taking $\mathbf{x} = \mathbf{x}_2$,

$$f(\mathbf{x}_1) - f(\mathbf{x}_2) \leq \langle G(\mathbf{x}_2) - \mathbf{z}_1, \mathbf{z}_1^* \rangle - \underbrace{\langle G(\mathbf{x}_1) - \mathbf{z}_1, \mathbf{z}_1^* \rangle}_{0} \leq \langle \mathbf{z}_2 - \mathbf{z}_1, \mathbf{z}_1^* \rangle$$

  - The other inequality can be done similarly using the Lagrange multiplier $\mathbf{z}_2^*$.

Lagrange Multiplier versus Derivative

- Rewrite the above proof (for the right inequality) as

$$\langle \mathbf{z}_1 - \mathbf{z}, \mathbf{z}_1^* \rangle \leq \inf_{\mathbf{x}\in\Omega, G(\mathbf{x})\leq \mathbf{z}} (\langle \mathbf{z}_1 - G(\mathbf{x}), \mathbf{z}_1^* \rangle) \leq \inf_{\mathbf{x}\in\Omega, G(\mathbf{x})\leq \mathbf{z}} (f(\mathbf{x}) - f(\mathbf{x}_1)) = \omega(\mathbf{z}) - \omega(\mathbf{z}_1).$$

- Similar bound for the left inequality.
- If the functional ω is Fréchet differentiable at $\mathbf{z}_1$, then

$$\omega'(\mathbf{z}_1) = -\mathbf{z}_1^*.$$

- The Lagrange multiplier $\mathbf{z}_1^*$ is the negative of the first-order derivative of the optimal objective with respect to the constraint $\mathbf{z}_1$.

Duality

- The duality principle
- Significance of the dual problem

Duality Principle

- Consider
  - ◇ The basic convex programming problem:

$$\begin{array}{ll}\text{Minimize} & f(\mathbf{x}),\\ \text{Subject to} & \mathbf{x} \in \Omega, \quad G(\mathbf{x}) \leq \mathbf{0},\end{array}$$

    where f, G, and Ω are convex s
  - ◇ The associate problem:

$$\begin{array}{ll}\text{Minimize} & f(\mathbf{x}),\\ \text{Subject to} & \mathbf{x} \in \Omega, \quad G(\mathbf{x}) \leq \mathbf{z},\end{array}$$

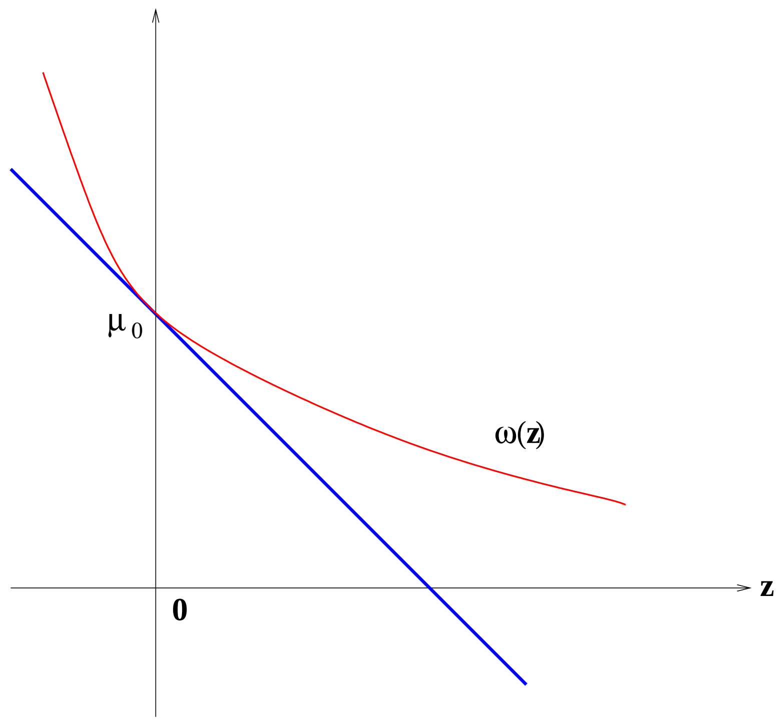
  - ◇ The subset Γ in Z:

$$\Gamma := \{\mathbf{z} \in Z | \text{There exists at least one } \mathbf{x} \in \Omega \text{ such that } G(\mathbf{x}) \leq \mathbf{z}\}.$$

  - ◇ The primal functional $\omega : \Gamma \to \mathbb{R}$ by

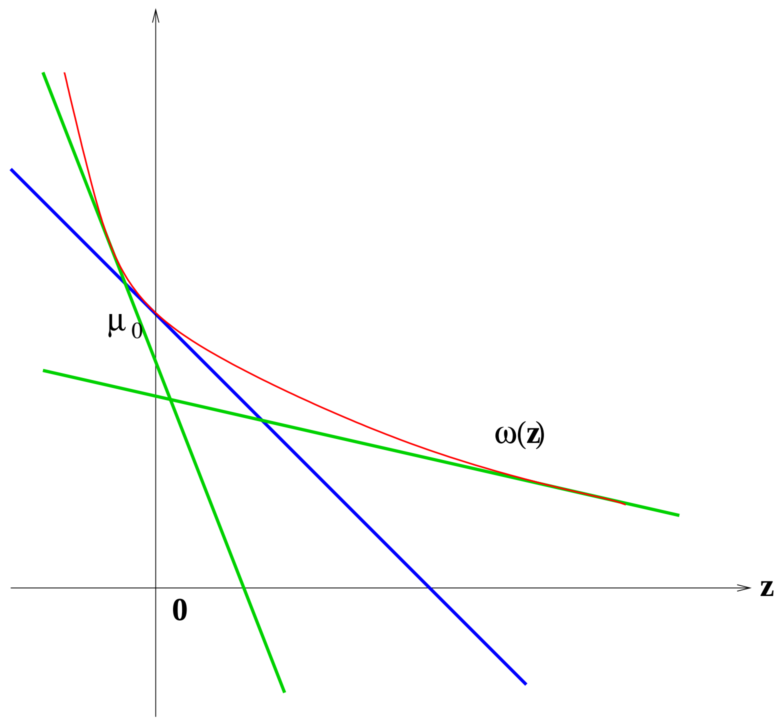
$$\omega(\mathbf{z}) := \inf\{f(\mathbf{x}) | \mathbf{x} \in \Omega, G(\mathbf{x}) \leq \mathbf{z}\}.$$

Geometric Meaning of Lagrange Multiplier



- There is a hyperplane tangent to ω at $\mathbf{z} = \mathbf{0}$ and lying below ω throughout Γ.
- Change the coordinate system so that the tangent line becomes horizontal.
    - ◇ ω attains its minimum at $\mathbf{0}$.
- The change of coordinate system can be achieved by adding an appropriate linear function $\langle \mathbf{z}, \mathbf{z}_0^* \rangle$ to $\omega(\mathbf{z})$, i.e., $\omega(\mathbf{z}) + \langle \mathbf{z}, \mathbf{z}_0^* \rangle$ is minimized at $\mathbf{z} = \mathbf{0}$.
    - ◇ The functional $\mathbf{z}_0^*$ is the Lagrange multiplier.

Geometric Meaning of Duality



- The duality principle is based on the observation that μ_0 is equal to the *maximum intercept* with the vertical axis of all closed hyperplanes that lie below $\omega(\mathbf{z})$.
- The maximum intercept is attained by the hyperplane corresponding to the Lagrange multiplier.

Dual Functional

- Corresponding to $\omega(\mathbf{z})$, define for every $\mathbf{z}^* \geq \mathbf{0}$ (i.e., over the positive cone in Z^*),

$$\phi(\mathbf{z}^*) := \inf_{\mathbf{x}\in\Omega}[f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}^*\rangle].$$

  - ◇ The dual functional $\phi(\mathbf{z}^*)$ is concave.
  - ◇ Can rewrite
  $$\phi(\mathbf{z}^*) = \inf_{\mathbf{z}\in\Gamma}[\omega(\mathbf{z}) + \langle \mathbf{z}, \mathbf{z}^*\rangle].$$
  - ◇ Given any $\mathbf{z}^* \geq \mathbf{0}$ and $\mathbf{z} \in \Gamma$,
  $$\begin{aligned} \phi(\mathbf{z}^*) &= \inf_{\mathbf{x}\in\Omega}[f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}^*\rangle] \leq \inf\{f(\mathbf{x}) + \langle \mathbf{z}, \mathbf{z}^*\rangle | G(\mathbf{x}) \leq \mathbf{z}, \mathbf{z} \in \Omega\} \\ &= \omega(\mathbf{z}) + \langle \mathbf{z}, \mathbf{z}^*\rangle. \end{aligned}$$
  - ◇ Given any $\mathbf{x}_1 \in \Omega$, let $\mathbf{z}_1 = G(\mathbf{x}_1)$. Then
  $$\begin{aligned} f(\mathbf{x}_1) + \langle G(\mathbf{x}_1), \mathbf{z}^*\rangle &\geq \inf\{f(\mathbf{x}) + \langle \mathbf{z}_1, \mathbf{z}^*\rangle | G(\mathbf{x}) \leq \mathbf{z}_1, \mathbf{z} \in \Omega\} \\ &= \omega(\mathbf{z}_1) + \langle \mathbf{z}_1, \mathbf{z}^*\rangle. \end{aligned}$$

Lagrange Duality

- Consider the basic convex programming problem where
  - f is a real-valued convex functional on a convex subset Ω in X.
  - G is a convex mapping from Ω into Z with a positive cone P.
  - There exists a point $\mathbf{x}_1 \in \Omega$ such that $G(\mathbf{x}_1) < \mathbf{0}$.
  - $\mu_0 := \inf\{f(\mathbf{x})|\mathbf{x} \in \Omega, G(\mathbf{x}) \leq \mathbf{0}\}$ is finite.
- Claim that
  - Equivalence:
    $$\inf_{G(\mathbf{x})\leq\mathbf{0},\mathbf{x}\in\Omega} = \max_{\mathbf{z}^*\geq\mathbf{0}} \phi(\mathbf{z}^*).$$
  - The maximum on the right is achieved by some $\mathbf{z}_0^* \geq \mathbf{0}$.
  - If the infimum on the left is achieved by some $\mathbf{x}_0 \in \Omega$, then
    $$\langle G(\mathbf{x}_0), \mathbf{z}_0^* \rangle = 0$$
    and $\mathbf{x}_0$ minimized $f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}_0^* \rangle$ for all $\mathbf{x} \in \Omega$.
    - Given any $\mathbf{z}^* \geq \mathbf{0}$, observe
      $$\inf_{\mathbf{x}\in\Omega}[f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}^* \rangle] \leq \inf_{\mathbf{x}\in\Omega, G(\mathbf{x})\leq\mathbf{0}}[f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}^* \rangle] \leq \inf_{\mathbf{x}\in\Omega, G(\mathbf{x})\leq\mathbf{0}} f(\mathbf{x}) = \mu_0.$$
    - But, it is already established earlier that there exists a $\mathbf{z}_0^* \geq \mathbf{0}$ such that
      $$\mu_0 = \inf_{\mathbf{x}\in\Omega}\{f(\mathbf{x}) + \langle G(\mathbf{x}), \mathbf{z}_0^* \rangle\}.$$

An Example

- Consider the quadratic programming problem:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} - \mathbf{b}^\top \mathbf{x} \\ \text{subject to} & A\mathbf{x} \leq \mathbf{c} \end{array}$$

  where $Q \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, $A \in \mathbb{R}^{m\times n}$, and $\mathbf{x} \in \mathbb{R}^n$.

- Assume that there is an $\mathbf{x}$ satisfying $A\mathbf{x} < \mathbf{c}$, the duality implies the equivalent problem:

$$\max_{\boldsymbol{\lambda}\geq \mathbf{0}} \underbrace{\min_{\mathbf{x}}\{\frac{1}{2}\mathbf{x}^\top Q\mathbf{x} - \mathbf{b}^\top\mathbf{x} + \boldsymbol{\lambda}^\top(A\mathbf{x} - \mathbf{c})\}}_{unconstrained}.$$

- The unconstrained problem has a closed form solution

$$\mathbf{x} = Q^{-1}(\mathbf{b} - A^\top\boldsymbol{\lambda}).$$

- The dual problem is also a quadratic programming problem in $\mathbb{R}^m$:

$$\max_{\boldsymbol{\lambda}\geq\mathbf{0}}\{-\frac{1}{2}\boldsymbol{\lambda}^\top P\boldsymbol{\lambda} - \boldsymbol{\lambda}^\top\mathbf{d} - \frac{1}{2}\mathbf{b}^\top Q^{-1}\mathbf{b}\}$$

  with $P := AQ^{-1}A^\top \in \mathbb{R}^{m\times m}$ and $\mathbf{d} := \mathbf{c} - AQ^{-1}\mathbf{b} \in \mathbb{R}^m$.

  - ⋄ The dual problem is easier to solve because $m < n$.

Applications

- Optimal Control
- Oil Drilling
- Farmer's Allocation

Optimal Control

- A system governed by

$$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t) + \mathbf{b}(t)u(t)$$

  - ◇ $A(t) \in \mathbb{R}^{n\times n}$, $\mathbf{x}(t), \mathbf{b}(t) \in \mathbb{R}^{n\times 1}$.
  - ◇ $u(t)$ is a scalar control.

- Given initial state $\mathbf{x}(t_0)$, seek control u_0 so as to

$$\begin{array}{rl} \text{minimize} & J = \dfrac{1}{2}\displaystyle\int_{t_0}^{t_1} u^2(t)dt \\ \text{subject to} & \mathbf{x}(t_1) \geq \mathbf{c}. \end{array}$$

- Physical interpretation:
  - ◇ Select a thrust program for a rocket which must exceed certain altitude and velocity limits in a given time.
  - ◇ Select a production program for a plant with constraints on the total amount produced in a given time.

Reformulation

- Let $\Phi(t;s)$ be the fundamental solution matrix of the homogeneous system,

$$\dot{\Phi}(t;s) = A(t)\phi(t;s), \quad \Phi(s;s) = I.$$

  - ◇ The solution to the inhomogeneous system can be written as

$$\mathbf{x}(t) = \Phi(t;t_0)\mathbf{x}(t_0) + \int_{t_0}^{t_1} \Phi(t;\tau)\mathbf{b}(\tau)u(\tau)d\tau.$$

- Rewrite the original control problem as

$$\begin{aligned} \text{minimize} \quad & J(u) = \frac{1}{2}\int_{t_0}^{t_1} u^2(t)dt \\ \text{subject to} \quad & Ku \geq \mathbf{d}, \end{aligned}$$

with

$$\begin{aligned} \mathbf{d} \quad &:= \quad \mathbf{c} - \Phi(t_1;t_0)\mathbf{x}(t_0), \\ Ku \quad &:= \quad \int_{t_0}^{t_1} \Phi(t_1;\tau)\mathbf{b}(\tau)u(\tau)d\tau \end{aligned}$$

  - ◇ This is a convex programming problem defined on $\mathcal{L}_2[t_0, t_1]$ with finite dimensional constraint.

- By using the duality theorem, the minimum objective value μ_0 is given by

$$\begin{aligned}\mu_0 &= \max_{\boldsymbol{\lambda}\geq\mathbf{0}} \min_u \{J(u) + \boldsymbol{\lambda}^\top(\mathbf{d} - Ku)\},\\ &= \max_{\boldsymbol{\lambda}\geq\mathbf{0}} \left\{\min_u \int_{t_0}^{t_1} \frac{1}{2}u^2(t) - \boldsymbol{\lambda}^\top \Phi(t_1;\tau)\mathbf{b}(\tau)u(\tau)d\tau\right\} + \boldsymbol{\lambda}^\top\mathbf{d}.\end{aligned}$$

  with $\boldsymbol{\lambda} \in \mathbb{R}^n$.

  - ◇ The minimization problem can easily be solved for $u(t)$.
  - ◇ The maximization problem is reduced to

$$\mu_0 = \max_{\boldsymbol{\lambda}\geq\mathbf{0}} \boldsymbol{\lambda}^\top Q\boldsymbol{\lambda} + \boldsymbol{\lambda}^\top\mathbf{d},$$

    with

$$Q := -\frac{1}{2}\int_{t_0}^{t_1} \Phi(t_1;\tau)\mathbf{b}(\tau)\mathbf{b}^\top(\tau)\Phi^\top(t_1;\tau)d\tau.$$

- The optimal control problem is reduced to a finite-dimensional maximization problem.

  - ◇ Once the optimal $\boldsymbol{\lambda}_0$ is determined, the optimal u_0 is given by

$$u_0(t) = \boldsymbol{\lambda}_0^\top\Phi(t_1;t)\mathbf{b}(t).$$

Oil Drilling

- Physical problem: An underground oil deposit of known quantity α has been located. Determine a long-term oil extraction program so as to maximize the total profit.
  - ◇ Let $x(t)$ denote the extraction rate.
    - ▷ $x(t)$ should be an integrable function over $[0, \infty)$.
  - ◇ Let $F(x)$ denote the profit rate associated with the extraction rate.
    - ▷ It is reasonable to assueme that F is increasing on $[0, \infty)$ with $F(0) = 0$.
    - ▷ It is common that the marginal return will diminish as more oil is on the market. Assume that F is strictly concave.
  - ◇ Let $v(t)$ represent the discount factor.
    - ▷ $v(t)$ is positive, strictly decreasing toward zero, and integrable over $[0, \infty)$.
- Formulation:

$$\begin{array}{ll} \text{maximize} & \displaystyle\int_0^\infty F(x(t))v(t)dt, \\ \text{subject to} & \displaystyle\int_0^\infty x(t)dt \leq \alpha, \quad x(t) \geq 0. \end{array}$$

Truncated Problem

- Assume temporarily that $x(t)$ is continuous and $x(t) = 0$ for $t \geq T$.
- The 1-dimensional Lagrangian is

$$L(x, \lambda) = \int_0^T F(x(t))v(t)dt - \lambda \left(\int_0^T x(t)dt - \alpha \right).$$

  - ◇ The minus sign is to reflect the fact that a concave functional is being maximized.
- The goal is to find $x_0(t) \geq 0$ and λ_0 such that

$$L(x_0, \lambda_0) = \min_{\lambda \geq 0} \max_{x \geq 0} L(x, \lambda).$$

- Stationary point of the above Lagrangian can be found by using differentiation.

Differentiation

- Suppose
  - $f : X \to \mathbb{R}$ is Frèchet differentiable and convex.
  - P is a convex cone in X.
- $\mathbf{x}_0 \in P$ minimizes f over P if and only if

$$\begin{aligned} \delta f(\mathbf{x}_0; \mathbf{x}) &\geq 0, \quad \text{for all } \mathbf{x} \in P \\ \delta f(\mathbf{x}_0; \mathbf{x}_0) &= 0. \end{aligned}$$

- Necessity:
  - If $\mathbf{x}_0$ minimized f over P, then

$$\begin{aligned} \frac{d}{ds} f(\mathbf{x}_0 + s(\mathbf{x} - \mathbf{x}_0))|_{s=0} &\geq 0, \quad \text{for all } \mathbf{x} \in P \\ \delta f(\mathbf{x}_0; \mathbf{x} - \mathbf{x}_0) &\geq \mathbf{0}. \end{aligned}$$

    - ▷ Taking $\mathbf{x} = \mathbf{x}_0/2$ yields $\delta f(\mathbf{x}_0; \mathbf{x}_0) \leq \mathbf{0}$.
    - ▷ Taking $\mathbf{x} = 2\mathbf{x}_0$ yields $\delta f(\mathbf{x}_0; \mathbf{x}_0) \geq \mathbf{0}$

- Sufficiency:
  - ◇ Given $\mathbf{x}_0$, any $\mathbf{x} \in P$ and $0 < \alpha < 1$, convexity implies that

$$\begin{aligned} f(\mathbf{x}_0 + \alpha(\mathbf{x} - \mathbf{x}_0)) &\leq f(\mathbf{x}_0) + \alpha(f(\mathbf{x} - \mathbf{x}_0), \\ f(\mathbf{x}) - f(\mathbf{x}_0) &\geq \frac{1}{\alpha}(f(\mathbf{x}_0 + \alpha(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0)). \end{aligned}$$

    - ▷ Taking $\alpha \to 0$ yields $f(\mathbf{x}) - f(\mathbf{x}_0) \geq \delta f(\mathbf{x}_0; \mathbf{x} - \mathbf{x}_0)$.
  - ◇ The conditions yield $\delta f(\mathbf{x}_0; \mathbf{x} - \mathbf{x}_0) \geq \mathbf{0}$.

Partial Solution

- Differentiating $L(x, \lambda_0)$ over $x \geq 0$, maximization is equivalent to

$$\begin{aligned}
\int_0^T \{F_x(x_0(t))v(t) - \lambda_0\} \, x(t)dt &\leq 0, \quad \text{for all } x(t) \geq 0, \\
\int_0^T \{F_x(x_0(t))v(t) - \lambda_0\} \, x_0(t)dt &= 0.
\end{aligned}$$

- Rewrite the conditions as

$$\begin{aligned}
x_0(t) &\geq 0, \\
\lambda_0 &> 0, \quad \text{(why?)} \\
\mu(t) &\geq 0, \quad \text{(Slack variable!)} \\
\mu(t)x_0(t) &\equiv 0, \\
\int_0^T x_0(t)dt &= \alpha, \\
F_x(x_0(t))v(t) - \lambda_0 + \mu(t) &= 0.
\end{aligned}$$

- Strict concavity of F implies that F_x is strictly decreasing. So F_x^{-1} exists and is strictly decreasing.
  - ◇ On any interval where $x_0(t) > 0$, it must be $\mu(t) = 0$ and hence $x_0(t) = F_x^{-1}(\lambda_0/v(t))$.
    - ▷ It follows that if $x_0(t) > 0$, then $x_0(t)$ is decreasing.

- A continuous solution $x(t)$, if exists, must be of the form

$$x_0(t) = \begin{cases} F_x^{-1}\left(\frac{\lambda_0}{v(t)}\right) & 0 \le t \le t_0, \\ 0 & t_0 \le t \le T. \end{cases}$$

  - ◇ At t_0, continuity implies $F_x^{-1}(\lambda_0/v(t_0)) = 0$. Thus,

$$\lambda_0 = F_x(0)v(t_0).$$

- Define

$$J(t_0) := \int_0^{t_0} F_x^{-1}\left(\frac{F_x(0)v(t_0)}{v(t)}\right) dt.$$

  - ◇ J is continuous.
  - ◇ It can be argued that $\lim_{t_0 \to 0} J(t_0) = 0$ and $\lim_{t_0 \to \infty} J(t_0) = \infty$.
  - ◇ It is possible to choose t_0 such that $J(t_0) = \alpha$.

- Can now choose $\mu(t)$ as

$$\mu(t) = \begin{cases} 0 & 0 \le t \le t_0, \\ \lambda_0 - F_x(0)v(t) & t_0 \le t \le T. \end{cases}$$

- The truncated problem is now completely solved.
- Any integrable function $y(t)$ for which $\int_0^\infty y(t)dt = \alpha$ can be approximated by a continuous function $x(t)$ vanishing outside a finite interval such that

$$\int_0^\infty \left(F(y(t)) - F(x(t))\right) v(t)dt < \epsilon.$$

  - ◇ The solution $x_0(t)$ for the truncated problem is optimal for the original problem.

Farmer's Allocation

- Determine a strategy on how much to store or to sell a certain crop being harvested for reinvestment in additional land or equipment to increase the production so as to maximize the total amount stored over a fixed period of time.
- Setup:
  - ◇ Denote

$$\begin{aligned} x_1(t) &= \text{rate of production}, \\ x_2(t) &= \text{rate of reinvestment}, \\ x_3(t) &= \text{rate of storage}. \end{aligned}$$

  - ◇ Want to

$$\begin{array}{rl} \text{maximize} & \displaystyle\int_0^T x_3(t)dt, \\ \text{subject to} & \dot{x}_1(t) = x_2(t), \quad x_1(0) = x_0 > 0, \\ & x_2(t) + x_3(t) = x_1(t), \\ & x_1(t), x_2(t), x_3(t) \geq 0. \end{array}$$

    - ▷ This is an infinite-dimensional linear programming problem.

Reformulation

- Rewrite the problem in terms of $x_2(t) \in C[0,T]$,

$$\begin{array}{rc} \text{maximize} & \int_0^T \left\{ x_0 + \int_0^t x_2(\tau)d\tau - x_2(t) \right\} dt \\ \text{subject to} & x_0 + \int_0^t x_2(\tau)d\tau \geq x_2(t) \\ & x_2(t) \geq 0 \end{array}$$

- The Lagrangian multiplier resides in the positive cone of the dual space $C^*[0,T] = NBV[0,T]$.
  - ◇ The multiplier v is a nondecreasing function of bounded variation.
  - ◇ The Lagrangian is given by

$$L(x_2, v) := \int_0^T \left\{ x_0 + \int_0^t x_2(\tau)d\tau - x_2(t) \right\} dt + \int_0^T \left\{ x_0 + \int_0^t x_2(\tau)d\tau - x_2(t) \right\} dv(t).$$

  - ◇ The duality principle implies that the multiplier v solves the problem

$$\min_{v \geq 0} \sup_{x_2 \geq 0} L(x_2, v).$$

Solution

- Rewrite the Lagrangian after integration by parts as

$$L(x_2, v) = x_0 T + \int_0^T \{(T - t - 1) + [v(T) - v(t)]\} x_2(t) dt + \int_0^T [x_0 - x_2(t)] dv(t).$$

- The supremum of $L(x_2, v)$ is finite over $x_2 \geq 0$ if and only if

$$\{T - t - 1 + v(T) - v(t)\} - dv(t) \leq 0, \quad \text{for all } t.(Why?)$$

  ◇ In this case, for each given $v(t)$,

$$\sup_{x_2 \geq 0} L(x_2, v) = x_0 T + \int_0^T x_0 dv(t).$$

- By inspection, the nondecreasing bounded variation $v(t)$ that minimizes the dual problem must satisfy the system

$$\begin{cases} v(t) = 0, & & t > T - 1, \\ \frac{dv}{dt} = -(v(t) + t + 1 - T), & v(T-1) = 0, & 0 < t < T - 1. \end{cases}$$

  ◇ With the optimal $v(t)$,

$$\sup_{x_2 \geq 0} L(x_2, v) = x_0 e^{T-1}.$$

  ◇ Still need to know the strategy $x_2(t)$.

- By checking, it turns out the (discontinuous) function

$$x_2(t) = \begin{cases} x_1(t), & 0 \le t \le T-1, \\ 0, & T-1 \le t \le T \end{cases}$$

happens to produce the same optimal objective value $x_0 e^{T-1}$.

  - ◇ The strategy represents a total reinvestment until $t = T-1$ and then total storage until $t = T$.
  - ◇ The above x_2 can be modified into continuous functions producing values arbitrarily close to the the optimal objective value.