MATH529 – Fundamentals of Optimization
Duality

Marco A. Montes de Oca

Mathematical Sciences, University of Delaware, USA

Example:

Maximize $3x + 4y$

subject to:

$x + y \leq 12$

$x + 4y \leq 42$

$x, y \geq 0$

Definitions

Definition (Supremum)

Let $f(\mathbf{x})$ be a real valued function on $C \subset \mathbb{R}^n$. If there is a smallest number $\beta \in \mathbb{R}$ such that $f(\mathbf{x}) \leq \beta$ for all $\mathbf{x} \in C$, then β is the *supremum* of $f(\mathbf{x})$ on C and write

$$\beta = \sup_{x \in C} f(\mathbf{x})$$

Example (1)

If $\mathbf{x}^\star$ is the global maximizer of $f(\mathbf{x})$ on C, then $\sup\limits_{x \in C} f(\mathbf{x}) = f(\mathbf{x}^\star)$.

Definitions

Example (2)

Let $f(\mathbf{x}) = \frac{1}{x_1^2 + x_2^2}$, where $C = \{(x_1, x_2) \,|\, x_1, x_2 \in \mathbb{R}^2 \setminus \{(0,0)\}\}$. Since $f(\mathbf{x})$ can be made as large as desired by letting $x_1 \to 0$ and $x_2 \to 0$ simultaneously, then there is no upper bound for $f(\mathbf{x})$ on C. Thus, strictly speaking, $\sup_{x \in C} f(\mathbf{x})$ does not exist. However, we will write $\sup_{x \in C} f(\mathbf{x}) = \infty$.

Example (3)

Let $f(x) = \dfrac{1}{1 + e^{-x}}$, where $C = \mathbb{R}$. In this case, $\sup_{x \in C} f(x) = 1$, even though here is no global maximizer on C.

Definitions

Definition (Infimum)

Let $f(\mathbf{x})$ be a real valued function on $C \subset \mathbb{R}^n$. If there is a largest number $\alpha \in \mathbb{R}$ such that $f(\mathbf{x}) \geq \beta$ for all $\mathbf{x} \in C$, then α is the *infimum* of $f(\mathbf{x})$ on C and write

$$\beta = \inf_{x \in C} f(\mathbf{x})$$

Duality

Let us formulate a general minimization problem as:

$$\text{Minimize } f(\mathbf{x})$$

$$\text{subject to:}$$

$$\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

The Lagrangian for this problem is therefore:

$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) - \boldsymbol{\mu}^T \mathbf{h}(\mathbf{x})$, with $\boldsymbol{\lambda} \in \mathbb{R}^m_+$ and $\boldsymbol{\mu} \in \mathbb{R}^p$, where m is the number of inequality constraints and p is the number of equality constraints.

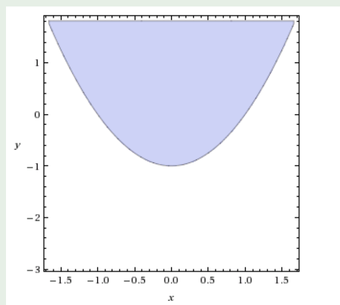
Duality

Definition (Primal function)

The primal function associated with the optimization problem above is:

$$L_p(\mathbf{x}) = \sup_{(\boldsymbol{\lambda},\boldsymbol{\mu})} L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu})$$

Example (Min $x^2 + y^2$, s.t. $1 - x^2 + y \geq 0$)



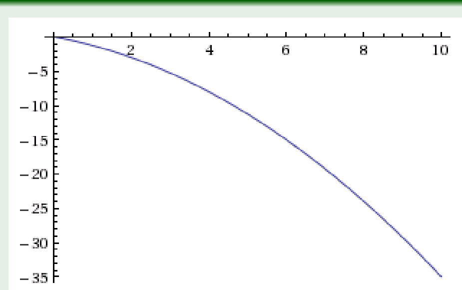
Duality

Definition (Dual function)

The dual function associated with the optimization problem above is:

$$L_d(\boldsymbol{\lambda},\boldsymbol{\mu}) = \inf_{\mathbf{x}} L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu})$$

Example (Min $x^2 + y^2$, s.t. $1 - x^2 + y \geq 0$)



Duality

The *primal problem* is to find

$$\min_{\mathbf{x}} L_p(\mathbf{x})$$

and the *dual problem* is to find

$$\max_{(\boldsymbol{\lambda},\boldsymbol{\mu})} L_d(\boldsymbol{\lambda},\boldsymbol{\mu})$$

Duality

Example 1: Find the dual function associated with:

Minimize $x + y$

subject to:

$x^2 + y^2 \leq 1$

Duality

Example 2:

$$\text{Minimize } \mathbf{c}^T\mathbf{x}$$

subject to:

$$A\mathbf{x} \geq \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

where A is an $m \times n$ matrix, $\mathbf{c} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$.

Duality

The Lagrangian is:

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}) &= \mathbf{c}^T\mathbf{x} + \boldsymbol{\lambda}^T(\mathbf{b} - A\mathbf{x}) = \\ & \mathbf{c}^T\mathbf{x} + \boldsymbol{\lambda}^T\mathbf{b} - \boldsymbol{\lambda}^T A\mathbf{x} = \\ & (\mathbf{c} - A^T\boldsymbol{\lambda})^T\mathbf{x} + \boldsymbol{\lambda}^T\mathbf{b} \end{aligned}$$

Thus, the dual function is:

$$L_d(\boldsymbol{\lambda}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) = \boldsymbol{\lambda}^T\mathbf{b} + \inf_{\mathbf{x}} \left[(\mathbf{c} - A^T\boldsymbol{\lambda})^T\mathbf{x}\right]$$

Now,

$$\inf_{\mathbf{x}} \left[(\mathbf{c} - A^T\boldsymbol{\lambda})^T\mathbf{x}\right]$$

is bounded only when $\mathbf{c} - A^T\boldsymbol{\lambda} \geq 0$.

Duality

Therefore, the dual problem can be formulated as follows:

Maximize $\mathbf{b}^T \boldsymbol{\lambda}$

subject to:

$$A^T \boldsymbol{\lambda} \leq \mathbf{c}$$

$$\boldsymbol{\lambda} \geq \mathbf{0}$$

Duality

Why do we care?

Duality

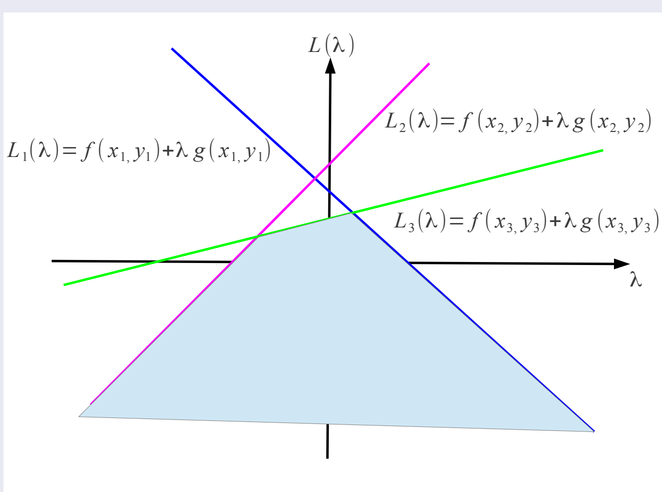
- In some cases, the dual problem is easier to solve than the original problem.
- In other cases, the solution of the dual problem provides a lower bound on the optimal value for the primal problem.
- Duality theory is used to motivate and develop optimization algorithms
- In economics, the dual may have an important economic meaning of its own.

Duality

First main property of dual functions:

Theorem (Concavity of $L_d(\boldsymbol{\lambda}, \boldsymbol{\mu})$)

The function $L_d(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ *is concave.*



Duality

Second main property of dual functions:

Theorem (Weak duality (Lower bounds for objective function))

For any feasible solution $\bar{\mathbf{x}}$ *and any* $\bar{\boldsymbol{\lambda}} \geq \mathbf{0}$ *and* $\bar{\boldsymbol{\mu}} \in \mathbb{R}^p$,
$L_d(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) \leq f(\bar{\mathbf{x}})$.

Proof:

By definition:

$$L_d(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) = \inf_{\mathbf{x}} L(\mathbf{x}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) = \inf_{\mathbf{x}} (f(\mathbf{x}) - \bar{\boldsymbol{\lambda}}^T \mathbf{g}(\mathbf{x}) - \bar{\boldsymbol{\mu}}^T \mathbf{h}(\mathbf{x})) \leq$$

$$f(\bar{\mathbf{x}}) - \bar{\boldsymbol{\lambda}}^T \mathbf{g}(\bar{\mathbf{x}}) - \cancelto{0}{\bar{\boldsymbol{\mu}}^T \mathbf{h}(\bar{\mathbf{x}})} \leq f(\bar{\mathbf{x}})$$

because at $\bar{\mathbf{x}}$, $\mathbf{g}(\bar{\mathbf{x}}) \geq \mathbf{0}$.

Duality

It follows that for the optimal $(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ and the optimal $\mathbf{x}^\star$, the difference $L_p(\mathbf{x}^\star) - L_d(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$, called *duality gap*, is minimal. (If the duality gap is 0, we talk about *strong duality*.)