Introduction to duality

Another way to look at optimization

Michel Bierlaire

Introduction to optimization and operations research

Duality

Motivation

- We take a different point of view about the optimization problem: the point of view of the person who defines the constraint.
- We introduce an important concept in optimization: duality.
- It will help to solve complex problems (e.g. shortest path problem).
- It will allow to calculate useful bounds (e.g. in discrete optimization).

Simple example

Deal

- Win 1 € per meter of altitude.
- Must stay in the Alps.

Model

x: position, $f(x)$: altitude at x

$$\max_x f(x) \text{ s.t. } x \in \text{ Alps}$$

Solution for the alpinist

Mont Blanc: 4807 €

Simple example

4807m

8848m

New deal

- Pay a fine if out of the Alps.
- What fine?

Model

x: position, $f(x)$: altitude at x,
$a(x)$=fine at x

$$\max_x f(x) - a(x)$$

Solution

$$a(x) = 4041\,€ \text{ if } x \notin \text{ Alps}$$

Duality

Primal problem

- Point of view of the alpinist
- Optimization under strict constraint

Dual problem

- Point of view of the billionaire
- Penalization of the constraints

Constraint relaxation

Primal problem

$$\min_{x \in \mathbb{R}^2} 2x_1 + x_2$$

subject to

$$\begin{aligned} 1 - x_1 - x_2 &= 0 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Solution: $(0, 1)$
Optimal value: 1

Constraint relaxation

$$\min_{x \in \mathbb{R}^2} 2x_1 + x_2 + \lambda(1 - x_1 - x_2)$$

subject to

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Dual problem

What is the value of λ?

$\lambda = 0$

$$\min_{x \in \mathbb{R}^2} 2x_1 + x_2 + \lambda(1 - x_1 - x_2) = 2x_1 + x_2$$

subject to

$$x_1 \geq 0$$
$$x_2 \geq 0$$

Solution: $(0, 0)$
Optimal value: $0 < 1$.

Comments

- We obtain a lower bound,
- but a strict one.

$\lambda = 2$

$$\min_{x \in \mathbb{R}^2} 2x_1 + x_2 + \lambda(1 - x_1 - x_2) = 2 - x_2$$

subject to

$$x_1 \geq 0$$
$$x_2 \geq 0$$

Solution: none
Optimal value: $-\infty < 1$.

Comments

- This value of λ generates an unbounded problem.
- It must be avoided.

$\lambda = 1$

$$\min_{x \in \mathbb{R}^2} 2x_1 + x_2 + \lambda(1 - x_1 - x_2) = x_1 + 1$$

subject to

$$x_1 \geq 0$$
$$x_2 \geq 0$$

Solution: $(0, x_2)$, $\forall x_2$
Optimal value: $1 = 1$.

Comments

- We obtain the same optimal value as for the primal.
- The same solution is optimal as well.

Lagrangian and dual problem

Motivation

- Rigorous definition of the concepts.
- The objective function including the penalty term is called the Lagrangian.
- The problem to find the values of the penalty coefficients is the dual problem.

Lagrangian

Primal problem

$$\min f(x) \;\; [f : \mathbb{R}^n \to \mathbb{R}]$$

subject to

$$h(x) = 0 \quad [h : \mathbb{R}^n \to \mathbb{R}^m]$$
$$g(x) \leq 0 \quad [g : \mathbb{R}^n \to \mathbb{R}^p]$$

Lagrangian

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$$
$$\lambda \in \mathbb{R}^m \;\; \mu \in \mathbb{R}^p$$

Dual function

Lagrangian

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$$

Dual function

$$q : \mathbb{R}^{m+p} \to \mathbb{R} : q(\lambda, \mu) = \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$$

Note

We want

$$\mu^T g(x) \geq 0 \text{ when } g(x) > 0.$$

Therefore, we impose $\mu \geq 0$.

Dual bound

Theorem 4.5

x^* is an optimal solution of the primal.
If $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$, $\mu \geq 0$, then

$$q(\lambda, \mu) \leq f(x^*).$$

Proof

$$\begin{aligned} q(\lambda, \mu) = \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu) \\ \leq L(x^*, \lambda, \mu) \\ = f(x^*) + \lambda^T h(x^*) + \mu^T g(x^*) \\ = f(x^*) + \mu^T g(x^*) \leq f(x^*) \end{aligned}$$

Dual bound

Corollary 4.6

x^* is an optimal solution of the primal.
x is a feasible solution of the primal.
If $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$, $\mu \geq 0$, then

$$q(\lambda, \mu) \leq f(x^*) \leq f(x)$$

Dual problem

Motivation

Find the best lower bound

Dual problem

Motivation

Find the best lower bound

$$\max_{\lambda,\mu} q(\lambda, \mu)$$

Dual problem

Motivation

Find the best lower bound

$$\max_{\lambda,\mu} q(\lambda, \mu)$$

subject to

$$\mu \geq 0$$

Dual problem

Motivation

Find the best lower bound

$$\max_{\lambda,\mu} q(\lambda, \mu)$$

subject to

$$\mu \geq 0$$

and

$$(\lambda, \mu) \in \{\lambda, \mu \mid q(\lambda, \mu) > -\infty\}.$$

Weak duality theorem

If

- x^* is the optimal solution of the primal,
- (λ^*, μ^*) is the optimal solution of the dual, then

$$q(\lambda^*, \mu^*) \leq f(x^*)\,.$$

Duality and feasibility

$$q(\lambda^*, \mu^*) \leq f(x^*)\,.$$

Corollary 4.10

If one problem is unbounded, the other one is infeasible.

Duality and feasibility

		Dual problem		
		Optimal	Unbounded	Infeasible
Primal problem	Optimal		NO	
	Unbounded	NO	NO	YES
	Infeasible		YES	

Optimality of primal and dual

Corollary 4.11

If $\exists\ x^*$, λ^*, μ^* such that

$$q(\lambda^*, \mu^*) = f(x^*),$$

then they are optimal.

Proof

Consider x primal-feasible. By weak-duality:

$$f(x) \geq q(\lambda^*, \mu^*) = f(x^*)$$

Consider (λ, μ) dual-feasible. By weak-duality:

$$q(\lambda^*, \mu^*) = f(x^*) \geq q(\lambda, \mu)$$

Duality and feasibility

		Dual problem		
		Optimal	Unbounded	Infeasible
Primal problem	Optimal	YES	NO	NO
	Unbounded	NO	NO	YES
	Infeasible	NO	YES	??

Duality in linear optimization

Motivation

- We consider now the specific case of linear optimization.
- We show that the dual problem has some nice properties in this context.

Linear optimization

p. 102

$$h(x) = b - Ax, g(x) = -x$$

$$\min_{x \in \mathbb{R}^n} c^T x$$

subject to

$$Ax = b$$

$$x \geq 0$$

$A \in \mathbb{R}^{m \times n}$
$b \in \mathbb{R}^m$
$c \in \mathbb{R}^n$

$$\begin{aligned} L(x, \lambda, \mu) &= c^T x + \lambda^T (b - Ax) - \mu^T x \\ &= \left(c - A^T \lambda - \mu\right)^T x + \lambda^T b \end{aligned}$$

Dual problem:

$$\max_{\lambda} \lambda^T b \text{ subject to } \mu = c - A^T \lambda \geq 0.$$

Equivalent formulation:

$$\min_{x} -b^T x \text{ subject to } A^T x \leq c.$$

Linear optimization

p. 103

$$\min_{x \in \mathbb{R}^m} -b^T x$$

subject to

$$A^T x \leq c$$

$A \in \mathbb{R}^{m \times n}$
$b \in \mathbb{R}^m$
$c \in \mathbb{R}^n$

$$g(x) = A^T x - c$$

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

Dual problem:

$$\max_{\lambda} -\mu^T c \text{ subject to } A\mu = b, \mu \geq 0.$$

Equivalent formulation:

$$\min_{x} c^T x \text{ subject to } Ax = b, x \geq 0.$$

Dual problem

Theorem 4.14

The dual of a linear optimization problem is another linear optimization problem. The dual can be derived for each possible specification.

Theorem 4.15

The dual of the dual is the primal.

Strong duality

Theorem 4.17

- Consider the primal problem and its dual.
- If one problem has an optimal solution, so does the other one,
- and the optimal value of their objective functions are the same.

Strong duality

Motivation

- An optimal solution of the primal is obtained when a basis is available such that all reduced costs are non negative.
- We show here that the same basis can be used to obtain an optimal solution of the dual problem.

Context

Data

$$A \in \mathbb{R}^{m \times n}$$

$$b \in \mathbb{R}^{m}$$

$$c \in \mathbb{R}^{n}$$

Context

Data

$$A \in \mathbb{R}^{m \times n}$$
$$b \in \mathbb{R}^{m}$$
$$c \in \mathbb{R}^{n}$$

Primal

$$\min_{x \in \mathbb{R}^n} c^T x$$

subject to

$$Ax = b$$
$$x \geq 0$$

Context

Data

$$A \in \mathbb{R}^{m \times n}$$
$$b \in \mathbb{R}^{m}$$
$$c \in \mathbb{R}^{n}$$

Primal

$$\min_{x \in \mathbb{R}^n} c^T x$$

subject to

$$Ax = b$$
$$x \geq 0$$

Dual

$$\max_{\lambda \in \mathbb{R}^m} \lambda^T b$$

subject to

$$A^T \lambda \leq c$$

Optimal solutions: primal

Assumptions

$$
\begin{aligned}
A &= (B|N) \\
B^{-1}b &\geq 0 \\
\bar{c} = c - A^T B^{-T} c_B &\geq 0
\end{aligned}
$$

Optimal solutions: primal

Assumptions

$$
\begin{aligned}
A &= (B|N) \\
B^{-1}b &\geq 0 \\
\bar{c} = c - A^T B^{-T} c_B &\geq 0
\end{aligned}
$$

Primal solution

$$x_B^* = B^{-1}b, \; x_N^* = 0.$$

Optimal solutions: dual

Assumptions

$$A = (B|N)$$
$$B^{-1}b \geq 0$$
$$\bar{c} = c - A^T B^{-T} c_B \geq 0$$

Dual solution

$$\lambda^* = B^{-T} c_B.$$

λ^* is dual-feasible

$$A^T \lambda^* = A^T B^{-T} c_B = c - \bar{c} \leq c$$

λ^* is dual-optimal

$$(\lambda^*)^T b = c_B^T B^{-1} b = c_B^T x_B^* = c^T x^*$$

Strong duality

- If the primal or the dual has an optimal solution,
- so does the other one,
- and the optimal objective values are equal.

Karush-Kuhn-Tucker optimality conditions

Theorem 6.13

- $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $h : \mathbb{R}^n \to \mathbb{R}^m$ continuously differentiable.
- x^* local optimal of the problem.
- If the constraints are "qualified" at x^*, there exists a unique $\lambda^* \in \mathbb{R}^m$, a unique $\mu^* \in \mathbb{R}^p$, $\mu^* \geq 0$, such that

$$\nabla_x L(x, \lambda^*, \mu^*) = 0,$$

  and

$$\mu_i^* g_i(x^*) = 0, \ i = 1, \ldots, p.$$

- Linear constraints are qualified.

Karush-Kuhn-Tucker optimality conditions

$$L(x, \lambda, \mu) = c^T x + \lambda^T (b - Ax) - \mu^T x.$$

$$\nabla_x L(x^*, \lambda^*, \mu^*) = c - A^T \lambda^* - \mu^* = 0. \text{ [Dual constraints].}$$

$$\mu^* = c - A^T \lambda^* \geq 0 \iff c - A^T B^{-T} c_B \geq 0. \text{ [Non negative reduced costs].}$$

$$\mu_i^* x_i^* = 0. \text{ [Complementarity slackness].}$$

$$(c_i - \sum_{j=1}^{m} a_{ji} \lambda_j^*) x_i^* = 0$$

Complementarity slackness

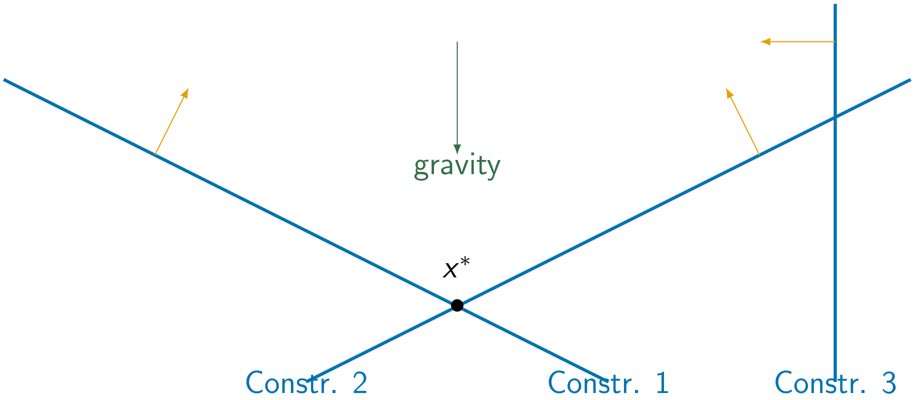
Primal problem

- If a constraint of the dual is not active, the corresponding primal variable is zero.
- If a primal variable is non zero, the corresponding dual constraint is active.

Dual problem

- If a constraint of the primal is not active, the corresponding dual variable is zero.
- If a dual variable is non zero, the corresponding primal constraint is active.

Physical interpretation



Optimization problem

Problem solved by mother nature

$$\min_{x \in \mathbb{R}^2} c^T x = x_2$$

subject to

$$a_{11}x_1 + a_{12}x_2 \geq b_1,$$
$$a_{21}x_1 + a_{22}x_2 \geq b_2,$$
$$a_{31}x_1 + a_{32}x_2 \geq b_3,$$

$$\min_{x \in \mathbb{R}^2} c^T x = x_2$$

subject to

$$a_1^T x \geq b_1,$$
$$a_2^T x \geq b_2,$$
$$a_3^T x \geq b_3,$$

where

$$a_i = \begin{pmatrix} a_{i1} \\ a_{i2} \end{pmatrix}.$$

Dual problem

$$\max_{\mu \in \mathbb{R}^3} b_1\mu_1 + b_2\mu_2 + b_3\mu_3,$$

subject to

$$\sum_{i=1}^{3} a_i\mu_i = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = c.$$

Forces

- Constraints: $a_i\mu_i$.
- Gravity: $-c = -\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
- Constraints of the dual: sum of forces is zero.

Complementarity slackness

$$\mu_i^*(a_i^T x^* - b_i) = 0, \; i = 1, 2, 3.$$

Only active constraints apply a force.

Summary

- Constraint relaxation and Lagrangian.
- Dual function: lower bound.
- Dual problem: best lower bound (weak duality).
- Linear optimization: strong duality.
- Complementarity slackness.