3 - Lagrangian duality

Mauro Passacantando

Department of Computer Science, University of Pisa
mauro.passacantando@unipi.it

Optimization Methods and Game Theory
Master of Science in Artificial Intelligence and Data Engineering
University of Pisa – A.Y. 2021/22

Lagrangian relaxation

Consider the general optimization problem

$$\begin{cases} \min \ f(x) \\ g_i(x) \leq 0 & i = 1, \dots, m \\ h_j(x) = 0 & j = 1, \dots, p \end{cases} \tag{P}$$

where $x \in \mathcal{D}$ and $v(P)$ denotes the optimal value.

The Lagrangian function $L : \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is

$$L(x, \lambda, \mu) := f(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x) + \sum_{j=1}^{p} \mu_j \, h_j(x)$$

Lagrangian relaxation and dual function

Definition

Given $\lambda \geq 0$ and $\mu \in \mathbb{R}^p$, the problem

$$\begin{cases} \min \ L(x, \lambda, \mu) \\ x \in \mathcal{D} \end{cases}$$

is called Lagrangian relaxation of (P) and
$\varphi(\lambda, \mu) = \inf\limits_{x \in \mathcal{D}} L(x, \lambda, \mu)$ is the Lagrangian dual function.

Dual function φ

- is concave because inf of affine functions w.r.t (λ, μ)
- may be equal to $-\infty$ at some point
- may be not differentiable at some point

Lagrangian relaxation and dual function

Lagrangian relaxation provides a lower bound to $v(P)$.

Theorem

For any $\lambda \geq 0$ and $\mu \in \mathbb{R}^p$, we have $\varphi(\lambda, \mu) \leq v(P)$.

Proof. If $x \in \Omega$, i.e. $g(x) \leq 0, h(x) = 0$, then

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x) \leq f(x),$$

hence

$$\varphi(\lambda, \mu) = \min_{x \in \mathcal{D}} \; L(x, \lambda, \mu) \leq \min_{x \in \Omega} \; L(x, \lambda, \mu) \leq \min_{x \in \Omega} \; f(x) = v(P)$$

$\square$

Lagrangian dual problem

Definition

The problem

$$\begin{cases} \max \ \varphi(\lambda, \mu) \\ \lambda \geq 0 \end{cases} \tag{D}$$

is called Lagrangian dual problem of (P) [and (P) is called primal problem].

- The dual problem (D) consists in finding the best lower bound of $v(P)$.
- (D) is always a convex problem, even if (P) is a non-convex problem.

Example 1 - Linear Programming.

Primal problem:

$$\begin{cases} \min \ c^{\mathsf{T}}x \\ Ax \geq b \end{cases} \tag{P}$$

Lagrangian function: $L(x,\lambda) = c^{\mathsf{T}}x + \lambda^{\mathsf{T}}(b - Ax) = \lambda^{\mathsf{T}}b + (c^{\mathsf{T}} - \lambda^{\mathsf{T}}A)x$
Dual function:

$$\varphi(\lambda) = \min_{x \in \mathbb{R}^n} L(x,\lambda) = \begin{cases} -\infty & \text{if } c^{\mathsf{T}} - \lambda^{\mathsf{T}}A \neq 0 \\ \lambda^{\mathsf{T}}b & \text{if } c^{\mathsf{T}} - \lambda^{\mathsf{T}}A = 0 \end{cases}$$

Dual problem:

$$\begin{cases} \max \ \varphi(\lambda) \\ \lambda \geq 0 \end{cases} \longrightarrow \begin{cases} \max \ \lambda^{\mathsf{T}}b \\ \lambda^{\mathsf{T}}A = c^{\mathsf{T}} \\ \lambda \geq 0 \end{cases} \tag{D}$$

is a linear programming problem.

Exercise 3.1. What is the dual of (D)?

Lagrangian dual problem

Example 2 - Least norm solution of linear equations.

Primal problem:

$$\begin{cases} \min \ \frac{1}{2}x^{\mathsf{T}}x \\ Ax = b \end{cases} \tag{P}$$

Lagrangian function: $L(x,\mu) = \frac{1}{2}x^{\mathsf{T}}x + \mu^{\mathsf{T}}(b - Ax)$.
Dual function: $\varphi(\mu) = \min_{x\in\mathbb{R}^n} L(x,\mu)$.

$L(x,\mu)$ is quadratic and strongly convex with respect to x, thus the global optimum is the stationary point:

$$\nabla_x L = x - A^{\mathsf{T}}\mu = 0 \Longleftrightarrow x = A^{\mathsf{T}}\mu,$$

hence $\varphi(\mu) = -\frac{1}{2}\mu^{\mathsf{T}}AA^{\mathsf{T}}\mu + b^{\mathsf{T}}\mu$.
Dual problem:

$$\begin{cases} \max \ -\frac{1}{2}\mu^{\mathsf{T}}AA^{\mathsf{T}}\mu + b^{\mathsf{T}}\mu \\ \mu \in \mathbb{R}^p \end{cases} \tag{D}$$

is an unconstrained convex quadratic programming problem.

Lagrangian dual problem

Exercise 3.2. Find the dual problem of a general convex quadratic programming problem

$$\begin{cases} \min\ \dfrac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x \\ Ax \leq b \end{cases} \tag{P}$$

where Q is a symmetric positive definite matrix.

Weak duality

Theorem (weak duality)

For any optimization problem (P), we have $v(D) \leq v(P)$.

Strong duality, i.e., $v(D) = v(P)$, does not hold in general.

Example. Consider the following (non-convex) problem with one variable:

$$\begin{cases} \min \ -x^2 \\ x - 1 \leq 0 \\ -x \leq 0 \end{cases} \tag{P}$$

It is easy to check that $v(P) = -1$.
The Lagrangian function is $L(x, \lambda) = -x^2 + \lambda_1(x - 1) - \lambda_2 x$, hence

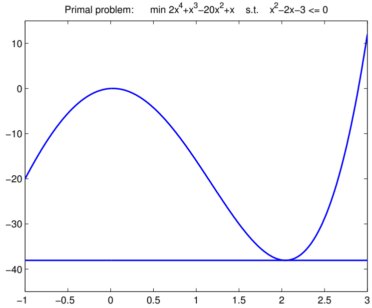
$$\varphi(\lambda) = \min_{x \in \mathbb{R}} L(x, \lambda) = -\infty \qquad \forall\, \lambda \in \mathbb{R}^2,$$

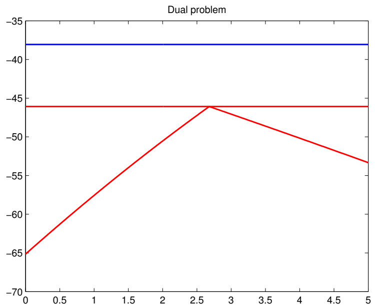
hence $v(D) = -\infty$.

Weak duality

Example. Consider the following (non-convex) problem with one variable:

$$\begin{cases} \min \ 2x^4 + x^3 - 20x^2 + x \\ x^2 - 2x - 3 \le 0 \end{cases}$$





Primal optimal solution $x^* \simeq 2.0427$, $v(P) \simeq -38.0648$.
Dual optimal solution $\lambda^* \simeq 2.68$, $v(D) \simeq -46.0838$.

Strong duality

Theorem (strong duality)

Suppose f, g, h are continuously differentiable, the primal problem

$$\begin{cases} \min \ f(x) \\ g(x) \leq 0 \\ h(x) = 0 \end{cases} \tag{P}$$

is convex, there exists a global optimum x^* and ACQ holds at x^*. Then:

- $v(D) = v(P)$
- (λ^*, μ^*) is optimal for (D) if and only if (λ^*, μ^*) is a KKT multipliers vector associated to x^*.

Strong duality

Proof. $L(x, \lambda, \mu)$ is convex with respect to x since (P) is convex.
Let (λ^*, μ^*) be any KKT multipliers vector associated to x^*. Then,

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0 \qquad \lambda^* \geq 0, \qquad (\lambda^*)^\mathsf{T} g(x^*) = 0.$$

Thus,

$$\begin{aligned} v(D) &\geq \varphi(\lambda^*, \mu^*) = \min_x L(x, \lambda^*, \mu^*) \underset{[L \text{ convex}]}{=} L(x^*, \lambda^*, \mu^*) \\ &= f(x^*) + (\lambda^*)^\mathsf{T} g(x^*) + (\mu^*)^\mathsf{T} h(x^*) = f(x^*) = v(P) \underset{[\text{weak duality}]}{\geq} v(D). \end{aligned}$$

Therefore, $v(P) = v(D)$ and (λ^*, μ^*) is optimal for (D).
Viceversa, if (λ^*, μ^*) is any optimal solution for (D), then

$$\begin{aligned} f(x^*) &= v(P) = v(D) = \varphi(\lambda^*, \mu^*) = \min_x L(x, \lambda^*, \mu^*) \leq L(x^*, \lambda^*, \mu^*) \\ &= f(x^*) + (\lambda^*)^\mathsf{T} g(x^*) + (\mu^*)^\mathsf{T} h(x^*) = f(x^*) + (\lambda^*)^\mathsf{T} g(x^*) \leq f(x^*), \end{aligned}$$

thus $(\lambda^*)^\mathsf{T} g(x^*) = 0$ and $\min_x L(x, \lambda^*, \mu^*) = L(x^*, \lambda^*, \mu^*)$, hence $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$, i.e., (λ^*, μ^*) is a KKT multipliers vector associated to x^*. $\square$

Strong duality

Strong duality may hold also for some non-convex problems.

Example. Consider the (non-convex) problem

$$\begin{cases} \min \ -x_1^2 - x_2^2 \\ x_1^2 + x_2^2 - 1 \leq 0 \end{cases}$$

We have $v(P) = -1$. The Lagrangian function is

$$L(x, \lambda) = -x_1^2 - x_2^2 + \lambda(x_1^2 + x_2^2 - 1) = (\lambda - 1)x_1^2 + (\lambda - 1)x_2^2 - \lambda.$$

The dual function is

$$\varphi(\lambda) = \begin{cases} -\infty & \text{if } \lambda < 1 \\ -\lambda & \text{if } \lambda \geq 1 \end{cases}$$

The dual problem is

$$\begin{cases} \max \ -\lambda \\ \lambda \geq 1 \end{cases}$$

hence its optimal solution is $\lambda^* = 1$ and $v(D) = -1$.

Exercises

Exercise 3.3. Consider the problem

$$\begin{cases} \min \ \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i \geq 1 \end{cases}$$

- Discuss existence and uniqueness of optimal solutions
- Find the optimal solution and the optimal value
- Write the dual problem
- Solve the dual problem and check whether strong duality holds

Exercise 3.4. Given $a, b \in \mathbb{R}$ with $a < b$, consider the problem

$$\begin{cases} \min \ x^2 \\ a \leq x \leq b \end{cases}$$

- Find the optimal solution and the optimal value for any a, b
- Solve the dual problem and check whether strong duality holds