Duality

Yuqian Zhang

Rutgers University

yqz.zhang@rutgers.edu

February 15, 2021

Overview

1. Lagrange dual problem
2. Weak and strong duality
3. KKT condition

Lagrangian

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\
& h_i(x) = 0, \quad i = 1, \cdots, p
\end{aligned}
$$

- Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ with $\mathrm{dom}(L) = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$

$$
L(x, \lambda, v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)
$$

- Lagrange dual function $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$
g(\lambda, v) = \inf_{x \in \mathcal{D}} L(x, \lambda, v) = \inf_{x \in \mathcal{D}} f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)
$$

Properties

$$L(x, \lambda, v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)$$

$$g(\lambda, v) = \inf_{x \in \mathcal{D}} L(x, \lambda, v) = \inf_{x \in \mathcal{D}} f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)$$

Independent of the convexity of f_i's, we have

- $L(x, \lambda, v)$ is linear and convex in λ and v
- $-g(\lambda, v)$ is convex in λ and v
- $g(\lambda, v)$ is concave independent of the convexity of f_i's

Lower bound property

$$L(x, \lambda, v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)$$

$$g(\lambda, v) = \inf_{x \in \mathcal{D}} L(x, \lambda, v) = \inf_{x \in \mathcal{D}} f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)$$

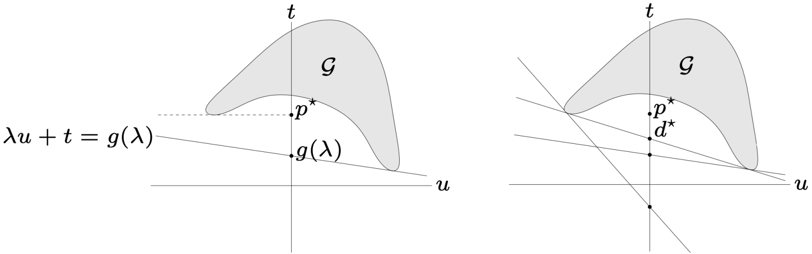
Lower bound property

If $\lambda_i \geq 0$, then $g(\lambda, v) \leq p^*$.

Geometric picture

interpretation for dual function:

$$g(\lambda) = \inf_{(u,t)\in\mathcal{G}} (t + \lambda u), \quad \mathcal{G}\left\{(f_1(x), f_0(x)) \mid x \in \mathcal{D}\right\}$$



- $g(\lambda) = \lambda u + t$ is (non-vertical) supporting hyperplane to $\mathcal{G}$
- hyperplane intersects t-axis at $t = g(\lambda)$

Primal and dual problem

Primal problem with optimal value p^*

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\ & h_i(x) = 0, \quad i = 1, \cdots, k \end{aligned}$$

Dual problem with optimal value d^*

$$\begin{aligned} \text{maximize} \quad & g(\lambda, v) \\ \text{subject to} \quad & \lambda \geq 0 \end{aligned}$$

Weak and strong duality

Weak duality $d^* \leq p^*$

- always holds (for convex and nonconvex problems)
- can be used to find nontrivial lower bounds for difficult problems

Strong duality $d^* = p^*$

- does not hold in general
- (usually) holds for convex problems
- Slater's condition

Theorem

Strong duality holds for a convex optimization if slater's condition holds.

Slater's constraint qualification

Strong duality holds for a convex problem

$$\begin{array}{rll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & Ax = b & \end{array}$$

if it is strictly feasible

$$\exists x \in \mathrm{int}\, \mathcal{D}: \quad \begin{array}{ll} f_i(x) < 0 & i = 1, \cdots, m, \\ Ax = b & \end{array}$$

- there exist many other types of constraint qualifications

Inequality form LP

Primal problems

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \end{aligned}$$

Dual problem

$$\begin{aligned} \text{maximize} \quad & -b^T \lambda \\ \text{subject to} \quad & A^T \lambda + c = 0 \end{aligned}$$

- from Slater's condition: $p^* = d^*$ if $A\tilde{x} < b$ for some $\tilde{x}$
- in fact, $p^* = d^*$ if primal and dual are feasible

Quadratic program

Primal problem, assume $P \in \mathbb{R}^n_{++}$

$$\begin{array}{rl} \text{minimize} & x^T P x \\ \text{subject to} & Ax \leq b \end{array}$$

Dual problem

$$\begin{array}{rl} \text{maximize} & -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \text{subject to} & \lambda \geq 0 \end{array}$$

- from Slater's condition: $p^* = d^*$ if $A\tilde{x} < b$ for some $\tilde{x}$
- in fact, $p^* = d^*$ always

Max-min interpretation

Consider inequality constrained optimization problem

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \end{aligned}$$

then

$$\begin{aligned} \sup_{\lambda \geq 0} L(x, \lambda) &= \sup_{\lambda \geq 0} \left(f_0(x) + \sum \lambda_i f_i(x) \right) \\ &= \begin{cases} f_0(x) & f_i(0) \leq 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

Max-min interpretation

$$p^* = \inf_x \sup_{\lambda \geq 0} L(x, \lambda), \quad d^* = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$$

Thus weak duality

$$\sup_{\lambda \geq 0} \inf_x L(x, \lambda) \leq \inf_x \sup_{\lambda \geq 0} L(x, \lambda)$$

strong duality

$$\sup_{\lambda \geq 0} \inf_x L(x, \lambda) = \inf_x \sup_{\lambda \geq 0} L(x, \lambda)$$

Above inequality holds for any function L.

Complementary slackness

Assume strong duality holds, x^* is primal optimal, (λ^*, v^*) is dual optimal

$$
\begin{aligned}
f_0(x^*) = g(\lambda^*, v^*) &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p v_i h_i(x) \right) \\
&\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \\
&\leq f_0(x^*)
\end{aligned}
$$

- x^* minimizes $L(x, \lambda^*, v^*)$
- $\lambda_i^* f_i(x^*) = 0$, known as complementary slackness

$$
\lambda_i^* > 0 \implies f_i(x^*) = 0, \quad f_i(x^*) < 0 \implies \lambda_i^* = 0
$$

Karush-Kuhn-Tucker (KKT) conditions

The following four conditions are called KKT conditions (for a problem with differentiable f_i and h_i)

1. primal constraints: $f_i \leq 0$ and $h_i = 0$
2. dual constraints: $\lambda \geq 0$
3. complementary slackness: $\lambda_i f_i(x) = 0$
4. gradient of Lagrangian w.r.t. x vanishes

$$\nabla f_0(x) + \sum \lambda_i \nabla f_i(x) + \sum v_i \nabla h_i(x) = 0$$

Karush-Kuhn-Tucker (KKT) conditions

If $\tilde{x}, \tilde{\lambda}, \tilde{v}$ satisfy KKT for a convex problem, then they are optimal:

- from complementary slackness: $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{v})$
- from 4th condition: $g(\tilde{\lambda}, \tilde{v}) = L(\tilde{x}, \tilde{\lambda}, \tilde{v})$

hence $f_0(\tilde{x}) = g(\tilde{\lambda}, \tilde{v})$.

Theorem

Every local minimum satisfies KKT, under regularity conditions.

Theorem

For convex optimizations problems satisfying Slater's condition, KKT is necessary and sufficient for optimality.

Examples

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^T P x + q^T x + r, \quad P \succeq 0 \\ \text{subject to} \quad & Ax = b \end{aligned}$$

- Slater's condition holds as long as $Ax = b$ is feasible
- $L(x, v) = \frac{1}{2}x^t P x + q^T x + r + v^T(Ax - b)$
- KKT conditions: $Ax^* = b$ and $\nabla_x L = Px^X + A^T v^* + q = 0$

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

Examples

$$\begin{aligned} \text{minimize} \quad & \sum_i x_i \log(x_i) \\ \text{subject to} \quad & Ax \leq b \\ & \mathbf{1}^T x = 1 \end{aligned}$$

Dual function

$$g(\lambda, v) = \inf_x \left(\sum_i x_i \log(x_i) + \lambda^T (Ax - b) + v(1^T x - 1) \right)$$

$$\nabla_{x_i} L(x, \lambda, v) = 1 + \log(x_i) + a_i^T \lambda + v$$

Examples

By first order optimality, the infimum for Lagrangian is obtained when

$$\nabla_{x_i} L(x, \lambda, v) = 1 + \log(x_i) + a_i^T \lambda + v = 0$$

then $x_i^* = e^{-1-a_i^T\lambda - v}$, then

$$\begin{aligned}
g(\lambda, v) &= \inf_x L(x, \lambda, v) = L(x^*, \lambda, v) \\
&= \sum_i \left(-1 - a_i^T \lambda - v\right) e^{-1-a_i^T\lambda - v} \\
&\quad + \lambda^T (A e^{-\mathbf{1} - A^T\lambda - v\mathbf{1}} - b) + v(\mathbf{1}^T e^{-\mathbf{1} - A^T\lambda - v\mathbf{1}} - 1) \\
&= -b^T\lambda - v - e^{-v-1} \sum_i e^{-a_i^T\lambda}
\end{aligned}$$

Examples

Primal Problem

$$\begin{aligned} \text{minimize} \quad & \sum_i x_i \log(x_i) \\ \text{subject to} \quad & Ax \leq b \\ & \mathbf{1}^T x = 1 \end{aligned}$$

Dual Problem

$$\begin{aligned} \text{maximize} \quad & -b^T\lambda - v - e^{-v-1}\sum_i e^{-a_i^T\lambda} \\ \text{subject to} \quad & \lambda \geq 0 \end{aligned}$$

Assume we have solved the dual problem and found λ^*, then the primal solution is $x_i^* = e^{-1-a_i^T\lambda^*-v^*}$

Examples

Primal Problem

$$\begin{aligned} \text{minimize} \quad & \sum_i f_i(x_i) \\ \text{subject to} \quad & a^T x = b \end{aligned}$$

Dual function

$$\begin{aligned} g(v) = \inf_x L(x, v) &= \inf_x \sum_i f_i(x_i) + v(a^T x - b) \\ &= -bv + \sum_i \inf_{x_i} \left(f_i(x_i) + v a_i x_i \right) \end{aligned}$$

Denote $\tilde{f}_i(v) = \sup_{x_i} \left(f_i(x_i) + v a_i x_i \right)$ is convex, then $-\tilde{f}_i(v)$ is concave

Examples

Primal Problem

$$\begin{aligned} \text{minimize} \quad & \sum_i x_i \log(x_i) \\ \text{subject to} \quad & Ax \leq b \\ & \mathbf{1}^T x = 1 \end{aligned}$$

Dual Problem: unstrained problem with only one variable v

$$\text{maximize} \quad -bv - \sum_i f_i(\tilde{v})$$

Examples

Primal Problem

$$\begin{aligned} \text{minimize} \quad & -3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3) \\ \text{subject to} \quad & x_1^2 + x_2^2 + x_3^2 = 1 \end{aligned}$$

KKT Condition

$$x_1^2 + x_2^2 + x_3^2 = 1 \qquad \begin{aligned} -6x_1 + 2 + 2vx_1 &= 0 \\ 2x_2 + 2 + 2vx_2 &= 0 \\ 4x_3 + 2 + 2vx_3 &= 0 \end{aligned}$$