Convex Optimization

Daniel Mao

Contents

Chapter 1

Linear Programming

1.1 Primal Problem and Dual Problem

DEFINITION 1.1 (Primal Problem)**.** Let $A \in \mathbb{R}^{m \times n}$. Let $b \in \mathbb{R}^m$. We define the **primal problem** to be the following.

$$\begin{aligned} \text{(LP)} \quad & \text{minimize} \quad && c^\top x \\ & \text{subject to} \quad && Ax = b \\ & && x \geq 0 \end{aligned}$$

DEFINITION 1.2 (Dual Problem)**.** We define the **dual problem** of the above primal problem to be the following.

$$\begin{aligned} \text{(LD)} \quad & \text{maximize} \quad && b^\top y \\ & \text{subject to} \quad && A^\top y + s = c \\ & && s \geq 0 \end{aligned}$$

1.2 Farkas' Lemma

LEMMA 1.3 (Farkas' Lemma)**.** Let $A \in \mathbb{R}^{m \times n}$. Let $b \in \mathbb{R}^m$. Then exactly one of the following systems has a solution.

1. $Ax = b$, $x \geq 0$.

2. $A^\top y \leq 0$, $b^\top y > 0$.

1.3 The Fundamental Theorem of Linear Programming

THEOREM 1.4 (The Fundamental Theorem of Linear Programming)**.** Every linear programming problem has exactly one of the following properties:

- The linear programming is infeasible.
- The linear programming is unbounded.
- The linear programming has an optimal solution.

1.4 Properties

PROPOSITION 1.5. If the feasible region of an LP problem is a pointed polyhedron, then

1. whenever the LP problem is feasible, it has a feasible solution that is an extreme point of the feasible region;
2. whenever the LP problem has optimal solution(s), it has an optimal solution that is an extreme point of the feasible region.

THEOREM 1.6 (Duality Theorem - 1)**.** If a LP has an optimal solution, then so does its LD and their optimal values are the same.

THEOREM 1.7 (Duality Theorem - 2)**.** If a LP and its LD both have feasible solutions, then they both have optimal solutions and their optimal values are the same.

Chapter 2

Minimizers

2.1 Local Minimizers and Global Minimizers

PROPOSITION 2.1. Let f be a proper convex function. Then any local minimizer of f is a global minimizer.

Proof Appoach 1.

Let f be a convex function.

Let x_0 be a local minimizer of f, if any.

Since x_0 is a local minimizer, $\exists \delta > 0$, $\forall x \in \text{ball}(x_0, \delta)$, we have $f(x) \geq f(x_0)$.

Since f is proper, $\text{dom}(f) \neq \emptyset$.

Let y be an arbitrary point in $\text{dom}(f)$.

Case 1. $y \in \text{ball}(x_0, \delta)$.

Since $y \in \text{ball}(x_0, \delta)$, and $\forall x \in \text{ball}(x_0, \delta), f(x) \geq f(x_0)$, we get $f(y) \geq f(x_0)$.

Case 2. $y \notin \text{ball}(x_0, \delta)$.

Define $\lambda := \delta / \|x - y\|$.

Since $y \notin \text{ball}(x_0, \delta)$, $\|x - y\| > 0$.

Since $\delta > 0$ and $\|x - y\| > 0$, we get $\lambda > 0$.

Since $y \notin \text{ball}(x_0, \delta)$, $\|x - y\| > \delta$.

Since $\delta < \|x - y\|$, $\lambda < 1$.

Define a point $z := \lambda y + (1 - \lambda) x$.

Since f is convex, $\text{dom}(f)$ is convex.

Since

$\square$

PROPOSITION 2.2. Any locally optimal point of a convex problem is globally optimal.

PROPOSITION 2.3. A point x is optimal if and only if it is feasible and for any feasible point y,

$$\nabla f_0(x) \cdot (y - x) \geq 0.$$

I forgot where this came from... and I don't know what it's talking about...

2.2 Main Results

THEOREM 2.4. Let f be a proper function from $\mathbb{E}$ to $\mathbb{R}^*$. Then

$$\operatorname{argmin}(f) = \big\{x \in \mathbb{E} : 0 \in \partial f(x)\big\}.$$

Proof.

$$\begin{aligned}
& x \in \operatorname{argmin}(f) \\
\iff & \forall y \in \mathbb{E}, f(x) \leq f(y) \\
\iff & \forall y \in \mathbb{E}, \langle 0, y - x \rangle + f(x) \leq f(y) \\
\iff & 0 \in \partial f(x).
\end{aligned}$$

□

THEOREM 2.5. Let f be a proper, convex, and lower semi-continuous function from $\mathbb{R}^d$ to $\mathbb{R}$. Let x be a point in $\mathbb{R}^d$. Then x is a global minimizer of f if and only if x is a fixed point of the proximal operator of f. i.e. $x = \operatorname{prox}_f(x)$.

Chapter 3

Duality

3.1 Definitions

DEFINITION 3.1 (Dual Problem).

3.2 Lagrangian Dual

3.2.1 Basics

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

where $x \in \mathcal{D} \subseteq \mathbb{R}^n$.

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

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

Lagrange Dual Function: $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$.

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu).$$

PROPOSITION 3.2. The Lagrange dual function is concave.

Proof. The Lagrange dual function is an infimum of an affine function and hence concave. □

PROPOSITION 3.3. If $\lambda \geq 0$, then $g(\lambda, \nu) \leq p^*$ where p^* denotes the optimal value of the primal problem.

Proof. Let $\bar{x}$ be an arbitrary feasible solution. Then

$$f_0(\bar{x}) \geq L(\bar{x}, \lambda, \nu) \geq \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = g(\lambda, \nu).$$

□

3.2.2 Dual of Linear Programming

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \geq 0 \end{array}$$

The Lagrangian is

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

The Lagrange dual function is

$$g(\lambda, \nu) = \begin{cases} -b^T \nu, & A^T \nu - \lambda + c = 0 \\ -\infty, & \text{otherwise} \end{cases}.$$

Note 1: g is linear on an affine domain: $\{(\lambda, \nu) : A^T \nu - \lambda + c = 0\}$ and hence concave.
Note 2: The Lower Bound Property says that if $\lambda = A^T \nu + c \geq 0$, then $p^* \geq -b^T \nu$.
Lagrange Dual Problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \geq 0 \end{array}$$

Standard form LP and its dual:

$$\text{(LP)} \quad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b,\ x \geq 0 \end{array}$$

$$\text{(Dual of LP)} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \geq 0 \end{array}$$

3.3 Weak Dual and Strong Dual

Weak Duality: $d^* \leq p^*$.

String Duality: $d^* = p^*$.

THEOREM 3.4 (Slater)**.** Consider an optimization problem

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

where $f_0, f_1, ..., f_m$ are all convex functions. Then the strong duality holds if there exists a point x^* in $\mathrm{ri}(\mathcal{D})$ where $\mathcal{D} := \mathrm{dom}(f_0) \cap \bigcap_{i=1}^{m} \mathrm{dom}(f_i)$ such that $f_i(x^*) < 0$ for $i = 1, ..., m$ and $Ax^* = b$.

THEOREM 3.5 (Complementary Slackness)**.** Consider an optimization problem and its dual:

$$\text{(Primal)} \quad \begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0 & i = 1, ..., m \\ & h_i(x) = 0 & i = 1, ..., p \end{array} \qquad \text{(Dual)} \quad \begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \geq 0 \end{array}$$

Let x be a feasible solution to the primal and (λ, ν) be a feasible solution to the dual. Then x and (λ, ν) are both optimal if and only if

$$\lambda_i f_i(x) = 0$$

for each $i = 1, ..., m$. i.e.,

$$\lambda_i > 0 \implies f_i(x) = 0, \text{ and } f_i(x) < 0 \implies \lambda_i = 0$$

for each $i = 1, ..., m$.

3.4 Weak Duality Theorem

THEOREM 3.6 (Weak Duality Theorem)**.** The duality gap is always greater than or equal to 0.

3.5 Perturbation

Primal

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

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq u_i & i = 1, ..., m \\ & h_i(x) = v_i & i = 1, ..., p \end{array}$$