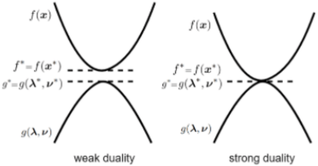
Optimization in Machine Learning

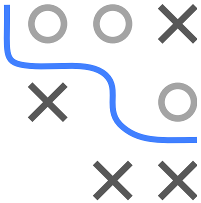
Duality in optimization



Learning goals

- Awareness of the concept of duality in optimization
- LP duality
- Weak and strong duality in LP

DUALITY: OVERVIEW



- Duality theory plays a fundamental role in (constrained) optimization. The concept of “duality" emerged in the context of LPs and dates back to the 1940s (works of Tucker and Wolfe).
- There are several different types of duality: LP duality, Lagrangian duality, Wolfe duality, Fenchel duality (which can lead to confusion).
- Key take-home message: The concepts of duality give you recipes to find **lower bounds** on your original “primal" constrained optimization problem. Under certain conditions, these lower bounds are actually identical to the optimal solution.
- Duality is also practical. It has been used to find **better algorithms** for solving constrained optimization problems

LP DUALITY: INTRODUCTORY EXAMPLE

Example:

A bakery sells brownies for 50 ct and mini cheesecakes for 80 ct each. The two products contain the following ingredients

	Chocolate	Sugar	Cream cheese
Brownie	3	2	2
Cheesecake	0	4	5

A student wants to minimize his expenses, but at the same time eat at least 6 units of chocolate, 10 units of sugar and 8 units of cream cheese.

MATHEMATICAL INTUITION

The example explained duality from an economic point of view. But what is the mathematical intuition behind duality?

Idea: In minimization problems one is often interested in **lower bounds** of the objective function. How could we derive a lower bound for the problem above?

If we “skillfully” multiply the three inequalities by factors and add factors (similar to a linear system), we can find a lower bound.

DUALITY

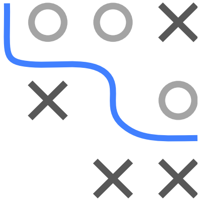
Dual problem:

$$\begin{aligned} \max_{\boldsymbol{\alpha} \in \mathbb{R}^m} \quad & g(\boldsymbol{\alpha}) := \boldsymbol{\alpha}^\top \mathbf{b} \\ \text{s.t.} \quad & \boldsymbol{\alpha}^\top \mathbf{A} \leq \mathbf{c}^\top \\ & \boldsymbol{\alpha} \geq 0 \end{aligned}$$

Primal problem:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & f(\mathbf{x}) := \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

DUALITY THEOREM



In general, the **weak duality theorem** applies to all feasible $\mathbf{x}, \alpha$

$$g(\alpha) = \alpha^\top \mathbf{b} \leq \mathbf{c}^\top \mathbf{x} = f(\mathbf{x})$$

The value of the dual function is therefore **always** a lower bound for the objective function value of the primal problem.

Proof:

$$\alpha^\top \mathbf{b} \overset{\mathbf{Ax} \geq \mathbf{b}}{\leq} \alpha^\top \mathbf{Ax} \overset{\alpha^\top \mathbf{A} \leq \mathbf{c}^\top}{\leq} \mathbf{c}^\top \mathbf{x}$$

ALTERNATIVE LP FORMULATION

Unfortunately, many slightly different (but ultimately equivalent) formulations of primal and dual LPs exist in the literature.

One common alternative with inequality and equality constraints is often formulated as follows. Let $\mathbf{c} \in \mathbb{R}^d$, $\mathbf{b} \in \mathbb{R}^l$, $\mathbf{A} \in \mathbb{R}^{l \times d}$, $\mathbf{h} \in \mathbb{R}^k$, and $\mathbf{G} \in \mathbb{R}^{k \times d}$.

Then the primal LP is defined as

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^d} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{G}\mathbf{x} = \mathbf{h} \\ & \mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned}$$