Duality

Pranav Kumar Kota

Introduction

Consider the following problem:

$$\begin{aligned}
&min\ f_0(\underline{x}) \\
&\text{s.t.} \\
&f_i(\underline{x}) <= 0, \quad i = 1, 2, ..r \\
&h_i(\underline{x}) = 0, \quad i = 1, 2, ..s
\end{aligned}$$

Let p^* be its optimum value. Can we obtain bounds on $p*$ such that $l < p* < u$?

As a primitive guess, we could bound p^* by $-\infty / +\infty$, but it would be of no use. We would like to have a tight bound.

An Alternate Problem

The problem considered above is constrained. To state the same optimisation problem without constraints, we could represent it as:

$$\begin{aligned}
&min\ f_0(\underline{x}) + \sum_{i=1}^{r} I(f_i(\underline{x})) + \sum_{i=1}^{s} I_0(h_i(\underline{x})) \\
&\text{where,} \\
&I(y) = \begin{cases} 0 \text{ if } y <= 0 \\ \infty \text{ otherwise} \end{cases} \\
&I_0(y) = \begin{cases} 0 \text{ if } y = 0 \\ \infty \text{ otherwise} \end{cases}
\end{aligned}$$

It is clear from the definitions of I and I_0, that for any x which is infeasible, i.e not satisfying the constraints of the original problem, the function value is ∞. Hence, minimisation is implicit within the constraint set.

Due to their nature, the functions I, and I_0 can be called the "displeasure" terms.

Displeasure functions

The functions I and I_0 as defined above are non-differentiable, which is called hard displeasure. This is rather inconvenient as differentiablility in optimisation problems is very helpful.

So, we now consider the following definitions:

$$I(y) = \lambda y$$
$$I_0(y) = \nu y$$
I and I_0 are differentiable

Lagrangian

Modifying the proposed alternative representation in the earlier section with soft displeasure, we get:

$$\min \; f_0(\underline{x}) + \sum_{i=1}^{r} \lambda_i (f_i(\underline{x})) + \sum_{i=1}^{s} \nu_i (h_i(\underline{x})), \; \lambda_i \geq 0$$

This problem is a proxy for the original problem, called the Primal, and is called the Lagrangian.

The Lagrangian w.r.t the Primal(P), is denoted as

$$\mathcal{L}(\underline{x}, \underline{\lambda}, \underline{\nu}) = f_0(x) + \sum_{i=1}^{r} \lambda (f_i(\underline{x})) + \sum_{i=1}^{s} \nu (h_i(\underline{x})), \; \lambda_i \geq 0$$

Properties of the Lagrangian

For any feasible $\underline{x}$,

$$f_0(\underline{x}) \geq \mathcal{L}(\underline{x}, \underline{\lambda}, \underline{\nu})$$

Since $p^* = \inf_{x} f_0(\underline{x})$,

$$p^* = \inf_{\underline{x}} f_0(\underline{x}) \geq \inf_{\text{feasible } \underline{x}} \mathcal{L}(\underline{x}, \underline{\lambda}, \underline{\nu})$$
$$\implies p^* \geq \inf_{\underline{x}} \mathcal{L}(\underline{x}, \underline{\lambda}, \underline{\nu}), \quad \forall \lambda_i \geq 0, \nu$$

We define $g(\lambda, \nu) = \inf_{\underline{x}} \mathcal{L}(\underline{x}, \underline{\lambda}, \underline{\nu})$

$$\implies p^* \geq g(\underline{\lambda}, \underline{\nu}), \quad \forall \lambda_i \geq 0, \nu$$

To get the best lower bound, we solve the following optimisation problem:

$$\max g(\underline{\lambda}, \underline{\nu})$$
$$\text{s.t. } \lambda_i \geq 0$$

This problem (D) is called the dual to the problem P.

$\underline{\lambda}$ and $\underline{\nu}$ are called dual variables or Lagrange multipliers.

Properties of the Dual Objective

$$
\begin{aligned}
g(\underline{\lambda}, \underline{\nu}) &= \inf_{\underline{x}} \mathcal{L}(\underline{x}, \underline{\lambda}, \underline{\nu}) \\
&= \inf_{\underline{x}} f_0(\underline{x}) + \sum_{i=1}^{r} \lambda_i(f_i(\underline{x})) + \sum_{i=1}^{s} \nu_i(h_i(\underline{x}))
\end{aligned}
$$

The dual objective g is the infimum of an affine function of λ and ν, i.e. the minimum of an affine(which is also concave).

$$\implies g(\underline{\lambda}, \underline{\nu}) \text{ is always a concave function}$$

Since the dual is the maximization of a concave function, it is always a convex optimisation problem.

Duality

Let d^* denote the optimal value of the dual problem.

Weak Duality

$$p^* \geq d^*$$

We began with trying to find bounds on the optimum value of the primal. We defined the Lagrangian, and showed that it's infimum lower bounds $p*$.

Since the dual problem is maximization of this lower bound (to make it tighter), $p^* \geq d^*$ always.

Strong Duality

This is the case when,

$$p^* = d^*$$

Illustration with Linear Programs

Consider the following linear optimization problem

$$\begin{aligned}
&\min \underline{c}^T \underline{x} \\
&\text{s.t} \\
&A\underline{x} = \underline{b} \\
&\underline{x} \geq 0
\end{aligned}$$

Equivalently, converting constraints to the notation used above

$$\begin{aligned}
&\min \underline{c}^T \underline{x} \\
&\text{s.t} \\
&A\underline{x} = \underline{b} \\
&-\underline{x} \leq 0
\end{aligned}$$

The Lagrangian ($\mathcal{L}(\underline{\lambda}, \underline{\nu})$) is given by

$$\begin{aligned}
&\underline{c}^T \underline{x} + \sum_{i=1}^{n} \lambda_i(-x_i) + \sum_{i=1}^{m} \nu_i(\underline{a_i}^T \underline{x} - b_i), \quad \underline{a_i} \text{ is the ith row vector of } A \\
&= \underline{c}^T \underline{x} - \underline{\lambda}^T \underline{x} + \underline{\nu}^T (A\underline{x} - \underline{b}) \\
&= (\underline{c}^T - \underline{\lambda}^T + \underline{\nu}^T A)\underline{x} - \underline{\nu b} \\
&= (c - \underline{\lambda} + A^T \underline{\nu})^T \underline{x} - \underline{\nu b}
\end{aligned}$$

The dual objective $g(\underline{\lambda}, \underline{\nu})$ is

$$\inf_{\underline{x}} \mathcal{L}(\underline{x}, \underline{\lambda}, \underline{\nu}) = \begin{cases} -\infty \text{ , if } \underline{c} - \underline{\lambda} + A^T \underline{\nu} \neq 0 \\ -\underline{\nu}^T \underline{b} \text{ , if } \underline{c} - \underline{\lambda} + A^T \underline{\nu} = 0 \end{cases}$$

Therefore, the Dual problem is

$$\begin{aligned}
&\max_{\underline{\lambda}, \underline{\nu}} g(\underline{\lambda}, \underline{\nu}) \\
&\text{s.t.} \\
&\lambda_i \geq 0 \\
&\underline{c} - \underline{\lambda} + A^T \underline{\nu} = 0
\end{aligned}$$

Equivalently,

$$\begin{aligned}
&\max_{\underline{\lambda}, \underline{\nu}} g(\underline{\lambda}, \underline{\nu}) \\
&\text{s.t. } \underline{c} + A^T \underline{\nu} \geq 0
\end{aligned}$$