# Two views of Duality: Lagrangians and Geometric

## 1 Linear Programs.

Recall last time, we discussed linear programs and their duals. Today, we discuss them in the context of a general manner of obtaining duals: Lagrangian ultipliers. We also discuss a proof of the strong duality theorem for linear programming.

| Primal LP | Dual LP |
|---|---|
| min $c.x$ | max $y^T b$ |
| $Ax \geq b$ | $y^T A \leq c$ |
| $x \geq 0$ | $y \geq 0$ |

## 2 Lagrangian Duals

Consider finding a feasible solution to the system:

$$f_i(x) \leq 0, i = 1, ..., m \tag{1}$$

The Lagrangian is written as follows:

$$L(x, \lambda) = \sum_{i=1}^{m} \lambda_i f_i(x)$$

where $\lambda_i \geq 0$ is called the Lagrangian multiplier associated with the $i$th inequality. It can can be viewed as a penalties for violating the constraint: $f_i(x) \leq 0$.

Notice that if there is no feasible solution for the constraint system, that the Lagrangian function can have an arbitrarily large value for any $x$.

Indeed, given $\lambda$ such that

$$\sum_i \lambda f_i(x) > 0$$

implies that there is no solution satisfiying the system 1; On the other hand, any solution satisfying the system should satisfy any positive linear combination of the constraints and thus there are no positive $f_i(x)$ and no positive setting of the $\lambda$'s will yield a positive value for the dual.

One can generalize the notion of dual for optimization problems[1] as follows:

[1] Equality constraints can be dealt with as well, where the lagrange multipliers will then be unrestricted.

$$\begin{aligned} \min \quad & f(x) && (2) \\ \text{subject to } f_i(x) \le 0, & \quad i = 1, ..., m && (3) \end{aligned}$$

The corresponding Lagrangian function will be

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

Here, in the case that there is a feasible primal solution with value $v$, the dual solution can only have nonzero values for $\lambda_i > 0$ with $f_i(x) = 0$. Moreover, the value of the best $\lambda$ will also be $v$.

We can also see that if there is a $\lambda$ where for all $x$, $L(x, \lambda) \ge \alpha$, that the value of the minimization function for problem 3-3 is at least $\alpha$; the optimal feasible solution for 3-3 yields this value for any dual value of $\lambda$.

The primal problem is to find an $x$ that minimizes the Lagrange function against any set of dual variables $\lambda$. The dual problem is to provide a $\lambda$ that maximizes the Lagrange function against any value of $x$. That is, the dual problem is to Find $\lambda$ which achieves $\min_\lambda \max_x L(x, \lambda)$.

For convex programs, these values will be the same. For any setup the dual provides a lower bound on the optimal solution of the primal. It is often used in that manner as a heuristic approach to solving constrained optimization problems.

### 2.1 Linear Program.

For a linear program, we have

$$\begin{aligned} \min \quad & c \cdot x \\ \text{subject to } b_i - a_i \cdot x \le 0, & \quad i = 1, ..., m \end{aligned}$$

And a Lagrangian formulation of

$$L(\lambda, x) = cx + \sum_i \lambda_i (b_i - a_i x_i).^2$$

We can rewrite the function as follows:

$$L(\lambda, x) = -(\sum_j x_j (a_j \lambda - c_j)) + b\lambda.$$

Translating back to a set of linear inequalities one gets the following problem.

$$\max b \cdot \lambda$$

[2] Note for any $x$, the minimizing $\lambda$ can only take nonzero values for $i$ where $(b_i - a_i x_i)$ is zero. This is the complementary slackness condition that we discussed before in the context of linear programs.

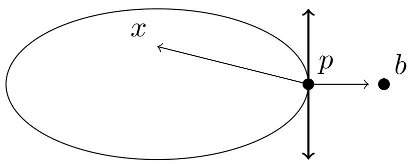



Figure 1: $(x-p)^T(b-p) < 0$ for $x$ in convex body.

$$a_j \cdot \lambda = c_j.$$
$$\lambda \geq 0$$

The linear programming dual of 4- 4!

That is the Lagrangian dual problem, of finding a lower bound for the Langrangian function for any $x$, is the linear programming dual.

Again, this technique applies more generally, but it is informative to see that the Lagrangian dual formulation is equivalent to linear programming. One might surmise that the nice part of some optimization problem may behave nicely with this formulation.

# 3 Strong Duality

We will begin a discussion of a the proof of strong duality here. We saw a proof based on experts for the special case of zero sum 2 person games. We discuss a geometric proof here. We begin with the notion of that a convex bodies can be separated using a hyperplane.

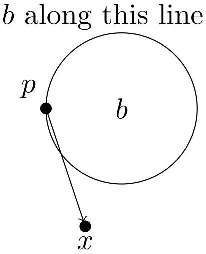
## 3.1 Convex Separator.

THEOREM 1
*For any convex body $A$, and a point $b$, either $b \in A$ or there exists a point, $p$, where $(x-p)^t(b-p) \leq 0$ for all $x \in A$.*

**Proof:** Choose $p$ as the closest point in $A$ to $b$ in the convex region.

For the sake of contradition, there is an $x \in A$, where $(x-p)^T(b-p) > 0$. For some intuition, note that the angle between $(x-p)$ and $(b-p)$ is less than 90.

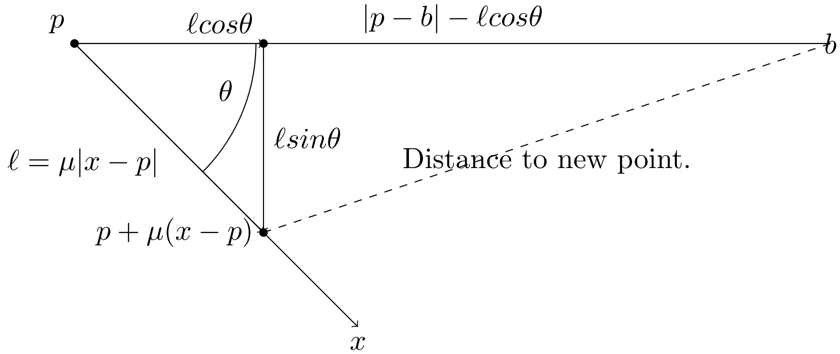
Since $A$ is convex every point between $p$ and $x$ is in $A$. Moreover, there must be a point closer to $b$ along this path. Again some intution. In the figure below, a portion of the line segment between $p$ and $x$ is inside the circle of radius $|p-b|$, and thus there is a closer point to $b$ along this line segment. This contradicts that $p$ is the closest point in $A$ to $p$.




To prove this formally, one can express the squared distance to $b$ from a point $p+(x-p)\mu$ (which for nonnegative $\mu \leq$ is some point between $p$ and $x$)

$$(|p-b| - \mu|x-p|cos\theta)^2 + (\mu|x-p|sin\theta)^2$$

where $\theta$ is the angle between $x-p$ and $b-p$. (See the figure below.)




Simplifying (and substituting that $\sin^2\theta = (1-\cos^2\theta)$, we obtain the following:

$$|p-b|^2 - 2\mu|p-b||x-p|cos\theta + (\mu|x-p|)^2.$$

Taking the derivative with respect to $\mu$ yields,
$-2|p-b||x-p|cos\theta + 2(\mu|x-p|)$.
which is negative for a small enough value of $\mu$ ( for positive $cos\theta$.)
**End of proof.**

### 3.2 Convex to Strong Duality.

We refer the reader to Goemans linear programming notes. (The proof there is does not have so much intuition. But it does do a translation from the simple geometric lemma above.)

What one should get from this picture is that there is a translation of the problem of solving an LP back and forth from the problem of finding a separating hyperplane for a convex region from a point.

The notion that there is a separating hyperplane follows from finding the point in the convex region that is "closest" to $b$.

One could proceed by identifying a closer point to $b$. This problem turns out to be finding a positive solution to a single linear equation; i.e., a dot product is positive. Thus, again, as with experts and the Lagrangian dual, one sees that satisfying many constraints can be approached by iteratively satisfying a linear combination of those constraints.