10-425/625: Introduction to Convex Optimization (Fall 2023)

Lecture 14: Duality in Linear Programs

Instructor:[1] *Matt Gormley* *October 11, 2023*

14.1 Summary: Rates of Convergence

Method	**Convergence Conditions**	**Suboptimality bound**	**Rate of Convergence (steps)**
Gradient Descent	smooth, possibly nonconvex	$O(1/\sqrt{k})$	$O(1/\epsilon^2)$
Gradient Descent	smooth, convex	$O(1/k)$	$O(1/\epsilon)$
Gradient Descent	smooth, strongly convex	$O(\gamma^k), \gamma \in (0,1)$	$O(\log(1/\epsilon))$
Subgradient Method	non-smooth, convex, G-Lipschitz, bounded subgradients	$O(1/\sqrt{k})$	$O(1/\epsilon^2)$
Projected Gradient Method	smooth, convex, constrained	$O(1/k)$	$O(1/\epsilon)$
Projected Gradient Method	smooth, strongly convex, constrained	$O(\gamma^k), \gamma \in (0,1)$	$O(\log(1/\epsilon))$
SGD	non-smooth, convex, bounded stochastic gradients	$O(1/\sqrt{k})$	$O(1/\epsilon^2)$

[1] These notes were originally written by Siva Balakrishnan for 10-725 Spring 2023 (original version: here) and were edited and adapted for 10-425/625.

14.2 What's next?

Now we'll depart the world of algorithms and return to talking about the structure of convex programs. Our focus will be on understanding the concept of *duality*. We'll see some uses of the concept of duality as we go along.

We will begin with a discussion of duality in linear programs. Often in LPs, the dual (also an LP) will be a nice reformulation of the original LP, so just writing down the dual will give you some insight into the original program. We'll also see that duality will give us an answer to a very basic question in optimization, given a candidate solution $\widehat{x}$ can we give a certificate of its optimality (we've done things like this before) and if it's not optimal can we give reasonable bounds on its sub-optimality, i.e. $f(\widehat{x}) - f(x^*)$.

14.3 Linear Programs

Linear programs (LPs) are a special sub-class of convex optimization problems. They were the focus of intense research during WWII, and the period after that. An LP is simply an optimization problem:

$$\begin{aligned} \min_x \; & c^T x \\ \text{subject to } & Ax = b \\ & Gx \leq h, \end{aligned}$$

where $c \in \mathbb{R}^d, A \in \mathbb{R}^{m\times d}, b \in \mathbb{R}^m, G \in \mathbb{R}^{r\times d}, h \in \mathbb{R}^r$.

14.3.1 Warm-up Examples

Suppose we want to lower-bound our objective value for some constrained optimization problem.

$$\begin{aligned} \min_{x,y} \quad & x + 3y \\ \text{subject to} \quad & x + y \geq 2 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

$$\begin{aligned} & x + y \geq 2 \\ + \quad & 2y \geq 0 \\ = \quad & x + 3y \geq 2 \end{aligned}$$

Lower bound $B = 2$

Here's a slightly more general example:

$$\begin{aligned} \min_{x,y} \quad & px + qy \\ \text{subject to} \quad & x + y \geq 2 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

$$\begin{aligned} & ax + ay \geq 2 \\ & bx \geq 0 \\ & cy \geq 0 \\ \text{Now let } & a + b = p \\ & a + c = q \\ & a, b, c \geq 0 \end{aligned}$$

Lower bound $B = 2a$, for any a, b, c satisfying above

What's the best we can do? Maximize our lower bound over all possible a, b, c:

$$\begin{aligned} \min_{x,y} \quad & px + qy \\ \text{subject to} \quad & x + y \geq 2 \\ & x, y \geq 0 \end{aligned}$$

Called primal LP

$$\begin{aligned} \max_{a,b,c} \quad & 2a \\ \text{subject to} \quad & a + b = p \\ & a + c = q \\ & a, b, c \geq 0 \end{aligned}$$

Called dual LP

Notice that the number of dual variables is number of primal constraints. This is by construction.

14.3.2 Duality in LPs

Definition 14.1 (Dual of a Linear Program)**.** *Let's start with the punchline by writing down a primal LP and its dual LP. Given* $c \in \mathbb{R}^d$*,* $A \in \mathbb{R}^{m \times d}$*,* $b \in \mathbb{R}^m$*,* $G \in \mathbb{R}^{r \times d}$*,* $h \in \mathbb{R}^r$*:*

Primal LP	Dual LP
$\min_x \quad c^T x$ *subject to* $\quad Ax = b$ $Gx \leq h$	$\max_{u,v} \quad -b^T u - h^T v$ *subject to* $\quad -A^T u - G^T v = c$ $v \geq 0$

The idea of duality will seem a bit strange at first. We're going to develop a different optimization program (the dual) whose value lower bounds the value of this linear program (which will now be called the primal).

We notice that, for any vector $u \in \mathbb{R}^m, v \in \mathbb{R}^r, v \geq 0$, and for any x which is feasible for the primal, we have that,

$$u^T(Ax - b) + v^T(Gx - h) \leq 0.$$

This can be re-written as:

$$(-A^T u - G^T v)^T x \geq -u^T b - v^T h.$$

Consequently, if we set $-A^T u - G^T v = c$, then we obtain that,

$$c^T x \geq -u^T b - v^T h.$$

So every u, v which satisfies the constraints that $v \geq 0$ and $-A^T u - G^T v = c$ gives us a lower bound $u^T b + v^T h$ on our primal objective value. So we could imagine trying to find the largest possible lower bound, i.e. we could solve the program:

$$\begin{aligned} \max_{u,v} \quad & -u^T b - v^T h \\ \text{subject to} \quad & -A^T u - G^T v = c \\ & v \geq 0. \end{aligned}$$

This program is called the dual of our original linear program. Lets make some quick observations:

1. The dual is also a linear program. It is a maximization program (in contrast to the primal which was a minimization program).

2. Each constraint in the primal, yields a variable in the dual. Conversely, each variable in the primal will yield a constraint in the dual (and typically we'll additionally have some non-negativity constraints).

3. By construction, if we denote the primal optimal value by p^*, and the dual optimal value by d^* then it is the case that, $p^* \geq d^*$. This is known as *weak duality.* It will turn out that under some additional conditions (say if the primal and dual problems are feasible) it will be the case that these two values are in fact equal – this is known as *strong duality* and we will revisit this later.

4. A useful exercise, is to rewrite the dual as a minimization LP, and then take its dual (can be done mechanically). What you will observe is that you will end up back at the primal (up to eliminating some variables, and switching signs again). Concisely, the dual of the dual LP is the primal LP. This fact also turns out to be true in more generality.

5. We will say that $p^* = \infty$ if the primal is infeasible (i.e. no x satisfies the constraints), and that $d^* = -\infty$ if the dual is infeasible. We will say that the primal is *unbounded* if $p^* = -\infty$ and the dual is unbounded if $d^* = \infty$.

   Weak duality then tells us the following facts: if the dual is unbounded, then the primal must be infeasible. Similarly, if the primal is unbounded then the dual must be infeasible.