A Crash Course on Linear Programs : Part 2[1]

- ***The Dual Linear Program.*** For every linear program there is another linear program which lives in a completely different space but has the same value! In approximation algorithms, the dual is often used to *design and analyze* "self-contained" algorithms for problems. By this, I mean algorithms which do not resort to solving LPs. In this note we brush up on the definitions.

- We begin with minimization programs on n variable. For convenience's sake, we will differentiate constraints as "non-trivial" inequalities and "non-negativity" constraints.

$$\begin{aligned} \mathsf{lp} := \text{minimize} \quad & \mathbf{c}^\top \mathbf{x} = \sum_{j=1}^{n} c_j x_j \\ & A\mathbf{x} \geq \mathbf{b}, \qquad A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m \\ & \mathbf{x} \in \mathbb{R}^n_{\geq 0} \end{aligned} \tag{Linear Program}$$

- ***The Lagrangean.*** The dual, which is not restricted to just linear programs but *any* program, starts with what is called the Lagrangean function named after the Italian-French mathematician Joseph-Louis Lagrange (aka Giuseppe Luis Lagrangia). The main idea of this is to "move all the constraints to the objective". Instead of moving all, we move the non-trivial ones. Let us introduce variables (called Lagrange/dual variables) $\mathbf{y}_i$ for each of the m constraints/rows of the matrix A. Given this m-dimensional variable vector $\mathbf{y}$, define

$$\mathcal{L}(\mathbf{y}) := \min_{\mathbf{x} \in \mathbb{R}^n_{\geq 0}} \left(\mathbf{c}^\top \mathbf{x} + \underbrace{\mathbf{y}^\top (\mathbf{b} - A\mathbf{x})}_{= \sum_{i=1}^{m} \mathbf{y}_i \cdot (\mathbf{b}_i - \mathbf{a}_i^\top \mathbf{x})} \right) \tag{Lagrangean}$$

One way to think about the above function is the following. For the time being assume $\mathbf{y}_i \geq 0$ and think of it as a rate at which we "penalize" $\mathbf{x}$ if it $\mathbf{x}$ doesn't satisfy the ith inequality, that is, $\mathbf{b}_i > \mathbf{a}_i^\top \mathbf{x}$. In that case, we multiply this "violation" by $\mathbf{y}_i$ and add it to the function. Since $\mathbf{x}$ is trying to "minimize" the term in the paranthesis, the $\mathbf{y}$'s perhaps nudge the $\mathbf{x}$ to becoming more feasible. The last line is really figurative and shouldn't be given much attention.

However, a few facts are to be observed.

Fact 1. Suppose $\mathbf{x}$ be any feasible solution to (Linear Program). Then, for *any* $\mathbf{y} \in \mathbb{R}^m_{\geq 0}$, we have $\mathcal{L}(\mathbf{y}) \leq \mathbf{c}^\top \mathbf{x}$. In particular, this is true if we take the optimal solution $\mathbf{x}^*$, and if we take the $\mathbf{y}$ which maximizes $\mathcal{L}(\mathbf{y})$. Therefore,

$$\max_{\mathbf{y} \in \mathbb{R}^m_{\geq 0}} \mathcal{L}(\mathbf{y}) \leq \mathsf{lp} \tag{1}$$

[1] *Lecture notes by Deeparnab Chakrabarty. Last modified : 18th Mar, 2022*
These have not gone through scrutiny and may contain errors. If you find any, or have any other comments, please email me at deeparnab@dartmouth.edu. Highly appreciated!

Proof. Because for a feasible $\mathbf{x}$ for (Linear Program), we have $(\mathbf{b} - A\mathbf{x}) \leq \mathbf{0}$ and thus $\mathbf{y}^\top(\mathbf{b} - A\mathbf{x}) \leq 0$ if $\mathbf{y} \geq \mathbf{0}$. Which in turn means $\mathcal{L}(\mathbf{y}) \leq \mathbf{c}^\top \mathbf{x}$. □

Fact 2. One can re-write (Lagrangean) as

$$\mathcal{L}(\mathbf{y}) = \begin{cases} \mathbf{y}^\top \mathbf{b} & \text{if } \mathbf{y}^\top A \leq \mathbf{c}^\top \\ -\infty & \text{otherwise} \end{cases}$$

Proof. Rearranging gives us $\mathcal{L}(\mathbf{y}) = \mathbf{y}^\top \mathbf{b} + \min_{\mathbf{x} \geq 0} \left(\mathbf{c}^\top - \mathbf{y}^\top A\right) \mathbf{x}$. If $(\mathbf{c}^\top - \mathbf{y}^\top A)$ has any coordinate i negative, then one would choose $\mathbf{x}_i$ to be as large a positive number and $\mathbf{x}_j = 0$ for all other coordinates to make the minimum be as negative a number as one wants. □

- ***The Dual LP and Weak Duality.*** The above two facts imply the following: one, that the maximization of $\mathcal{L}(\mathbf{y})$ can be written as a linear program itself, and two, the value of this linear program is a *lower* bound on the LP value. This linear program is called the *Dual LP*.

$$\begin{aligned} \mathsf{dual} := \text{maximize} \quad & \mathbf{b}^\top \mathbf{y} = \sum_{i=1}^{m} b_i y_i & \text{(Dual Program)} \\ & A^\top \mathbf{y} \leq \mathbf{c}, \\ & \mathbf{y} \in \mathbb{R}^m_{\geq 0} \end{aligned}$$

$$\mathsf{dual} \leq \mathsf{lp} \qquad \text{(Weak Duality)}$$

A couple of remarks about the dual. One, the dual is a *maximization* LP while the original LP, which is called the *primal* LP, was a minimization one. Therefore the dual value of *any feasible* dual solution is a lower bound on the value of the primal LP; this is a very important fact that will be used in algorithm design and analysis. Second, for every variable x_j in the primal there is a constraint in the dual, and for every constraint in the primal there is a variable y_i in the dual. Writing the dual LP is a completely *mechanical* process, but experience tells me it takes some time getting used to; the inexperienced reader is urged to look at the following illustrations and then try taking duals of every LP they see (in particular, take dual of the dual).

- ***Two Illustrations.*** Consider the following LP on $n = 3$ variables having $m = 2$ constraints apart from the non-negativity constraints.

$$\begin{aligned} \mathsf{lp} := \text{minimize} \quad & 2x_1 + 3x_2 - x_3 & \text{(Illus-primal)} \\ & x_1 + x_2 - x_3 \geq 3, & \text{(P1)} \\ & x_3 - 2x_1 \geq 0, & \text{(P2)} \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Before reading further, can you see any simple *lower bound* on lp? To me, I can see that the LP objective is at least $x_1 + x_2 - x_3$ which is at least 3 by (P1). Therefore, surely $\mathsf{lp} \geq 3$. Anything larger doesn't immediately leap to the eye (the adverb "immediately" is important). Ok, let's take the dual now.

In the dual LP, we have *two* variables, let's call them y_1 and y_2 corresponding to primal constraints (P1) and (P2). The objective of the dual LP is to maximize a linear combination of y_1 and y_2, and the coefficients are simply the RHS of the corresponding primal constraints. Thus, it is $3y_1 + 0y_2 = 3y_1$.

There is a dual constraint for each primal variable; therefore, there will be three constraints. Let me show how to figure out the *dual* constraint on (y_1, y_2) corresponding to primal variable x_1. We first figure out which primal constraints x_1 appears in; the corresponding dual variables will appear in the dual constraint. In this case, x_1 appears in both, and so both y_1 and y_2 will appear. Furthermore, the coefficient of y_1 will be the coefficient of x_1 in (P1), and similarly the coefficient of y_2 will be coefficient of x_1 in (P2). This forms the LHS of the dual constraint, which in this case is $y_1 - 2y_2$. The inequality of the constraint is $\leq$, and the RHS of the constraint is the coeffecient of x_1 in the primal objective. And therefore, the dual constraint is $y_1 - 2y_2 \leq 1$. We can similarly write the dual constraints corresponding to x_2 and x_3 (do you want to try before reading ahead?) Finally, we add non-negativity constraints on y_1 and y_2, and this finishes the dual. Pretty mechanical.

$$\begin{aligned}
\mathsf{dual} := \text{maximize} \quad & 3y_1 && \text{(Illus-Dual)} \\
& y_1 - 2y_2 \leq 2, && \text{(D1)} \\
& y_1 \leq 3, && \text{(D2)} \\
& y_2 - y_1 \leq -1 && \text{(D3)} \\
& y_1, y_2 \geq 0 &&
\end{aligned}$$

First note that $y_1 = 3$ and $y_2 = 2$ is a feasible solution with $\mathsf{dual} = 9$. And therefore, by (1), we also get that the optimum value lp of (Illus-primal) is at least 9 as well. Indeed, it is precisely 9 since $(x_1 = 0, x_2 = 3, x_3 = 0)$ achieves that value; but this was not immediate before the dual, right? Good. Let's move to a second and more abstract illustration.

Consider the LP relaxation for the vertex cover problem that we have seen before. Here it is.

$$\begin{aligned}
\mathsf{lp}(G) := \text{minimize} \quad & \sum_{v \in V} c(v) x_v && && \text{(Vertex Cover LP)} \\
& x_u + x_v \geq 1, && \forall (u,v) \in E && (2) \\
& x_v \geq 0, && \forall v \in V &&
\end{aligned}$$

What is the dual of the above LP? Do you want to try writing it before peeking ahead? The dual has a variable y_e per edge of the graph (since the primal has a constraint per edge). The constraint is simply the sum of these y_e's since the RHS of the primal constraints is 1. There is a dual constraint per vertex $v \in V$ since there is a primal variable for every $v \in V$. The constraint corresponding to v is a linear combination of all dual variables y_e such that the primal variable x_v appears in the eth constraint. In particular, it is the sum of all the y_e's for e incident on v. The RHS of the constraint is $c(v)$ since that is the coefficient of x_v int he primal LP. And finally, we have non-negativity constraints on y_e's. Done.

$$\begin{aligned}
\mathsf{dual}(G) := \text{maximize} \quad & \sum_{e \in E} y_e && && \text{(Vertex Cover Dual)} \\
& \sum_{e : v \in e} y_e \leq c(v), && \forall v \in V && (3) \\
& y_e \geq 0, && \forall e \in E &&
\end{aligned}$$

- ***Strong Duality.*** Here is one of the most magical theorems out there.

> **Theorem 1** (Strong Duality)**.** $\mathsf{dual} = \mathsf{lp}$

Proof. (Sketch) We provide a proof to give an idea of how such a theorem is proven. Indeed, we consider the special case of *non-degenerate* feasible regions. That is, the feasible region is full dimensional and every basic feasible solution $\mathbf{x}$ has exactly n constraints holding with equality, and the rest hold with strict inequality. This assumption is not needed, but it gets to the essence of the proof.

Consider an optimal bfs $\mathbf{x}^*$ (recall, such a solution always exists) and let B be the corresponding basis. So, $B\mathbf{x}^* = \mathbf{b}_B$, that is $\mathbf{a}_i^\top \mathbf{x}^* = \mathbf{b}_i$ for $i \in B$ (we abuse B to denote rows and the index of the rows), and the rows of B span $\mathbb{R}^n$. In particular, the cost vector $\mathbf{c}$ can be uniquely written as a linear combination of the basis vectors; $\mathbf{c} = \sum_{i \in B} y_i \mathbf{a}_i$.

Now consider a candidate solution $\mathbf{y}$ to (Dual Program with equalities) where $\mathbf{y}_i = y_i$ for $i \in B$ and $\mathbf{y}_j = 0$ for $j \notin B$. Observe (a) by definition $\mathbf{y}^\top A = \mathbf{c}^\top$, and (b) $\mathbf{c}^\top \mathbf{x}^* = \sum_{i \in B} y_i \mathbf{a}_i^\top \mathbf{x}^* = \sum_{i \in B} y_i \mathbf{b}_i$. It seems as if we have found a feasible solution $\mathbf{y}$ to the dual LP whose objective equals $\mathbf{c}^\top \mathbf{x}^*$. Since we already have established weak-duality, this equality would prove theorem. The only nub is that we haven't establishes $\mathbf{y} \geq 0$; indeed, we have also not really used $\mathbf{x}^*$ is the *optimal solution*. We do so next.

We claim that all the $y_i \geq 0$ which would complete the proof of the theorem. Suppose not, and say $y_1 < 0$. Consider a vector $\mathbf{v} \in \mathbb{R}^n$ in the *null space* of $B \setminus \{1\}$ such that $\mathbf{a}_1^\top \mathbf{v} > 0$ and $\mathbf{a}_i^\top \mathbf{v} = 0$ for $i \in B \setminus \{1\}$. This exists since $\mathbf{a}_1$ is linearly independent of $B \setminus \mathbf{a}_1$. Now choose $\theta > 0$ small enough such that $\mathbf{a}_j^\top(\theta \mathbf{v}) > \mathbf{b}_j$ for all $j \notin B$; this is where we are using the non-degeneracy assumption. By design, therefore, $\mathbf{x}' = \mathbf{x}^* + \theta \mathbf{v}$ is feasible. And, $\mathbf{c}^\top \mathbf{x}' - \mathbf{c}^\top \mathbf{x}^* = \theta \mathbf{c}^\top \mathbf{v}$. However,

$$\mathbf{c}^\top \mathbf{v} = y_1 \underbrace{\mathbf{a}_1^\top \mathbf{v}}_{>0} + \sum_{i=2}^{m} y_i \underbrace{\mathbf{a}_i^\top \mathbf{v}}_{=0} \quad < 0$$

since $y_1 < 0$. This contradicts $\mathbf{x}^*$ is the optimum solution, completing the proof of strong duality. □

- ***Complementary Slackness.*** A very interesting feature about the mirroring is captured by the following observation which, due to its importance, is given a name called *complementary slackness*. It says, a dual variable is *positive* in an optimal dual solution only if the corresponding *primal constraint* must be tight, that is hold with equality, in any optimal primal solution. Similarly, a primal variable is *positive* in an optimal primal solution only if the corresponding *dual constraint* is tight.

> **Lemma 1** (Complementary Slackness.)**.** Let $\mathbf{x}^*$ be *any* optimal solution of (Linear Program). Let $\mathbf{y}^*$ be any optimal solution of (Dual Program with equalities). Then, $\mathbf{y}_j^* > 0 \Rightarrow \mathbf{a}_j^\top \mathbf{x} = \mathbf{b}_j$ and $\mathbf{x}_i^* > 0 \Rightarrow \mathbf{y}^\top \mathbf{A}_i = \mathbf{c}_i$. Her $\mathbf{A}_i$ is the ith column of the matrix A.

Proof. For brevity's sake, let's call $\mathbf{x}^*$ simply $\mathbf{x}$ and $\mathbf{y}^*$ simply $\mathbf{y}$. By Strong Duality, we know that $\mathbf{c}^\top \mathbf{x} = \mathbf{y}^\top \mathbf{b}$, since $(\mathbf{x}, \mathbf{y})$ are optimal solutions. We also know that $\mathbf{c}^\top \geq \mathbf{y}^\top A$. Therefore, since $\mathbf{x} \geq 0$, we get $\mathbf{c}^\top \mathbf{x} \geq \left(\mathbf{y}^\top A\right) \mathbf{x}$. And so,

$$\mathbf{y}^\top \mathbf{b} = \mathbf{c}^\top \mathbf{x} \geq \left(\mathbf{y}^\top A\right) \mathbf{x} \;\; \Rightarrow \;\; \mathbf{y}^\top \mathbf{b} \geq \mathbf{y}^\top (A\mathbf{x}) \;\; \Rightarrow \;\; \mathbf{y}^\top (A\mathbf{x} - \mathbf{b}) \leq 0$$

On the other hand $A\mathbf{x} \geq \mathbf{b}$, or in other words if we define the m-dimensional vector $\mathbf{v} := A\mathbf{x} - \mathbf{b}$, $\mathbf{v}_j \geq 0$ for all $1 \leq j \leq m$. Thus, we get $\sum_{j=1}^{m} \mathbf{y}_j \mathbf{v}_j \leq 0$ while $\mathbf{y}_j \geq 0$ and $\mathbf{v}_j \geq 0$.

There is only *one* possibility : we must have $\sum_{j=1}^{m} \mathbf{y}_j \mathbf{v}_j = 0$. And therefore, whenever $\mathbf{y}_j > 0$ we *must* have $\mathbf{v}_j = 0$, that is, $\mathbf{a}^\top j = \mathbf{b}_j$.

Since $\mathbf{y}^\top (A\mathbf{x} - \mathbf{b}) = 0$, we also get that $\mathbf{c}^\top \mathbf{x} = \left(\mathbf{y}^\top A\right) \mathbf{x}$. That is, $\left(c^\top - \mathbf{y}^\top A\right) \mathbf{x} = 0$. Again, if we define the n-dimensional vector $\mathbf{w} := \mathbf{c} - A^\top \mathbf{y}$, then we get $\mathbf{w}^\top \mathbf{x} = 0$ while both $\mathbf{w}$ and $\mathbf{x}$ are non-negative. This would mean that $\mathbf{x}_i > 0 \Rightarrow \mathbf{w}_i = 0$, that is, $\mathbf{y}^\top \mathbf{A}_i = \mathbf{c}_i$. □

- ***The Dual of a Maximization LP.*** The same procedure using the Lagrangean function can be used to write the dual of a maximization LP as well. So, if the primal LP is

$$\begin{aligned} \mathsf{lp} := \text{maximize} \quad & \mathbf{c}^\top \mathbf{x} = \sum_{j=1}^{n} c_j x_j \\ & A\mathbf{x} \leq \mathbf{b}, \qquad A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m \\ & \mathbf{x} \in \mathbb{R}^n_{\geq 0} \end{aligned} \tag{Max Linear Program}$$

  Then the dual LP also has variables $\mathbf{y} \in \mathbb{R}^m$ corresponding to the constraints in the primal. It is a *minimization* LP, and the constraints are of the "$\geq$" type. Weak duality asserts that the value of the dual is *at least* the value of the maximizing primal, and strong duality implies they are equal.

$$\begin{aligned} \mathsf{dual} := \text{minimize} \quad & \mathbf{b}^\top \mathbf{y} = \sum_{i=1}^{m} b_i y_i \\ & A^\top \mathbf{y} \geq \mathbf{c}, \\ & \mathbf{y} \in \mathbb{R}^m_{\geq 0} \end{aligned} \tag{Min Dual Program}$$

- ***The Dual with Equality Constraints.*** Sometimes the primal LP has equality constraints. In that case, the corresponding dual variables are "free"; that is, they don't have any non-negativity constraint and are allowed to be free. Once again, this is not hard to see if one treats the equality constraint as two sets of *inequality* constraints, and then writes the dual. In particular, if the primal LP is

$$\begin{aligned} \mathsf{lp} := \text{minimize} \quad & \mathbf{c}^\top \mathbf{x} = \sum_{j=1}^{n} c_j x_j \\ & A\mathbf{x} \geq \mathbf{b}, \qquad A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m \\ & P\mathbf{x} = \mathbf{q}, \qquad P \in \mathbb{R}^{k \times n}, \mathbf{q} \in \mathbb{R}^k \\ & \mathbf{x} \in \mathbb{R}^n_{\geq 0} \end{aligned} \tag{Linear Program with Equalities}$$

  then its dual has two sets of variables $\mathbf{y} \in \mathbb{R}^m$ corresponding to A and $\mathbf{z} \in \mathbb{R}^k$ corresponding to P. The program is

$$\begin{aligned} \mathsf{dual} := \text{maximize} \quad & \mathbf{b}^\top \mathbf{y} + \mathbf{q}^\top \mathbf{z} \\ & A^\top \mathbf{y} + P^\top \mathbf{z} \leq \mathbf{c}, \\ & \mathbf{y} \in \mathbb{R}^m_{\geq 0}, \mathbf{z} \in \mathbb{R}^k \end{aligned} \tag{Dual Program with equalities}$$

  Note that $\mathbf{z}$ has no non-negativity constraints.

Notes

Since this is not a course on linear programming, my notes will be short because the alternative is to be extremely long. All I will say is that everyone who studies linear programming has a favorite source which enlightened them. For me it was this beautiful text [1] by Bertsimas and Tsitsiklis.

References

[1] D. Bertsimas and J. Tsitsiklis. *Introduction to Linear Optimization*. Athena-Scientific, 1997.