week04-01–duality

March 13, 2024

1 George McNinch Math 87 - Spring 2024

2 Week 4

3 Duality and linear programming

3.1 Duality

Consider a linear program $\mathcal{L}$ – in *standard form* – determined by data $(\mathbf{c} \in \mathbb{R}^{1 \times n}, A \in \mathbb{R}^{r \times n}, \mathbf{b} \in \mathbb{R}^r)$ that seeks to `maximize` the value $\mathbf{c} \cdot \mathbf{x}$ of the *objective* function determined by $\mathbf{c} \in \mathbb{R}^{1 \times n}$. The linear program considers non-negative vectors $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \geq \mathbf{0}$, subject to inequality constraints $A\mathbf{x} \leq \mathbf{b}$ determined by $A \in \mathbb{R}^{r \times n}$ and $\mathbf{b} \in \mathbb{R}^r$, and maximizes the objective function.

We are going to associate to $\mathcal{L}$ a *new* linear program $\mathcal{L}'$. The linear program $\mathcal{L}'$ we are going to describe is called the *dual* linear program to $\mathcal{L}$; in this context, $\mathcal{L}$ is referred to as the *primal* linear program.

We first just give the rule describing $\mathcal{L}'$ before trying to motivate the construction.

3.2 Definition of $\mathcal{L}'$

The linear program $\mathcal{L}'$ dual to $\mathcal{L}$ is determined by the data $(\mathbf{b}^T, \mathbf{A}^T, \mathbf{c}^T)$ and it seeks to `minimize` the value $\mathbf{b}^T \cdot \mathbf{y}$; note that $\mathbf{b}^T \in \mathbb{R}^{1 \times r}$ – so in this case the variable is a non-negative vector $\mathbf{y} \in \mathbb{R}^r$, $\mathbf{y} \geq \mathbf{0}$ subject to the inequality constraint $A^T\mathbf{y} \geq \mathbf{c}^T$ (note the direction of the inequality!).

Speaking roughly, in the dual $\mathcal{L}'$, each of $\mathbf{c}$, A, and $\mathbf{b}$ have been replaced by their transpose, $\mathbf{c}$ and $\mathbf{b}$ have "swapped roles", and the inequality direction in the constraint has been reversed.

3.3 A first example

So, we are perhaps a bit skeptical that this construction can be useful. Let's recall our first example of a linear program $\mathcal{L}$ – the carpenter constructing tables and bookshelves. Recall that the goal in that case was to maximize

$$\mathbf{c} \cdot \begin{bmatrix} t \\ b \end{bmatrix} \quad \text{such that} \quad A \cdot \begin{bmatrix} t \\ b \end{bmatrix} \leq \mathbf{r}$$

where t and b respectively denote the number of tables and bookshelves constructed by the carpenter – so that $\begin{bmatrix} t \\ b \end{bmatrix} \geq \mathbf{0}$ - and where

$$\mathbf{c} = \begin{bmatrix} 25 & 30 \end{bmatrix}, \qquad A = \begin{bmatrix} 5 & 4 \\ 20 & 30 \end{bmatrix}, \qquad \mathbf{r} = \begin{bmatrix} 120 \\ 690 \end{bmatrix}.$$

Thus $\mathbf{c} \cdot \begin{bmatrix} t \\ b \end{bmatrix}$ denotes the *profit*, and the condition $A \cdot \begin{bmatrix} t \\ b \end{bmatrix} \leq \mathbf{r}$ reflects resource usage constraints (labor and lumber).

Thus we are trying to maximize the profit, keeping in mind that lumber and labor resources are limited.

Now let's describe the dual linear program $\mathcal{L}'$. We introduce "dual variables" corresponding to the entries of $\mathbf{r}$ – we'll call them y_L and y_W – which we view as *unit prices* of the indicated resources. In particular, $\mathbf{y} = \begin{bmatrix} y_L \\ y_W \end{bmatrix} \geq \mathbf{0}$. Thus the dual objective function

$$\mathbf{r}^T \cdot \mathbf{y} = 120 y_L + 690 y_W$$

repesents the cost of our labor and wood, and $\mathcal{L}'$ seeks to minimize this cost, subject to the inequality constraint given by

$$A^T \mathbf{y} \geq \mathbf{c}^T \iff \begin{bmatrix} 5 & 20 \\ 4 & 30 \end{bmatrix} \begin{bmatrix} y_L \\ y_W \end{bmatrix} \geq \begin{bmatrix} 25 \\ 30 \end{bmatrix} \iff \begin{cases} 5 y_L + 20 y_W \geq 25 \\ 4 y_L + 30 Y_W \geq 30 \end{cases}$$

Thus, we seek the unit prices of the resources which minimize the usage of labor and wood, keeping in mind the amount of profit we'd get from using the labor and wood on making a table (5 hours and 20 feet for 25 dollars) or a bookshelf (4 hours and 30 feet for 30 dollars).

There remains of course the question:

Why do we want to consider the dual linear program?

Probably the best answer is that the solutions to dual linear programs give you a way to reason about what happens when the constraints (the vector $\mathbf{b}$) vary.

As a first question, we'd like to know: to what extent are the solutions to $\mathcal{L}$ and the solutions to $\mathcal{L}'$ related?? We are going to give some partial answers to this question.

3.4 Dual of the dual?

Recall that $\mathcal{L}$ is determined by $(\mathbf{c}, A, \mathbf{b})$, and $\mathcal{L}'$ is determined by $(\mathbf{b}^T, \mathbf{A}^T, \mathbf{c}^T)$.

We'd like to understand the dual of $\mathcal{L}'$. Of course, strictly speaking we haven't even *defined* the dual of $\mathcal{L}'$, because $\mathcal{L}'$ is a `minimize`-ing linear program and speaking strictly we've only defined the dual for `maximize`-ing linear programs (in standard form)!

However, the task of finding an optimal solution to the linear program $\mathcal{L}'$ – a `minimize`-ing linear program – is equivalent to the task of finding an optimal solution to the linear program determined by $(-\mathbf{b}^T, -\mathbf{A}^T, -\mathbf{c}^T)$ which seeks to `maximize` its objective function subject to the constraint $-\mathbf{A}^T \mathbf{y} \leq -\mathbf{c}^T$.

Since $(B^T)^T = B$ for any matrix B, if we view $\mathcal{L}'$ as the primal linear program (with the preceding caveat for how to view it as a `maximize`-ing linear program) then its dual $(\mathcal{L}')'$ is determined by $(\mathbf{c}, A, \mathbf{b})$ and it is then clear that the dual $(\mathcal{L}')'$ identifies with $\mathcal{L}$ (subject to the same caveat!).

Succintly: the dual of the dual is the primal linear program.

3.5 Some theorems relating solutions of the primal and the dual

By a *feasible point* of a linear system, we just mean a point in the feasible region of the linear system. In general, the constraints of a linear program might be *inconsistent* and thus have no feasible points.

Suppose that the feasible region is non-empty. If this feasible region is *bounded*, there is always an optimal solution. (If the region is unbounded in the direction of the gradient of the objective function, there will be no maximal solution).

Consider a primal linear program $\mathcal{L}$ determined as before by the data $(\mathbf{c}, A, \mathbf{b})$ and let $\mathcal{L}'$ be the dual linear system.

We are going to formulate three theorems.

3.6 Weak Duality:

The first result says that on feasible points, the objective function of $\mathcal{L}$ is always bounded by that of $\mathcal{L}'$:

> **Theorem** (Weak duality): Let $\mathbf{x} \in \mathbb{R}^n$ be a feasible point for $\mathcal{L}$ and $>$ let $\mathbf{y} \in \mathbb{R}^r$ be a feasible point for $\mathcal{L}'$. Then
> $$\mathbf{c}\mathbf{x} \leq \mathbf{y}^T \cdot A \cdot \mathbf{x} \leq \mathbf{b}^T \mathbf{y}.$$
> In particular,
> $$\max_{\mathbf{x}}(\mathbf{c}\mathbf{x}) \leq \min_{\mathbf{y}}(\mathbf{b}^T \mathbf{y}).$$

Proof: Note that $A \cdot \mathbf{x} \leq \mathbf{b}$ since $\mathbf{x}$ is a feasible point for $\mathcal{L}$. Since the vector $\mathbf{y}$ is non-negative, multiplication with $\mathbf{y}$ doesn't change inequalities, and it follows that

$$\mathbf{y}^T \cdot A \cdot \mathbf{x} \leq \mathbf{y}^T \cdot \mathbf{b}.$$

Similarly, since $\mathbf{y}$ is feasible for $\mathcal{L}'$, we have $A^T \cdot \mathbf{y} \geq \mathbf{c}^T$. Since $\mathbf{x}$ is non-negative, find that

$$\mathbf{x}^T \cdot A^T \cdot \mathbf{y} \geq \mathbf{x}^T \cdot \mathbf{c}^\mathbf{T}$$

Taking the transpose preserves the inequality (and reverse the order of matrix multiplication!!); thus

$$\mathbf{y}^T \cdot A \cdot \mathbf{x} \geq \mathbf{c} \cdot \mathbf{x}.$$

Combining inequalities, we now see that

$$\mathbf{c} \cdot \mathbf{x} \leq \mathbf{y}^T \cdot A \cdot \mathbf{x} \leq \mathbf{y}^T \cdot \mathbf{b} = \mathbf{b}^T \cdot \mathbf{y}.$$

QED

3.7 Strong duality

The second theorem relates the values of the objective functions for $\mathcal{L}$ and $\mathcal{L}'$ at optimal solutions.

Theorem (Strong duality): If $\mathbf{x}^*$ is an optimal solution for $\mathcal{L}$ and if $\mathbf{y}^*$ is an optimal solution for $\mathcal{L}'$, then

$$\mathbf{c}\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*.$$

Conversely, if $\mathbf{x}^*$ is a feasible point for $\mathcal{L}$, if $\mathbf{y}^*$ is a feasible point for $\mathcal{L}'$ and if $\mathbf{c} \cdot \mathbf{x}^* = \mathbf{b}^T \cdot \mathbf{y}^*$ then $\mathbf{x}^*$ is an optimal solution for $\mathcal{L}$ and $\mathbf{y}^*$ is an optimal solution for $\mathcal{L}'$.

The proof is omitted because it involves a bit more linear algebra than we want to wade into for the moment.

3.8   Complementary slackness

Theorem (Complementary Slackness): Let $\mathbf{x}$ be a feasible point for $\mathcal{L}$ and $\mathbf{y}$ a feasible point for $\mathcal{L}'$. Then, $\mathbf{x}$ is optimal for $\mathcal{L}$ and $\mathbf{y}$ is optimal for $\mathcal{L}'$ if and only if

$$(\clubsuit) \quad (\mathbf{b} - A\mathbf{x})^T \cdot \mathbf{y} = 0 \quad \text{and} \quad (\mathbf{y}^T A - \mathbf{c}) \cdot \mathbf{x} = 0.$$

Remark: The Theorem can be formulated in terms of the so-called *slack variables* z_i $1 \leq i \leq r$ which may be defined by the equation

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_r \end{bmatrix} = \mathbf{z} = \mathbf{b} - A \cdot \mathbf{x}$$

(There the dual slack variables are defined in the same way). We'll revisit the slack variables in our discussion of *dual prices* in a subsequent notebook.

Sketch of proof (of complementary slackness): Since $\mathbf{x}$ and $\mathbf{y}$ are feasible, the weak duality Theorem tells us that

$$(\heartsuit) \quad \mathbf{c} \cdot \mathbf{x} \leq \mathbf{y}^T \cdot A\mathbf{x} \leq \mathbf{b}^T \cdot \mathbf{y}.$$

Now if $\mathbf{x}$ and $\mathbf{y}$ are both optimal, the strong duality Theorem tells us that

$$\mathbf{c}\mathbf{x} = \mathbf{b}^T\mathbf{y}.$$

In particular, equality holds everywhere in $(\heartsuit)$ and we find that

$$\mathbf{c} \cdot \mathbf{x} = \mathbf{y}^T \cdot A \cdot \mathbf{x} = \mathbf{b}^T \cdot \mathbf{y}.$$

It follows that

$$\mathbf{y}^T \cdot A \cdot \mathbf{x} - \mathbf{c} \cdot \mathbf{x} = 0 \implies (\mathbf{y}^T \cdot A - \mathbf{c}) \cdot \mathbf{x} = 0.$$

Similarly,

$$0 = \mathbf{y}^\mathbf{T} \cdot A \cdot \mathbf{x} - \mathbf{b}^T \cdot \mathbf{y} = \mathbf{x}^T \cdot A^T \cdot \mathbf{y} - \mathbf{b}^T \cdot \mathbf{y} = (\mathbf{x}^T A^T - \mathbf{b}^T) \cdot \mathbf{y} = (A\mathbf{x} - \mathbf{b})^T \cdot \mathbf{y}.$$

This confirms that $(\clubsuit)$ holds.

Conversely, given $(\clubsuit)$ the preceding calculations show that equality holds in $(\heartsuit)$, and the result follows by the strong duality theorem. **QED**

These results and ideas allow us to "play the primal or dual linear program off one another" to try to find an optimal solution.

In the next notebook, we'll see some useful application!!