15-451/651: Design & Analysis of Algorithms November 3, 2022
Lecture #16: Linear Programming III last changed: November 3, 2022

In this lecture we discuss the general notion of Linear Programming *Duality*, a powerful tool that can allow us to solve some linear programs easier, gain theoretical insights into the properties of a linear program, and has many more applications that we might see later in the course. We will show how duality connects to some topics we have already seen, like minimax optimal strategies in zero-sum games.

Objectives of this lecture

In this lecture, we will

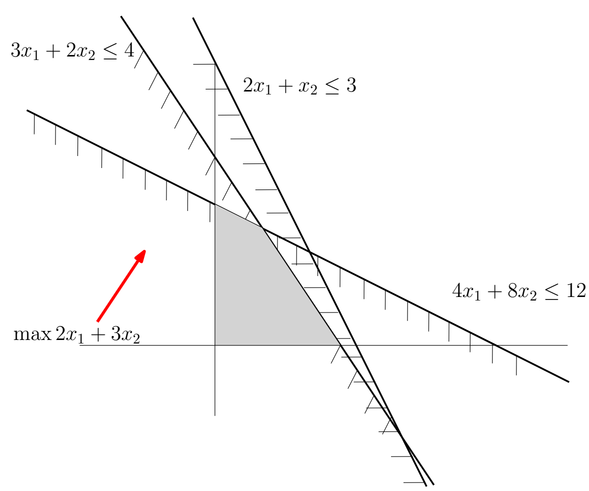
- Motivate and define the idea of the dual of a linear program
- See a general method for converting any linear program into its dual program
- Learn some powerful theorems that tell us about the behavior of a linear program and its dual
- See how duality can teach us about minimax optimal strategies for zero-sum games

1 The Dual Program as an Upper Bound

Consider the following LP which is written in standard form.

$$\begin{aligned} \text{maximize } & 2x_1 + 3x_2 \\ \text{s.t.} \quad & 4x_1 + 8x_2 \leq 12 \\ & 2x_1 + x_2 \leq 3 \\ & 3x_1 + 2x_2 \leq 4 \\ & x_1, x_2 \geq 0 \end{aligned} \tag{1}$$

Here it is in a diagram which shows each constraint, the feasible region shaded in gray, and the objective direction as a red arrow.



Rather than try to solve the LP using an algorithm directly, we are going to do an experiment. Lets see if we can figure out some bounds on the objective value, and use those to hone in on the optimal value. How can we bound the objective? Well the only other information that we have are the constraints, so lets use them! Lets refer to the optimal objective value as OPT.

- First, since $x_1, x_2 \geq 0$, we can notice that the left-hand equation of the first constraint $(4x_1 + 8x_2)$ must be bigger than the objective $(2x_1 + 3x_2)$, in other words, we can write

$$\underbrace{2x_1 + 3x_2}_{\text{objective function}} \leq \underbrace{4x_1 + 8x_2 \leq 12}_{\text{first constraint}}$$

Note that the left-hand side is the objective function and the right hand side is the first constraint. So we therefore know that OPT ≤ 12. That's a start. Can we get a tighter bound?

- Yes, the first constraint is more than double the objective function, so we can write a tighter bound by using just half of the constraint

$$\underbrace{2x_1 + 3x_2}_{\text{objective function}} \leq \underbrace{\tfrac{1}{2}(4x_1 + 8x_2) \leq 6}_{\text{half of the first constraint}}$$

This gives us a bound of OPT ≤ 6. Can we do better? Maybe by combining multiple constraints!

- We want to combine some constraints such that we get as close to a coefficient of 2 for x_1 and a coefficient of 3 for x_2. By inspection, we can see that if we add the first and second constraint, we will have $6x_1 + 9x_2$, which is exactly three times our objective function, so lets try one third of that combination.

$$\underbrace{2x_1 + 3x_2}_{\text{objective function}} \leq \underbrace{\tfrac{1}{3}((4x_1 + 8x_2) + (2x_1 + x_2)) \leq 5}_{\text{one third of the first two constraints}}$$

Note that we get 5 on the right-hand side because we sum the right-hand sides of the original constraints to get $12 + 3 = 15$, then take one third of it. So we know that OPT ≤ 5.

In each of these cases we take a positive linear combination of the constraints, looking for better and better bounds on the maximum possible value of $2x_1 + 3x_2$. Why positive? Because if we multiply by a negative value, the sign of the inequality changes.

1.1 The tightest possible bound

After playing with this experiment for a bit, the natural question that arises is how do we find the tightest lower bound that can be achieved with this method, i.e., by writing down a a linear combination of the constraints? This is just another algorithmic problem, and we can systematically solve it, by letting y_1, y_2, y_3 be the (unknown) coefficients of our linear combination. So our goal is to write the combination like so

$$y_1(4x_1 + 8x_2) + y_2(2x_1 + x_2) + y_3(3x_1 + 2x_2) \leq 12y_1 + 3y_2 + 4y_3,$$

such that the value on the right-hand side is as small as possible. This sounds like just another linear program! To bound the original objective function, we require that the coefficients of x_1 add up to at least 2, and the coefficients of x_2 add up to at least 3. We can write these requirements down as a linear program.

$$\begin{aligned} \text{minimize} \quad & 12y_1 + 3y_2 + 4y_3 \\ \text{s.t.} \quad & 4y_1 + 2y_2 + 3y_3 \geq 2 \\ & 8y_1 + y_2 + 2y_3 \geq 3 \\ & y_1, y_2, y_3 \geq 0 \end{aligned} \tag{2}$$

This is indeed an LP! We refer to this LP (2) as the "**dual**" and the original LP (1) as the "**primal**". We designed the dual to serve as a method of constructing an upper bound on the optimal value of the primal, so if y is a feasible solution for the dual and x is a feasible solution for the primal, then $2x_1 + 3x_2 \leq 12y_1 + 3y_2 + 4y_3$.

This serves as an upper bound, but what happens if we make it tight? If we can find two feasible solutions $\mathbf{x}$ and $\mathbf{y}$, that make these equal, then we know we have found the provably optimal values of these LPs. In this case the feasible solutions $x_1 = \frac{1}{2}, x_2 = \frac{5}{4}$ and $y_1 = \frac{5}{16}, y_2 = 0, y_3 = \frac{1}{4}$ give us a value and matching upper bound of 4.75, which therefore must be the optimal value.

Exercise: The dual of the dual

The dual LP is a *minimization* LP, where the constraints are of the form $f_i(\mathbf{x}) \geq c_i$. You can try to give *lower* bounds on the optimal value of this LP by taking positive linear combinations of these constraints. E.g., argue that

$$12y_1 + 3y_2 + 4y_3 \geq 4y_1 + 2y_2 + 3y_2 \geq 2$$

(since $y_i \geq 0$ for all i) and

$$12y_1 + 3y_2 + 4y_3 \geq 8y_1 + y_2 + 2y_3 \geq 3$$

and

$$12y_1 + 3y_2 + 4y_3 \geq \frac{2}{3}(4y_1 + 2y_2 + 3y_2) + (8y_1 + y_2 + 2y_3) \geq \frac{4}{3} + 3 = 4\frac{1}{3}.$$

Formulate the problem of finding the best lower bound obtained by linear combinations of the given inequalities as an LP. Show that the resulting LP is the same as the primal LP (1), in other words, the dual of the dual gives you back the primal LP you started with.

Exercise: Another dual of the dual

Consider the "primal" LP below on the left:

$$\begin{aligned} \text{maximize} \quad & 7x_1 - x_2 + 5x_3 \\ \text{s.t.} \quad & x_1 + x_2 + 4x_3 \leq 8 \\ & 3x_1 - x_2 + 2x_3 \leq 3 \\ & 2x_1 + 5x_2 - x_3 \leq -7 \\ & x_1, x_2, x_3 \geq 0 \end{aligned} \qquad \begin{aligned} \text{minimize} \quad & 8y_1 + 3y_2 - 7y_3 \\ \text{s.t.} \quad & y_1 + 3y_2 + 2y_3 \geq 7 \\ & y_1 - y_2 + 5y_3 \geq -1 \\ & 4y_1 + 2y_2 - y_3 \geq 5 \\ & y_1, y_2, y_3 \geq 0 \end{aligned}$$

Show that the problem of finding the best upper bound obtained using linear combinations of the constraints can be written as the LP above on the right (the "dual" LP). Also, now formulate the problem of finding a lower bound for the dual LP. Show that you get the primal LP back again.

Exercise: The dual from not-standard form

In the examples above, the maximization LPs had constraints of the form $lhs_i \leq rhs_i$, and the rhs were all scalars, so taking positive linear combinations gave us *blah* $\leq$ *number*, i.e., an *upper* bound as we wanted. However, suppose the primal LP has some "nice" constraints $lhs_i \leq rhs_i$ and others are "not nice" $lhs_i \geq rhs_i$, e.g., like the left one below. Show that the dual has non-positive variables for the non-nice constraints. For example,

$$
\begin{aligned}
\text{maximize } & 7x_1 - x_2 + 5x_3 \\
\text{s.t. } & x_1 + x_2 + 4x_3 \leq 8 \\
& 3x_1 - x_2 + 2x_3 \geq 3 \\
& x_1, x_2, x_3 \geq 0
\end{aligned}
\qquad
\begin{aligned}
\text{minimize } & 8y_1 + 3y_2 \\
\text{s.t. } & y_1 + 3y_2 \geq 7 \\
& y_1 - y_2 \geq -1 \\
& 4y_1 + 2y_2 \geq 5 \\
& y_1 \geq 0, y_2 \leq 0
\end{aligned}
$$

Another way is to replace $lhs_i \geq rhs_i$ by the equivalent constraint $(-lhs_i) \leq (-rhs_i)$ and get an LP with only nice constraints. Show that the dual for this LP is equivalent to the dual for the original.

2 A General Formulation of the Dual

Consider the examples/exercises above. In all of them, we started off with a "primal" maximization LP:

$$
\begin{aligned}
\text{maximize } & \mathbf{c}^T \mathbf{x} \\
\text{subject to } & A\mathbf{x} \leq \mathbf{b} \\
& \mathbf{x} \geq \mathbf{0},
\end{aligned}
\tag{3}
$$

The constraint $\mathbf{x} \geq \mathbf{0}$ is just short-hand for saying that the $\mathbf{x}$ variables are constrained to be non-negative.[1] And to get the best upper bound we generated a "dual" minimization LP:

$$
\begin{aligned}
\text{minimize } & \mathbf{r}^T \mathbf{y} \\
\text{subject to } & P\mathbf{y} \geq \mathbf{q} \\
& \mathbf{y} \geq \mathbf{0},
\end{aligned}
\tag{4}
$$

The important thing is: this matrix P, and vectors $\mathbf{q}, \mathbf{r}$ are not just any vectors. Look carefully: $P = A^T$. $\mathbf{q} = \mathbf{c}$ and $\mathbf{r} = \mathbf{b}$. The dual is in fact:

Claim: The dual of a linear program

The dual of the standard form LP (4) is

$$
\begin{aligned}
\text{minimize } & \mathbf{y}^T \mathbf{b} \\
\text{subject to } & \mathbf{y}^T A \geq \mathbf{c}^T \\
& \mathbf{y} \geq \mathbf{0},
\end{aligned}
\tag{5}
$$

And if you take the dual of (5) to try to get the best lower bound on this LP, you'll get (4). *The dual of the dual is the primal.* The dual and the primal are best upper/lower bounds you can obtain as linear combinations of the inputs.

The natural question is: maybe we can obtain better bounds if we combine the inequalities in more complicated ways, not just using linear combinations. Or do we obtain optimal bounds using just linear combinations? In fact, we get optimal bounds using just linear combinations, as the next theorems show.

[1] We use the convention that vectors like $\mathbf{c}$ and $\mathbf{x}$ are column vectors. So $\mathbf{c}^T$ is a row vector, and thus $\mathbf{c}^T\mathbf{x}$ is the same as the inner product $\mathbf{c} \cdot \mathbf{x} = \sum_i c_i x_i$. We often use $\mathbf{c}^T\mathbf{x}$ and $\mathbf{c} \cdot \mathbf{x}$ interchangeably. Also, $\mathbf{a} \leq \mathbf{b}$ means component-wise inequality, i.e., $a_i \leq b_i$ for all i.

2.1 The Theorems

When we derived the dual linear program, we used it as a means to provide upper bounds on the primal LP. We can formally prove that it indeed does just that. This fact is called *weak duality*.

Theorem 1: Weak Duality

If $\mathbf{x}$ is a feasible solution to the primal LP (4) and $\mathbf{y}$ is a feasible solution to the dual LP (5) then

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

Proof. This follows by applying the constraints of the primal and dual LPs in (4) and (5) and the fact that $\mathbf{x} \geq 0$ and $\mathbf{y} \geq 0$. Since $\mathbf{y}^T A \geq \mathbf{c}^T$, we can plug this into the objective $\mathbf{c}^T\mathbf{x}$ and get

$$\mathbf{c}^T\mathbf{x} \leq (\mathbf{y}^T A)\mathbf{x}.$$

Now we can move the brackets (associativity), and use the fact that $A\mathbf{x} \leq b$, to get

$$(\mathbf{y}^T A)\mathbf{x} = \mathbf{y}^T(A\mathbf{x}) \leq \mathbf{y}^T b.$$

□

The amazing (and deep) result here is to show that the dual actually gives a perfect upper bound on the primal (assuming some mild conditions).

Theorem 2: Strong Duality Theorem

Suppose the primal LP (4) is feasible (i.e., it has at least one solution) and bounded (i.e., the optimal value is not ∞). Then the dual LP (5) is also feasible and bounded. Moreover, if $\mathbf{x}^*$ is the optimal primal solution, and $\mathbf{y}^*$ is the optimal dual solution, then

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

In other words, the maximum of the primal equals the minimum of the dual.

We will not prove Theorem 2 in this course, though the proof is not difficult. But let's give a geometric intuition of why this is true in the next section. Why is this useful? If I wanted to prove to you that $\mathbf{x}^*$ was an optimal solution to the primal, I could give you the solution $\mathbf{y}^*$, and you could check that $\mathbf{x}^*$ was feasible for the primal, $\mathbf{y}^*$ feasible for the dual, and they have equal objective function values.

This relationship is like in the case of *s*-*t* flows: the max flow equals the minimum cut. Or like in the case of zero-sum games: the payoff for the optimal strategy of the row player equals the (negative) of the payoff of the optimal strategy of the column player. Indeed, both these things are just special cases of strong duality!

2.2 Using duality to determine feasibility and boundedness

In addition to helping us bound feasible solutions to our LPs, duality can also be used as a tool to determine when certain programs are feasible or infeasible, or perhaps show that they are bounded or unbounded.

- If the primal is feasible and bounded, strong duality tells us that the dual is also feasible and bounded.
- Suppose the primal (maximization) problem is unbounded. What can duality tell us? Weak duality says $\mathbf{c}^T x \leq \mathbf{b}^T y$... If there existed any feasible $\mathbf{y}$ for the dual, this would imply that the primal is bounded, and hence by the contrapositive, if the primal is unbounded, then the dual *must be infeasible*. (This should make intuitive sense, the point of the dual was to provide an upper bound on the primal. If the primal is unbounded, then we can't find an upper bound.)

- By the exact same logic (reversed), if the dual is unbounded, since the primal is a lower bound on the dual, the primal must be infeasible.
- Can both the primal and dual be unbounded? No, because as the two previous points show, if one of them is unbounded, then the other is infeasible, and if a program is infeasible, it certainly can not be unbounded.

We can use these facts to represent all of the possible situations in a table like so:

		Dual		
		Inf	**F&B**	**Unb**
Primal	**Inf**	✓	X	✓
	F&B	X	✓	X
	Unb	✓	X	X

Here, **Inf** means infeasible, **F&B** means feasible and bounded, and **Unb** means unbounded. The only scenario that duality does not cover for us is the top-left cell, can an LP and its dual both be infeasible?

Exercise: Both primal and dual can be infeasible

Find an LP that is infeasible such that its dual is also infeasible.

Remark: Usefulness

This table has some very useful implications. If we have an LP for some problem, we might want to prove conditions on when it is feasible or infeasible. Directly proving that the LP is infeasible might be too difficult. Instead, if we can write the dual program and give a proof that the dual is unbounded, then we have indirectly proven that the primal is infeasible! A useful trick.

2.3 The Geometric Intuition for Strong Duality

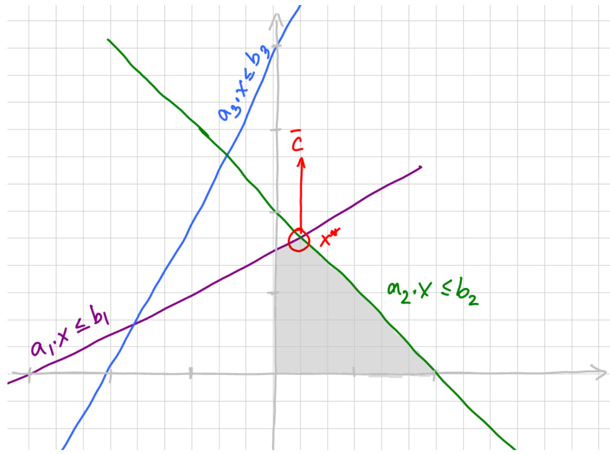
To give a geometric view of the strong duality theorem, consider an LP of the following form:

$$\begin{aligned} \text{maximize } & \mathbf{c}^T\mathbf{x} \\ \text{subject to } & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

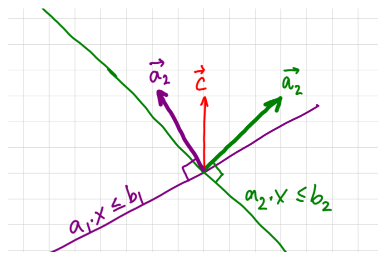
For concreteness, let's take the following 2-dimensional LP:

$$\begin{aligned} \text{maximize } & x_2 \\ \text{subject to } & -x_1 + 2x_2 \leq 3 \\ & x_1 + x_2 \leq 2 \\ & -2x_1 + x_2 \leq 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

If $\mathbf{c} := (0, 1)$, then the objective function wants to maximize $\mathbf{c} \cdot \mathbf{x}$, i.e., to go as far up in the vertical direction as possible. As we have already argued before, the optimal point $\mathbf{x}^*$ must be obtained at the intersection of two constraints for this 2-dimensional problem (n tight constraints for n dimensions). In this case, these happen to be the first two constraints.



If $\mathbf{a}_1 = (-1, 2), b_1 = 3$ and $\mathbf{a}_2 = (1, 1), b_2 = 2$, then $\mathbf{x}^*$ is the (unique) point $\mathbf{x}$ satisfying both $\mathbf{a}_1 \cdot \mathbf{x} = b_1$ and $\mathbf{a}_2 \cdot \mathbf{x} = b_2$. Indeed, we're being held down by these two constraints. Geometrically, this means that $\mathbf{c} = (0, 1)$ lies "between" these the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ that are normal (perpendicular) to these constraints.



Consequently, $\mathbf{c}$ can be written as a positive linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$. (It "lies in the cone formed by $\mathbf{a}_1$ and $\mathbf{a}_2$.") I.e., for some positive values y_1 and y_2,

$$\mathbf{c} = y_1 \mathbf{a}_1 + y_2 \mathbf{a}_2.$$

Great. Now, take dot products on both sides with $\mathbf{x}^*$. We get

$$\begin{aligned} \mathbf{c} \cdot \mathbf{x}^* &= (y_1 \, \mathbf{a}_1 + y_2 \, \mathbf{a}_2) \cdot \mathbf{x}^* \\ &= y_1(\mathbf{a}_1 \cdot \mathbf{x}^*) + y_2(\mathbf{a}_2 \cdot \mathbf{x}^*) \\ &= y_1 b_1 + y_2 b_2 \end{aligned}$$

Defining $\mathbf{y} = (y_1, y_2, 0, \ldots, 0)$, we get

$$\text{optimal value of primal} = \mathbf{c} \cdot \mathbf{x}^* = \mathbf{b} \cdot \mathbf{y} \geq \text{value of dual solution } \mathbf{y}.$$

The last inequality follows because

- the $\mathbf{y}$ we found satisfies $\mathbf{c} = y_1 \mathbf{a}_1 + y_2 \mathbf{a}_2 = \sum_i y_i \mathbf{a}_i = A^T \mathbf{y}$, and hence $\mathbf{y}$ satisfies the dual constraints $\mathbf{y}^T A \geq \mathbf{c}^T$ by construction.

In other words, $\mathbf{y}$ is a feasible solution to the dual, has value $\mathbf{b} \cdot \mathbf{y} \leq \mathbf{c} \cdot \mathbf{x}^*$. So the *optimal* dual value cannot be less. Combined with weak duality (which says that $\mathbf{c} \cdot \mathbf{x}^* \leq \mathbf{b} \cdot \mathbf{y}$), we get strong duality

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

Above, we used that the optimal point was constrained by two of the inequalities (and that these were not the non-negativity constraints). The general proof is similar: for n dimensions, we just use that the optimal point is constrained by n tight inequalities, and hence $\mathbf{c}$ can be written as a positive combination of n of the constraints (possibly some of the non-negativity constraints too).

3 Example: Zero-Sum Games

Consider a 2-player zero-sum game defined by an n-by-m payoff matrix R for the row player. That is, if the row player plays row i and the column player plays column j then the row player gets payoff R_{ij} and the column player gets $-R_{ij}$. To make this easier on ourselves (it will allow us to simplify things a bit), let's assume that all entries in R are positive (this is really without loss of generality since as pre-processing one can always translate values by a constant and this will just change the game's value to the row player by that constant). We saw we could write this as an LP:

- Variables: $v, p_1, p_2, \ldots, p_n$.
- Maximize v,
- Subject to:

  $p_i \geq 0$ for all rows i,

  $\sum_i p_i = 1$,

  $\sum_i p_i R_{ij} \geq v$, for all columns j.

To put this into the form of (4), we can replace $\sum_i p_i = 1$ with $\sum_i p_i \leq 1$ since we said that all entries in R are positive, so the maximum will occur with $\sum_i p_i = 1$, and we can also safely add in the constraint $v \geq 0$. We can also rewrite the third set of constraints as $v - \sum_i p_i R_{ij} \leq 0$. This then gives us an LP in the form of (4) with

$$\mathbf{x} = \begin{bmatrix} v \\ p_1 \\ p_2 \\ \ldots \\ p_n \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ 1 \end{bmatrix}, \text{and } A = \left[\begin{array}{c|ccc} 1 & & & \\ 1 & & -R^T & \\ \ldots & & & \\ 1 & & & \\ \hline 0 & 1 & \ldots & 1 \end{array}\right].$$

I.e., maximizing $\mathbf{c}^T\mathbf{x}$ subject to $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

We can now write the dual, following (5). Let $\mathbf{y}^T = (y_1, y_2, \ldots, y_{m+1})$. We now are asking to minimize $\mathbf{y}^T\mathbf{b}$ subject to $\mathbf{y}^T A \geq c^T$ and $\mathbf{y} \geq \mathbf{0}$. In other words, we want to:

- Minimize y_{m+1},
- Subject to:

  $y_1 + \ldots + y_m \geq 1$,

  $-y_1 R_{i1} - y_2 R_{i2} - \ldots - y_m R_{im} + y_{m+1} \geq 0$ for all rows i,

  or equivalently,

  $y_1 R_{i1} + y_2 R_{i2} + \ldots + y_m R_{im} \leq y_{m+1}$ for all rows i.

So, we can interpret y_{m+1} as the value to the row player, and $y_1, \ldots, y_m$ as the randomized strategy of the column player, and we want to find a randomized strategy for the column player that minimizes y_{m+1} subject to the constraint that the row player gets *at most* y_{m+1} no matter what row he plays. Now notice that we've only required $y_1 + \ldots + y_m \geq 1$, but since we're minimizing and the R_{ij}'s are positive, the minimum will happen at equality.

Notice that the fact that the maximum value of v in the primal is equal to the minimum value of y_{m+1} in the dual follows from strong duality. Therefore, the minimax theorem is a corollary to the strong duality theorem.