Lecture 5 - Duality

Optimization and Approximation - ENS M1

Nicolas Bousquet

1 *Introduction to duality*

1.1 Definition

Consider the following LP:

$$\max c^t x$$

subject to

$$Ax \leq b \text{ and } x \geq 0$$

The *dual* of this linear program is the following LP:

$$\min b^t y$$

subject to

$$y^t A \geq c \text{ and } y \geq 0$$

Since this second problem has a name, it is quite convenient to give a name to the first one: it is called the *primal*. In the following, the primal is denoted by (P) and the dual is denoted by (D)

Note that, the variables of the dual, called *dual variables*, correspond to the constraints in the primal. Indeed we multiply the matrix A by the left (recall that we want to "buy" the constraints, so each variable of the dual will provide the "price" of a constraint). And similarly, each variable of the primal corresponds to a constraint in the dual (we want to be sure that for each possible production it is better for the producer to sell constraints than to produce something).

This correspondence constraint - variable is the core of the duality in LP. You have to keep it in mind at any time !

Example Let us consider the "Dovetail" example:

$$\max_{(x_1,x_2)\in\mathbb{R}^2} 3 \cdot x_1 + 2 \cdot x_2.$$

subject to

$$\begin{aligned} 3 \cdot x_1 + x_2 \leq\ & 18 \\ x_1 + x_2 \leq\ & 9 \\ x_1 \leq\ & 7 \\ x_2 \leq\ & 6 \\ x_1, x_2 \geq\ & 0 \end{aligned}$$

The dual of this LP is:

$$\min_{(y_1,y_2,y_3,y_4)\in\mathbb{R}^4} 18 \cdot y_1 + 9 \cdot y_2 + 7y_3 + 6y_4.$$

subject to

$$\begin{aligned} 3y_1 + y_2 + y_3 \geq\ & 3 \\ y_1 + y_2 + y_4 \geq\ & 2 \\ y_1, y_2, y_3, y_4 \geq\ & 0 \end{aligned}$$

1.2 Interpretation of duality for production

So far we have introduced a new LP in a way that might seem completely arbitrary. We will see later that this dual LP actually has some strong links in a mathematical way with the primal one. But let us first exhibit a natural "interpretation" of the dual linear program.

In the Dovetail example, the first constraint correspond to the amount of wood available for the company. Instead of using it for its own production, the company might decide to sell this wood. There is then a natural question: when is it interesting for the company to sell this wood? and at which price? Let us denote by y_1 the price at which the company decides to sell one unit (cubic meter) of wood. Similarly, the company might also decide to sell some "production time" on its machine. Let us denote by y_2 the price at which it decides to sell one unit of production time. It might also decide to sell long tail boxes (y_3) and short ones (y_4).

The company decides to do that, but does not want to lose money (business is business). In particular by producing $100,000$ boxes of long tails, the company earns $3,000$\$. And in order to do that, the company "uses" 3 units of wood, 1 unit of machine and 1 unit of long boxes. So we should have $3y_1 + y_2 + y_3 \geq 3$, otherwise it would not be interesting for the company to sell these items / capacities but rather to produce, they will earn more money. Similarly, for short tails we have the constraint $y_1 + y_2 + y_4 \geq 2$. Which exactly gives us the constraints of the Dual LP !

Assume finally that you want to buy the ressources of the company. Which price would you pay for each ressource in order to be sure that the company will sell them to you? Your goal is indeed to minimize the total price of $18y_1 + 9y_2 + 7y_3 + 6y_4$... And you want to be sure that the company is willing to sell you the ressources, ie you must satisfy the above constraints.

1.3 General shape of the dual

We have seen the definition of dual when the LP is in canonical form. Though, in many cases, the problem is not in this form. We have seen that it is always possible to transform a LP in canonical form. Using this transformation, we'll see that every type of inequalities and of variables have their equivalent in the dual. Let us first see what happen if we have $\geq$ inequalities or equalities in the constraints.

Assume that (P) has the following constraint: $a_i^t x \geq b_i$. By multiplying by minus one this inequality we obtain: $-a_i^t x \leq -b_i$. In the dual, we create some a non-negative variable y_i such that the coefficient of y_i in the objective function is $-b_i$ and the coefficient of y_i in the j-th constraint (of the dual) is $-a_{i,j}$. Now let us replace y_i by a new variable y_i' such that y_i' is non-positive, the coefficient of y_i' in the objective function is b_i and the coefficient of y_i' in the j-th constraint of the dual is $a_{i,j}$. A point is a solution of this new LP if and only if it is a solution of the original one with $y_i' = -y_i$.

So an **at least inequality of the primal is transformed into a non-positive variable of the dual**.

Let us now consider the case where we have an equality constraint $a_i^t x = b_i$ (it is for instance the case if the LP is given in standard form). We have to transform this equality into inequalities. As we have already seen, this inequality can be transformed into the two following inequalities constraints:

$$a_i^t x \geq b_i$$
$$a_i^t x \leq b_i$$

So, in the dual (D) we create two variables, denoted by y_i^+ and y_i^- such that y_i^+ is non-negative (it corresponds to the first of these two constraints) and y_i^- is non-positive (according to what we've just seen). By construction of the dual, in the objective function, both have coefficient b_i. Let us now look at the constraints of the dual. For the j-th constraint of the dual, the coefficients of y_i^+ and of y_i^- are the same: it is the coefficient $a_{i,j}$. So in every constraint, we have $a_{i,j}(y_i^+ + y_i^-)$.

Consider now the linear program (D_2) where the variables y_i^+ and y_-^- are replaced by a unique variable y_i which is unconstrained. The coefficient of y_i in the objective function is b_i and the coefficient of y_i for every constraint is the coefficient of both y_i^+ and y_-^- in the corresponding constraint. Let us show that any solution of (D_2) can be transformed into a solution of (D) and conversely.

Let y be a solution of (D). By putting $y_i' = y_i^+ + y_i^-$ we obtain a solution of (D_2) with the same objective value satisfying all the constraints. Conversely, if y' is a solution of (D_2), then by putting $y_i^+ = y_i'$ or $y_i^- = y_i'$ depending on the sign of y_i', we obtain a solution of (D) with the objective value. So the optimal value of both LP are the same. So **an equality constraint is transformed into a unconstrained variable in the dual**.

We can use the same type of arguments to show that each non-negative variable leads to a $\geq$ constraints, each non-positive variable leads to a $\leq$ constraints and each unconstrained variable leads to an equality constraint (do it!). Nevertheless, we wont' prove it there since it will be a direct consequence of the fact that the dual of a dual is the primal. So, to conclude, we have the following correspondence:

$Primal$	$Dual$
max	min
Vector of constraint	Objective function
Objective function	Vector of constraint
Variables	Constraints
Constraints	Variables
Constraint $\leq$	Variable ≥ 0
Constraint $\geq$	Variable ≤ 0
Constraint $=$	Variable unconstrained
Variable ≥ 0	Constraint $\geq$
Variable ≤ 0	Constraint $\leq$
Variable unconstrained	Constraint $=$

We will not prove these "equivalences" during the lectures. Some of them will be treated during the exercises sessions.

Exercise 1. *Show that the dual of a non-negative variable is a $\geq$ constraint.*
Show that the dual of a non-positive variable is a $\leq$ constraint.
Show that the dual of a $\leq$ constraints is a non negative variable.

2 Duality theorems

2.1 Dual of dual

Theorem 1. *Dual of dual is primal.*

Proof. Consider a LP (P) in canonical form and its dual (D):

$$(P) \quad \begin{aligned} & z_p = \max c^T x \\ & \text{such that} \\ & Ax \leq b \\ & x \geq 0 \end{aligned} \qquad\qquad (D) \quad \begin{aligned} z_d = \min b^T y \\ \text{such that} \\ A^T y \geq c \\ y \geq 0 \end{aligned}$$

Let us rewrite the dual in order to obtain a canonical form:

$$\begin{aligned} -z_d = -\max -b^T y \\ \text{such that} \\ -A^T y \leq -c \\ y \geq 0 \end{aligned}$$

Note that optimal value of this LP is the opposite of the optimal value of the dual and the optimal points are the same. So finally, we can write the dual of this LP and obtain

$$\begin{aligned} -(z_d)_d = -\min -c^T x' \\ \text{such that} \\ (-A^T)^T x' \geq -b \\ x' \geq 0 \end{aligned} \qquad \Leftrightarrow \qquad \begin{aligned} z_p = \max c^T x \\ \text{such that} \\ Ax \leq b \\ x \geq 0 \end{aligned}$$

which concludes the proof. $\square$

2.2 Weak and strong duality

Consider a LP (P) and its dual (D) in the following form:

$$(P)\quad \begin{aligned} & z_p = \max c^T x \\ & \text{such that} \\ & Ax \leq b \\ & x \geq 0 \end{aligned} \qquad\qquad (D)\quad \begin{aligned} z_d = \min b^T y \\ \text{such that} \\ A^T y \geq c \\ y \geq 0 \end{aligned}$$

Recall that:

- If we consider a maximization problem and the function f is unbounded in F we say that the optimal value is $+\infty$. If F is empty we say that the optimal value is $-\infty$.
- If we consider a minimization problem and the function f is unbounded in F we say that the optimal value is $-\infty$. If F is empty we say that the optimal value is $+\infty$.

In both cases, if f is bounded, we say that the optimization problem has *an optimal solution* (in the sense that the optimal solution can be reached).

Theorem 2 (Weak Duality). *Assume that both (P) and (D) are feasible. The objective value of any point of (D) is larger or equal to the objective value of any point of (P). In particular we have:*

$$z_p \leq z_d.$$

Proof. By assumption, (P) has some feasible solution x and (D) has some feasible solution y. Since they are feasible, we have

$$Ax \leq b \Rightarrow b^T y \geq (x^T A^T) y$$

since y is non-negative. Moreover we have

$$A^T y \geq c \Rightarrow (y^T A) x \geq c^T x$$

Since x is non-negative. Note that the $(x^T A^T) y = (y^T A) x$ since the transpose of a real number is itself. So finally we obtain

$$b^T y \geq c^T x$$

which is satisfied for any pair of points x, y such that x is in the feasible region of (P) and y in the feasible region of (D). So in particular this inequality holds for the optimal points of respectively (P) and (D). And then we have the conclusion. □

The proof of the theorem shows that as soon as we have a feasible primal solution x and a feasible dual solution y, we obtain an upper bound ($b^T y$) for the primal optimal objective value z_p, and a lower bound (cx) for the dual optimal objective value z_d. From this it follows that if the primal problem is unbounded (*i.e.* has an unbounded solution), then the dual problem is infeasible (it is the only way for a minimization function to be infinite). Similarly, if the dual problem is unbounded, the primal problem is infeasible (it is the only way for a maximization function to reach $-\infty$). Note that this explains the conventions given before the weak duality theorem.

The weak duality theorem can be strengthen into the following theorem called the strong duality theorem.

Theorem 3 (Strong Duality). *If the $z_p \neq -\infty$ we have:*

$$z_p \geq z_d.$$

The principle of the proof is the following: given an optimal point x of (P), we want to find some y such that y is in the feasible region of (D) and such that the value of y is the same as the value of x. The weak duality theorem then ensures that the two values z_p and z_d are equal. Note that we will not prove the y is an optimal point, but this fact follows from the Weak Duality Theorem.

Proof. The LP can be written in the following form when we add the slack variables.

$$(P) \qquad \begin{aligned} & z_p = \max c^T x \\ & \text{such that} \\ & \begin{pmatrix} A & I \end{pmatrix} \begin{pmatrix} x \\ x_s \end{pmatrix} = b \\ & x \geq 0 \end{aligned}$$

where x_s is the set of slack variables. Recall the LP with the slack variables has the same optimal value as the LP without the slack variables. Assume that the Simplex Method has pivoted to an optimal solution given by the basis B. Let NB be the complement of B. Let us denote by B the matrix (be careful two notations for the same thing) of the columns of A corresponding to the columns of the variables of the basis B. Let A_N be the other columns of the matrix. Let us denote by x_B the vector of variables of B and by x_N the vector of variables of NB. We have:

$$\begin{pmatrix} A & I \end{pmatrix} \begin{pmatrix} x \\ x_s \end{pmatrix} = Bx_B + A_N x_N = b \Leftrightarrow x_B = B^{-1}b - B^{-1}A_N x_N$$

Moreover the objective function can be rewritten in the following way:

$$c_B^T x_B + c_N^T x_N = c_B^T B^{-1} b + (c_N^T - c_B^T B^{-1} A_N) x_N$$

The simplex algorithm terminates if the coefficients of the non-basis variables in the objective function are non-positive (indeed in this form, the objective function is expressed in terms of non-basic variables). So we have,

$$c_N^T - c_B^T B^{-1} A_N \leq 0$$

Moreover, we have the following inequality: $c_B^T - c_B^T B^{-1} B \leq 0$. So if we "combine" these two inequalities into a single inequality (recall that the matrix A is the matrix A_N plus the columns of the matrix B)

$$c^T - c_B^T B^{-1} A \leq 0$$

In particular, if we restrict this equation to the variables of x_S (or differently to the columns corresponding to the vertices of x_S) we have:

$$0^T - c_B^T B^{-1} I \leq 0$$

Indeed the initial vector c was completed with 0 when we add slack variables (the objective function is a function of the original variables). Now let us consider

$$y^T = c_B^T B^{-1}$$

Let us show that y is a feasible solution of (D) and then we will show that the value of the objective function of (D) at point y equals z_p. First we have

$$-c_B^T B^{-1} I \leq 0 \Leftrightarrow y \geq 0$$

Moreover we have:

$$c^T - c_B^T B^{-1} A \leq 0 \Rightarrow y^T A \geq c^T \Rightarrow A^T y \geq c$$

So the point y satisfies the constraints of (D). Now let x^* be the optimal point of (P) which corresponds to the basic feasible solution B. All the non-basic coefficients of x^* are equal to zero (by definition of B). So in other words, we have $x_N^* = 0$. So the objective function on x^* equals $c_B^T B^{-1} b$. Let us now compute the value of the objective function of (D) at point y:

$$b^T y = b^T (B^{-1})^T c_B$$

which is precisely the value of the objective function in the point x^* which achieves the proof of the Strong Duality Theorem. □

Note that the Strong Duality theorem ensures that if a solution is feasible in the primal then its dual solution is feasible in the dual if and only if it is an optimal solution of the primal. What we mean by this is the following: if we look at the "prices" for each constraint given by some basic feasible solution of the primal, then it does not satisfy the constraints of the dual except if the solution is optimal.

2.3 Unbounded or unfeasible Linear Programs

Recall that by convention we have:

$$\min\{c^T x | x \in F = \varnothing\} = +\infty \qquad \text{and} \qquad \max\{c^T x | x \in F = \varnothing\} = -\infty$$

In the previous section, we have seen that if both (P) and (D) are feasible then $z_p = z_d$. One may ask what happens in the other cases. This section is devoted to understand what happens in these cases. Let us show the following:

Lemma 4. *If (P) is infeasible then (D) is unbounded or infeasible.*

Proof. The proof is based on the Farkas' Lemma. Up to an addition of the slack variables, we can assume that we are looking for a solution of

$$\begin{pmatrix} A & I \end{pmatrix} \begin{pmatrix} x \\ x_s \end{pmatrix} = b$$
$$x, x_s \geq 0$$

in the primal (P) with the objective function $\max c^t x$. So in the dual we are looking for:

$$A^t y \geq c$$

with unconstrained variables y_i and with the objective function $\min b^t y$.

Assume now that (D) has a solution. So there exists $z \geq 0$ such that $A^t z \geq c$. The Farkas' Lemma ensures moreover that there exists y such that

$$y^t \begin{pmatrix} A & I \end{pmatrix} \geq 0 \text{ and } b^t y < 0$$

Indeed, the equation $\begin{pmatrix} A & I \end{pmatrix} \binom{x}{x_s} = b, x, x_s \geq 0$ has no solution so the above equation has a solution. The set of "constraint" for y (which is the set of columns) can be divided into two parts: $A^t y \geq 0$ and $y \geq 0$. So in particular $y \geq 0$. Moreover $z + \lambda y$ is a solution (D) as long as $\lambda \geq 0$. Indeed $A^t(z + \lambda y) = A^t z + \lambda A^t y \geq b$ and $z + \lambda y \geq 0$. Let us finally show that the objective function of the dual tends to infinity when λ tends to infinity. If $\lambda \to +\infty$ then the objective function tends to $-\infty$ (since $b^t y < 0$) which achieves the proof.

So if (P) is infeasible then if (D) is feasible then (D) is unbounded. Note that we have not proved that it can be unbounded or infeasible, it will be done in exercise. □

As we have already mentioned, we have:

Lemma 5. *The dual of a unbounded linear program is not feasible.*

Proof. Assume that the linear program (P) is unbounded. By weak duality theorem, we know that the value of any solution of (D) (if it exists) is larger than any solution of (P). So in particular, if (P) is unbounded it would mean that any solution of (D) is infinite which is a non-sense. So if (P) is unbounded then (D) is not feasible. □

In the following array, p^* denotes the optimal value of the primal while d^* denotes the optimal value of the dual. A 0 means that this case cannot happen while a 1 ensures that this case can happen.

	$p^* = -\infty$	p bounded	$p = +\infty$
$d^* = -\infty$	1	0	0
d^* bounded	0	1	0
$d^* = +\infty$	1	0	1

Exercise 2. 1. *Find a linear program (P) such that both (P) and its dual (D) are not feasible.*

2. *Find a linear program (P) such that (P) is not feasible and (D) is unbounded.*

3 Complementary slackness and certificate of optimality

Let us finally prove the following theorem:

Theorem 6 (Complementary Slackness). *Let x be a feasible solution to the primal and y be a feasible solution to the dual. Then x is an optimal point of (P) and y is an optimal point of (D) if and only if the conditions of Complementary Slackness hold:*

$$(b_i - \sum_{j=1}^{n} a_{ij}x_j)y_i = x_{n+i}y_i = 0 \text{ for } i = 1, 2, \ldots, m$$

where x_{n+i} is the slack variable of the i-th constraint of (P) and

$$(\sum_{i=1}^{m} a_{ji}y_i - c_j)x_j = y_{m+j}x_j = 0 \text{ for } i = 1, 2, \ldots, n$$

where y_{m+j} is the slack variable of the j-th constraint of (D).

Interpretation Let us try to interpret this inequality before proving it. Consider the first inequality. It means that there are two options, either the constraint is tight or the value of the dual variable corresponding to this equality equals 0. And it exactly corresponds to what we have seen before in that lecture. Indeed, either the constraint is tight and we are done, or the constraint is not tight. But in that case, the shadow price of that constraint is equal to zero. And then the optimal solution of the dual for that constraint equal zero !

The other inequality can be interpreted similarly.

Proof of Theorem 6 By the weak duality theorem we have

$$c^t x \leq y^t Ax \leq b^t y$$

where the mid-equality is correct since $Ax \leq b$. So we have:

$$(c^t - y^t A)x \leq 0$$
$$y^t(b - Ax^t) \geq 0$$

Moreover, the Strong duality theorem ensures that the solutions are optimal if and only if both equalities equal 0. Note that the value of the i-th coordinate of $(b - Ax^t)$ is precisely the value of the slack variable of the i-th variable. The fact that this coordinate is equal to zero precisely ensures that either $(b_i - \sum_{j=1}^{n} a_{ij}x_j)$ or y_i equals 0. The second case is handled symmetrically using the first equation.

Certificate of optimality Let us consider the "Dovetail" example:

$$\max_{(x_1,x_2)\in\mathbb{R}^2} 3 \cdot x_1 + 2 \cdot x_2.$$

subject to

$$\begin{aligned} 3 \cdot x_1 + x_2 \leq&\quad 18 \\ x_1 + x_2 \leq&\quad 9 \\ x_1 \leq&\quad 7 \\ x_2 \leq&\quad 6 \\ x_1, x_2 \geq&\quad 0 \end{aligned}$$

Its dual is the following LP.

$$\min_{(y_1,y_2,y_3,y_4)\in\mathbb{R}^4} 18 \cdot y_1 + 9 \cdot y_2 + 7y_3 + 6y_4.$$

subject to

$$\begin{aligned} 3y_1 + y_2 + y_3 \geq & \quad 3 \\ y_1 + y_2 + y_4 \geq & \quad 2 \\ y_1, y_2, y_3, y_4 \geq & \quad 0 \end{aligned}$$

Assume that someone is claim that the optimal solution is $(3, 6)$. Can we find an (efficient) way of checking if it is true or not without computing the optimal solution? The answer is positive ! One can easily check that two constraints are tight for these values ($x_2 \leq 6$ and $x_1 + x_2 \leq 9$) and that all the other constraints are satisfied.

Now we can use the complementary slackness to find a contradiction. Indeed it ensures that, since the first and the third constraint are not tight, the dual variable should be equal to 0. And then $y_1^* = y_3^* = 0$ in an optimal solution of the dual y^* of the dual. We moreover now that, since x_t and x_c are positive, the corresponding constraints of the dual have to be tight. So we must have:

$$\begin{aligned} y_2^* = & \quad 3 \\ y_2^* + y_4^* = & \quad 2 \end{aligned}$$

So the corresponding solution of the dual would be $(0, 3, 0, -1)$, which does not satisfy the non negativity constraints, a contradiction ! Si $(3, 6)$ was not a solution of the primal...