LP. Lecture 5

Chapter 5: duality theory

- motivation
- the dual problem
- weak and strong duality
- the dual of LP problems in other forms

Motivation

Associated to every LP problem (P) there is another "mirrored" LP problem (D). Here (D) is called the dual problem of (P), and (P) is called the primal problem. It turns out that the dual problem of (D) is (P)! (Double mirroring!)

LP problems occur in couples: one primal and one dual problem.

The duality theory is useful because:

- the dual problem can be used to quickly give bounds of the optimal value of an LP problem
- instead of solving an LP problem (P) one may solve the dual (D). One will get a solution of (P) "for free"! This can be more efficient.

The dual problem

Consider the LP problem (P), the primal problem, given by

$$
\begin{array}{llll}
\text{(P)} & \max & \sum_{j=1}^{n} c_j x_j & \\
& \text{s.t.} & & \\
& & \sum_{j=1}^{n} a_{ij} x_j \leq b_i & \text{for } i = 1, \ldots, m \\
& & x_j \geq 0 & \text{for } j = 1, \ldots, n.
\end{array}
$$

We define the dual problem (D) like this:

$$
\begin{array}{llll}
\text{(D)} & \min & \sum_{i=1}^{m} b_i y_i & \\
& \text{s.t.} & & \\
& & \sum_{i=1}^{m} y_i a_{ij} \geq c_j & \text{for } j = 1, \ldots, n \\
& & y_i \geq 0 & \text{for } i = 1, \ldots, m.
\end{array}
$$

Rules to remember:

	x_1	$\dots$	x_n	
y_1	$a_{1,1}$	$\dots$	$a_{1,n}$	b_1
$\vdots$		$\vdots$		$\vdots$
y_m	$a_{m,1}$	$\dots$	$a_{m,n}$	b_m
	c_1	$\dots$	c_n	

Now, let $A = [a_{ij}]$ be the coefficient matrix.

Observe:

- (D): the variables are associated to the rows in A, while the constraints are attached to the columns in A
- (P): reversed! Then: the variables are associated to the columns in A, while the constraints are associated to the rows in A
- b_i-s make up the right-hand side in (P), but are included in the objective function in (D)
- c_j-s are a part of the objective function in (P), but constitute the right-hand side in (D)

- the constraints in (D) are $\geq$
- (D) is also a LP problem. We will soon rewrite it to its "standard form".

We will first give an important result which is the motivation for duality: any feasible solution of an LP problem is the source of a bound of the optimal value in the dual.

Theorem 5.1: (Weak duality) *If* $(x_1, \dots, x_n)$ *is feasible in (P) and* $(y_1, \dots, y_m)$ *is feasible in (D) we have*

$$\sum_{j=1}^{n} c_j x_j \leq \sum_{i=1}^{m} b_i y_i.$$

Proof: From the constraints in (P) and (D) we have

$$\sum_{j=1}^{n} c_j x_j \leq \sum_{j=1}^{n} (\sum_{i=1}^{m} y_i a_{ij}) x_j = \sum_{i=1}^{m} y_i \sum_{j=1}^{n} a_{ij} x_j \leq \sum_{i=1}^{m} y_i b_i.$$

$\square$

Example 1:

$$
\begin{array}{llllll}
\text{(P)} & \text{maximize} & 5x_1 + & 6x_2 + & 8x_3 & \\
& \text{assuming that} & & & & \\
& & x_1 + & 2x_2 + & 3x_3 & \leq 5 \\
& & 4x_1 + & 5x_2 + & 6x_3 & \leq 11 \\
& & & x_1, x_2, x_3 \geq 0. & &
\end{array}
$$

$$
\begin{array}{lllll}
\text{(D)} & \text{minimize} & 5y_1 + & 11y_2 & \\
& \text{assuming that} & & & \\
& & y_1 + & 4y_2 & \geq 5 \\
& & 2y_1 + & 5y_2 & \geq 6 \\
& & 3y_1 + & 6y_2 & \geq 8 \\
& & & y_1, y_2 \geq 0. &
\end{array}
$$

We now see that, for instance, $(y_1, y_2) = (1, 1)$ is an feasible solution in (D), and the corresponding value of the objective function in (D) is $5 + 11 = 16$. Then, the optimal value in (P) can not be more than 16. On the other side $(x_1, x_2, x_3) = (0, 0, 5/3)$ is feasible in (P) with corresponding value $\eta = 40/3 \approx 13.33$. So, the optimal value η^* in (P) must lie between 13.33 and 16.

How about $x^* = (x_1, x_2, x_3) = (1/2, 0, 3/2)$ and $y^* = (y_1, y_2) = (1/3, 7/6)$? We have that $\sum_{j=1}^{3} c_j x_j^* = 29/2$ and $\sum_{i=1}^{2} b_i y_i^* = 29/2$. But then it follows from weak duality that x^* is optimal in (P) and that y^* is optimal in (D)!

Weak duality gives a principle for showing optimality, or "almost-optimality".

Interpretation of (D): any feasible x in (P) satisfies $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$ and therefore also a nonnegative linear combination of these:

$$(*) \quad \sum_{i=1}^{m} y_i (\sum_{j=1}^{n} a_{ij}x_j) \leq \sum_{i=1}^{m} y_i b_i$$

Here y_i is a nonnegative multiplier for inequality nr. i.

If we also choose the y_i-s so that $\sum_{i=1}^{m} y_i a_{ij} \geq c_j$ the left side in $(*)$ will be $\geq \sum_{j=1}^{n} c_j x_j$. Then we have an upper bound for the optimal value, η^* is (P), namely $\sum_{i=1}^{m} y_i b_i$. We would like to have the best possible bound, which means lowest possible, and this gives the problem

$$\min\{\sum_{i=1}^{m} y_i b_i : \sum_{i=1}^{m} y_i a_{ij} \geq c_j \ \text{ for alle } j, \ \ y_i \geq 0 \ \text{ for alle } i\}$$

which is the dual problem!

Strong duality

Natural question: Weak duality implies that optimal value in (P) $\leq$ optimal value in (D). Can we have a *strict inequality* here? The answer is, among other things, important for testing of optimality. The answer is: *no*, except in very special situations. We have:

Theorem 5.2: (Strong duality) *If (P) has an optimal solution* $x^* = (x_1^*, \ldots, x_n^*)$*, then (D) has an optimal solution* $y^* = (y_1^*, \ldots, y_m^*)$ *so that*

$$\sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_i^* .$$

□

Consequence: (P) and (D) have the same optimal value when (P) has an optimal solution.

Will later discuss the situation where (P), and sometimes (D), is unbounded, or if neither (P) or (D) is feasible (this can happen, but not for "interesting problems").

Strong duality can be proved short via the simplex algorithm, especially in matrix notation. But to increase understanding we will stick to component notation and study closer what happens in (P) and (D) during a simplex pivot.

Pivot, primal and dual

Example: $m = 2$, $n = 3$. Introducing slack variables z_j in (D) and writing also (D) as a problem of maximization. In dictionary form:

$$
\text{(P)} \quad \begin{array}{rcrcrcrcr}
\eta & = & 0 & + & 4x_1 & + & x_2 & + & 3x_3 \\
\hline
w_1 & = & 1 & - & x_1 & - & 4x_2 & & \\
w_2 & = & 3 & - & 3x_1 & + & x_2 & - & x_3
\end{array}
$$

$$
\text{(D)} \quad \begin{array}{rcrcrcr}
-\xi & = & 0 & - & y_1 & - & 3y_2 \\
\hline
z_1 & = & -4 & + & y_1 & + & 3y_2 \\
z_2 & = & -1 & + & 4y_1 & - & y_2 \\
z_3 & = & -3 & & & + & y_2
\end{array}
$$

Note the " negative-transpose property" on the right side:

$$\begin{bmatrix} 0 & 4 & 1 & 3 \\ 1 & -1 & -4 & 0 \\ 3 & -3 & 1 & -1 \end{bmatrix} \leftrightarrow \begin{bmatrix} 0 & -1 & -3 \\ -4 & 1 & 3 \\ -1 & 4 & -1 \\ -3 & 0 & 1 \end{bmatrix}$$

Pivoting now in (P): x_3 into basis and w_2 out of basis. Do corresponding pivot in (D): x_3 corresponds to z_3 and w_2 corresponds to y_2. So, in (D) y_2 goes into basis and z_3 out of basis.

Note: the pivot is carried out in the regular way (switching roles + row operations) even though we "accidentally" don't have an feasible basis solution in (D). Result:

$$
(\mathrm{P})\quad \begin{array}{rcrcrcrcr}
\eta & = & 9 & - & 5x_1 & + & 4x_2 & - & 3w_2 \\
\hline
w_1 & = & 1 & - & x_1 & - & 4x_2 & & \\
x_3 & = & 3 & - & 3x_1 & + & x_2 & - & w_2
\end{array}
$$

$$
(\mathrm{D})\quad \begin{array}{rcrcrcr}
-\xi & = & -9 & - & y_1 & - & 3z_3 \\
\hline
z_1 & = & 5 & + & y_1 & + & 3z_3 \\
z_2 & = & -4 & + & 4y_1 & - & z_3 \\
y_2 & = & 3 & & & + & z_3
\end{array}
$$

Observe again that the negative-transpose property holds. In particular we see that the value of the primal solution equals the value of the dual solution. But the dual solution is not feasible.

New pivot: in (P): x_2 in and w_1 out. Corresponding pivot in (D): y_1 in and z_2 out. Result:

(P)

$$\begin{array}{rcrcrcrcr} \eta & = & 10 & - & 6x_1 & - & w_1 & - & 3w_2 \\ \hline x_2 & = & 0.25 & - & 0.25x_1 & - & 0.25w_1 & & \\ x_3 & = & 3.25 & - & 3.25x_1 & - & 0.25w_1 & - & w_2 \end{array}$$

(D)

$$\begin{array}{rcrcrcr} -\xi & = & -10 & - & 0.25z_2 & - & 3.25z_3 \\ \hline z_1 & = & 6 & + & 0.25z_2 & + & 3.25z_3 \\ y_1 & = & 1 & + & 0.25z_2 & + & 0.25z_3 \\ y_2 & = & 3 & & & + & z_3 \end{array}$$

Can now see that:

- the negative transpose property still holds
- optimal solution in (P), and therefore:
- for the first time the dual basis solution is feasible

Lemma PIV: (Pivot in (P) and (D)) *Assume that every pivot is done in both (P) and (D) so that if x_j replaces w_i in the primal basis, y_i will replace z_j in the dual basis. Then the negative-transpose property will hold in each iteration.* □

Exercise: prove Lemma PIV by checking the following:

(P)

b	$\dots$	a	
$\vdots$		$\vdots$	
d	$\dots$	c	

$\xrightarrow{\text{pivot}}$

$-b/a$	$\dots$	$1/a$	
$\vdots$		$\vdots$	
$d - bc/a$	$\dots$	c/a	

(D)

	$-b$	$\dots$	$-d$
	$\vdots$		$\vdots$
	$-a$	$\dots$	$-c$

$\xrightarrow{\text{pivot}}$

	b/a	$\dots$	$-d + bc/a$
	$\vdots$		$\vdots$
	$-1/a$	$\dots$	$-c/a$

Proof of strong duality:

From Lemma PIV it follows that in every iteration k we have a primal basis solution x^k and a dual basis solution y^k with the same value of the corresponding objective functions, which means that:

$$\sum_{j=1}^{n} c_j x_j^k = \sum_{i=1}^{m} b_i y_i^k.$$

The primal simplex algorithm terminates with an feasible basis variable x^* and this happens when all the coefficients in front of the nonbasic variables in (P) are nonpositive.

But by Lemma PIV this means that the corresponding dual basis solution y^* is feasible (the basis variables are nonnegative). As wished: $\sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_i^*$. □

Complementary slack

Shall study an optimality condition in LP; called complementary slack. Assume that $x = (x_1, \dots, x_n)$ is a feasible solution in (P) and that $y = (y_1, \dots, y_m)$ is a feasible solution in (D). (Whether they are basic solutions or not is of no importance now.)

Question: what is required for x to be optimal in (P) and y optimal in (D)?

Analysis: Since (P) and (D) have the same optimal value (consequence of strong duality) we see that: x and y are both optimal (according to (P) and (D)) if and only if

$$(*) \quad \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i.$$

But, from the constraints we get that (as in the proof for weak duality)

$$\sum_{j=1}^{n} c_j x_j \le \sum_{j=1}^{n} (\sum_{i=1}^{m} y_i a_{ij}) x_j = \sum_{i=1}^{m} y_i \sum_{j=1}^{n} a_{ij} x_j \le \sum_{i=1}^{m} y_i b_i.$$

So $(*)$ holds if and only if

- $\sum_{i=1}^{m} y_i a_{ij} = c_j$ if $x_j > 0$, and
- $\sum_{j=1}^{n} a_{ij} x_j = b_i$ if $y_i > 0$.

These two conditions are called complementary slack.

We have therefore shown the following result:

Theorem 5.3: (Complementary slack) *Assume that $x = (x_1, \ldots, x_n)$ is a feasible solution in (P) and that $y = (y_1, \ldots, y_m)$ is a feasible solution in (D). Let $(w_1, \ldots, w_m)$ be the corresponding primal slack variables, and $(z_1, \ldots, z_n)$ be the corresponding dual slack variables.*
Then x is optimal in (P) and y is optimal in (D) if and only if

$$x_j z_j = 0 \quad \text{for } j = 1, \ldots, n,$$
$$w_i y_i = 0 \quad \text{for } i = 1, \ldots, m.$$

□

Complementary slack therefore says: if there is slack in an inequality (the slack variably is positive) in one of the problems, the corresponding dual variable has to be zero.

Complementary slack is therefore an optimality property. Note that these conditions are nonlinear equations:

$$x_j z_j = 0 \quad (j \leq n).$$

This is *the nonlinearity of linear optimization* !! This makes LP more difficult to solve that linear equations. But this nonlinearity is still fairly simple, which may explain why LP problems can be solved so efficiently.

By the way: in interior point methods for LP, one uses Newton's method for solving a modified set of equations which consists of the original equations from (P) and (D) (where the slack variables are introduced), in addition to complementary slack).

"Schemes" for LP algorithms.

About algorithms for LP.

From Theorem 5.3 we can see that solving an LP problem consists of fulfilling three properties at once

- 1. primal feasibility,
- 2. dual feasibility, and
- 3. complementary slack.

One gets different algorithms by making sure that two of these properties hold in each iteration, while one strides for the third one to hold as well; then the problem is solved.

- The algorithm we have studied fulfills 1 and 3 and aims at 2; it is often called the primal simplex algorithm.
- Another possibility is to fulfill 1 and 2 and aim at 3; this results in so called primal-dual algorithms. (Both simplex and "non-simplex" algorithms).

The dual simplex algorithm:

- Fulfills properties 2 and 3, and aims at 1.
- Often used if it is easy to find a dual feasible initial solution, because then one does not have to do the Phase I problem (in primal simplex). Used for 'reoptimization": have solved a problem and will solve a new problem where we have added e.g. another constraint
- May be used for Phase 1: just insert another objective function so that initial dictionary is dual feasible!!
- Also used often if a problem has more constraints than variables; this reduces the number of pivots and is faster.
- corresponds to using the primal simplex algorithm on the dual problem, and this can be uses to perform the algorithm directly in the primal dictionary. Based on that the initial solution is dual feasible (coefficients in front of nonbasic variables are nonpositive).
- See section 6.6 and 6.7 for further details.

The dual simplex algorithm: example

$$\begin{array}{rcrcrcrcr} \eta & = & 12 & - & 4x_1 & - & x_2 & - & x_3 \\ \hline x_4 & = & -4 & + & 3x_1 & - & 11x_2 & + & x_3 \\ x_5 & = & 3 & - & x_1 & + & 3x_2 & - & 2x_3 \end{array}$$

1. dual pivot: x_4 leaves and x_3 enters (as $+x_3$ and $1/1 < 4/3$).

$$\begin{array}{rcrcrcrcr} \eta & = & 8 & - & x_1 & - & 12x_2 & - & x_4 \\ \hline x_3 & = & 4 & - & 3x_1 & + & 11x_2 & + & x_4 \\ x_5 & = & -5 & + & 5x_1 & - & 19x_2 & - & 2x_4 \end{array}$$

2. dual pivot: x_5 leaves and x_1 enters (as $+x_1$).

$$\begin{array}{rcrcrcrcr} \eta & = & 7 & - & 0.2x_5 & - & 15.8x_2 & - & 1.4x_4 \\ \hline x_3 & = & 1 & - & 0.6x_5 & - & 0.4x_2 & - & 0.2x_4 \\ x_1 & = & 1 & + & 0.2x_5 & + & 3.8x_2 & + & 0.4x_4 \end{array}$$

Dual feasible, so go on with primal pivots: done right away!

Duality, other forms

Our standard form of (P) and (D) is:

$$
\begin{array}{llrll}
\text{(P)} & \max & \sum_{j=1}^{n} c_j x_j & & \\
& \text{subject to} & & & \\
& & \sum_{j=1}^{n} a_{ij} x_j & \leq b_i & \text{for } i = 1, \ldots, m \\
& & x_j & \geq 0 & \text{for } j = 1, \ldots, n.
\end{array}
$$

$$
\begin{array}{llrll}
\text{(D)} & \min & \sum_{i=1}^{m} b_i y_i & & \\
& \text{subject to} & & & \\
& & \sum_{i=1}^{m} y_i a_{ij} & \geq c_j & \text{for } j = 1, \ldots, m \\
& & y_i & \geq 0 & \text{for } i = 1, \ldots, m.
\end{array}
$$

In matrix form:

$$\text{(P)} \qquad \max\{c^T x : Ax \leq b,\ x \geq O\}$$

$$\text{(D)} \qquad \min\{b^T y : A^T y \geq c,\ y \geq O\}.$$

One may meet LP problems in other forms. But: every LP problem may be rewritten in the form (P). To do so, certain techniques are needed:

- each equations is written as two inequalities
- $\min f = -\max(-f)$
- a free variable x is replaced by $x^+ - x^-$ where $x^+, x^- \geq 0$

One may then (if desirable) find the dual problem (since the primal now has the "right" form) and write this in the simplest form possible.

It is important to practise the techniques to

- write any LP problem on the form (P), and
- find the dual of any LP problem.

It is recommended to use the matrix form in this rewriting of the problem.

We then need to work on partitioned matrices, see section on this in the linear algebra book (MAT1120). In particular, we need a rule for matrix multiplication:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} = \begin{bmatrix} A_{11}x^1 + A_{12}x^2 \\ A_{21}x^1 + A_{22}x^2 \end{bmatrix}$$

Another useful rule is for the transpose of a partitioned matrix:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^T = \begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix}$$

Example:

$$\max\{c^T x^1 + d^T x^2 : A^1 x^1 \geq b^1,\ A^2 x^1 + A^3 x^2 \leq b^2,\ x^1, x^2 \geq O\}$$

Here the variables are x^1 og x^2 (suitable vectors). We may write this in the form (P):

$$\max\{\begin{bmatrix} c \\ d \end{bmatrix}^T \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} : \begin{bmatrix} -A^1 & O \\ A^2 & A^3 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} \leq \begin{bmatrix} -b^1 \\ b^2 \end{bmatrix},\ x^1, x^2 \geq O\}$$

Then the dual may be determined and, finally, one sees if the dual may be simplified.

This, and related, examples are given on the blackboard. (For instance, where a variable vector x is free (that is, no sign constraint) and is replaces by $x' - x''$ where $x', x'' \geq O$.)

Last comment on this, a connection between the primal and the dual:

- an equation in one of the problems corresponds to a free variable in the other problem,
- an inequality in one of the problems corresponds to a nonnegative variable in the other problem.