Duality

The discussion in these notes closely follows Chapter 4 in Bertsimas & Tsitsiklis, Introduction to Linear Optimization, Athena Scientific, 1997.

Consider an optimisation problem of constructing a minimum cost diet that meets a set of nutritional requirements, from a given basket of foodstuffs whose unit prices are specified. This can be posed as an LP.

Now consider someone who wants to sell nutritional supplement pills, and who wants to know how to price them. If they set the price too high, then people can meet all their nutritional requirements from foodstuffs and won't buy the pills.

How are the consumer's cost minimisation problem and the pill manufacturer's profit maximisation problem related? That is the question we will study now.

Consider the standard form problem

$$\begin{array}{ll} \min & \underline{c}^T \underline{x} \\ \text{Subj. to} & A\underline{x} = \underline{b} \\ & \underline{x} \geq \underline{0} \end{array} \quad , \ \underline{x} \in \mathbb{R}^n , \ \cancel{\underline{c} \in \mathbb{R}^n}$$

$A \in \mathbb{R}^{m \times n}$, $\underline{b} \in \mathbb{R}^m$, $\underline{c} \in \mathbb{R}^n$ are given

We call this the primal problem.

We can think of each constraint as being associated with a "resource" (e.g. vitamin A), and the resource as being available at some unit price (the price of a vitamin A pill).

So we can consider a "relaxation" of the problem where the constraints may not be satisfied, and so there is additional cost of acquiring the resources to ensure the constraints are met. This is a penalty associated with violating the constraints.

Thus, we have the modified problem:

$$\begin{array}{ll} \min & \underline{c}^T \underline{x} + \underline{p}^T (\underline{b} - A\underline{x}) \\ \text{subj. to} & \underline{x} \geq \underline{0} \end{array} \quad , \ \underline{x} \in \mathbb{R}^n , \ \cancel{\underline{p} \in \mathbb{R}^m}$$

where $\underline{p}^T = (p_1 \ p_2 \ \cdots \ p_m)$

and p_i is the unit cost of the i^{th} resource (which could be +ve or −ve).

Let $g(\underline{p})$ be the min. cost of this "relaxed" problem

It is easy to see that, for any choice of the price vector $\underline{p}$, the minimum cost $g(\underline{p})$ of the relaxed problem is a lower bound on that of the original problem.

Indeed, for any feasible $\underline{x}$, i.e., any $\underline{x}$ satisfying the constraints $A\underline{x} = \underline{b}$, $\underline{x} \geq \underline{0}$, the cost functions are identical:

$$\underline{c}^T \underline{x} = \underline{c}^T \underline{x} + \underline{p}^T (\underline{b} - A\underline{x})$$

In addition, there are values of $\underline{x}$ that are feasible for the relaxed problem but not the original problem.

Since $g(\underline{p})$ is a lower bound on the value of the primal problem for every price vector $\underline{p}$, so is the maximum of $g(\underline{p})$ over $\underline{p}$.

When is the resulting lower bound tight, i.e., when is the maximum of $g(\underline{p})$ over $\underline{p}$ equal to the minimum of $\underline{c}^T \underline{x}$ over feasible $\underline{x}$?

It turns out that this is always the case in LP problems, and is the main result of duality theory.

We now study this in a bit more detail.

Recall the definition of $g(\underline{p})$

$$g(\underline{p}) = \min_{\underline{x} \geq \underline{0}} \underline{c}^T \underline{x} + \underline{p}^T (\underline{b} - A\underline{x})$$

$$= \underline{p}^T \underline{b} + \min_{\underline{x} \geq \underline{0}} (\underline{c}^T - \underline{p}^T A)\underline{x}$$

Now,

$$\min_{\underline{x} \geq \underline{0}} (\underline{c}^T - \underline{p}^T A)\underline{x} = \begin{cases} 0 & , \text{ if } \underline{c}^T - \underline{p}^T A \geq \underline{0} \\ -\infty & , \text{ otherwise} \end{cases}$$

Thus, $g(\underline{p}) = \begin{cases} \underline{p}^T \underline{b} & , \text{ if } \underline{c}^T - \underline{p}^T A \geq \underline{0} \\ -\infty & , \text{ otherwise} \end{cases}$

The problem of maximising $g(\underline{p})$, i.e., finding the best lower bound on the primal problem, is called the dual problem.

Thus, we see that the dual problem is

$$\begin{aligned} &\max \quad \underline{p}^T \underline{b}, \quad \text{over } \underline{p} \in \mathbb{R}^m \\ &\text{subject to } \underline{p}^T A \leq \underline{c}^T. \end{aligned}$$

Note that the dual problem is also an LP. As written above it is not in standard form, but it can be put in standard form.

The main thing to take away from the discussion so far is how to write the dual LP for a primal LP given in standard form.

We could do a similar derivation for the other standard form of an LP. Consider the primal

$$(P) \quad \begin{array}{ll} \min & \underline{c}^T \underline{x} \quad \text{over } \underline{x} \in \mathbb{R}^n \\ \text{subj. to} & A\,\underline{x} \geq \underline{b} \end{array}$$

The dual of this problem is:

$$(D) \quad \begin{array}{ll} \max & \underline{p}^T \underline{b}, \quad \text{over } \underline{p} \in \mathbb{R}^m \\ \text{subj. to} & \underline{p}^T A = \underline{c}^T, \\ & \underline{p} \geq 0 \end{array}$$

More generally, we can write down the dual of an LP by inspection, even if it is not in standard form.

There is a dual variable p_i associated with each primal constraint $\underline{a}_i^T \underline{x} = (\geq, \leq)\ \underline{b}_i$.

Depending on the form of the constraint, the dual variable p_i is either free, or constrained to be ≥ 0 or ≤ 0 respectively.

Likewise, there is a dual constraint $\underline{p}^T A_j = c_j \ (\geq c_j, \leq c_j)$ associated with each primal variable x_j, depending on whether it is free or constrained to be ≤ 0 or ≥ 0 respectively.

The following summary from the book of Bertsimas & Tsitsiklis captures all variants.

In the preceding example, we started with the equality constraint $\mathbf{Ax} = \mathbf{b}$ and we ended up with no constraints on the sign of the price vector $\mathbf{p}$. If the primal problem had instead inequality constraints of the form $\mathbf{Ax} \geq \mathbf{b}$, they could be replaced by $\mathbf{Ax} - \mathbf{s} = \mathbf{b}$, $\mathbf{s} \geq \mathbf{0}$. The equality constraint can be written in the form

$$[\mathbf{A} \mid -\mathbf{I}] \begin{bmatrix} \mathbf{x} \\ \mathbf{s} \end{bmatrix} = \mathbf{0},$$

which leads to the dual constraints

$$\mathbf{p}'[\mathbf{A} \mid -\mathbf{I}] \leq [\mathbf{c}' \mid \mathbf{0}'],$$

or, equivalently,

$$\mathbf{p}'\mathbf{A} \leq \mathbf{c}', \quad \mathbf{p} \geq \mathbf{0}.$$

Also, if the vector $\mathbf{x}$ is free rather than sign-constrained, we use the fact

$$\min_{\mathbf{x}}(\mathbf{c}' - \mathbf{p}'\mathbf{A})\mathbf{x} = \begin{cases} 0, & \text{if } \mathbf{c}' - \mathbf{p}'\mathbf{A} = \mathbf{0}', \\ -\infty, & \text{otherwise}, \end{cases}$$

to end up with the constraint $\mathbf{p}'\mathbf{A} = \mathbf{c}'$ in the dual problem. These considerations motivate the general form of the dual problem which we introduce in the next section.

In summary, the construction of the dual of a primal minimization problem can be viewed as follows. We have a vector of parameters (dual variables) $\mathbf{p}$, and for every $\mathbf{p}$ we have a method for obtaining a lower bound on the optimal primal cost. The dual problem is a maximization problem that looks for the tightest such lower bound. For some vectors $\mathbf{p}$, the corresponding lower bound is equal to $-\infty$, and does not carry any useful information. Thus, we only need to maximize over those $\mathbf{p}$ that lead to nontrivial lower bounds, and this is what gives rise to the dual constraints.

4.2 The dual problem

Let $\mathbf{A}$ be a matrix with rows $\mathbf{a}_i'$ and columns $\mathbf{A}_j$. Given a *primal* problem with the structure shown on the left, its *dual* is defined to be the maximization problem shown on the right:

$$\begin{array}{lll}
\text{minimize} & \mathbf{c}'\mathbf{x} & \\
\text{subject to} & \mathbf{a}_i'\mathbf{x} \geq b_i, & i \in M_1, \\
& \mathbf{a}_i'\mathbf{x} \leq b_i, & i \in M_2, \\
& \mathbf{a}_i'\mathbf{x} = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3,
\end{array}
\qquad
\begin{array}{lll}
\text{maximize} & \mathbf{p}'\mathbf{b} & \\
\text{subject to} & p_i \geq 0, & i \in M_1, \\
& p_i \leq 0, & i \in M_2, \\
& p_i \text{ free}, & i \in M_3, \\
& \mathbf{p}'\mathbf{A}_j \leq c_j, & j \in N_1, \\
& \mathbf{p}'\mathbf{A}_j \geq c_j, & j \in N_2, \\
& \mathbf{p}'\mathbf{A}_j = c_j, & j \in N_3.
\end{array}$$

Notice that for each constraint in the primal (other than the sign constraints), we introduce a variable in the dual problem; for each variable in the primal, we introduce a constraint in the dual. Depending on whether the primal constraint is an equality or inequality constraint, the corresponding dual variable is either free or sign-constrained, respectively. In addition, depending on whether a variable in the primal problem is free or sign-constrained, we have an equality or inequality constraint, respectively, in the dual problem. We summarize these relations in Table 4.1.

PRIMAL	minimize	maximize	DUAL
constraints	$\geq b_i$	≥ 0	**variables**
	$\leq b_i$	≤ 0	
	$= b_i$	free	
variables	≥ 0	$\leq c_j$	**constraints**
	≤ 0	$\geq c_j$	
	free	$= c_j$	

Table 4.1: Relation between primal and dual variables and constraints.

If we start with a maximization problem, we can always convert it into an equivalent minimization problem, and then form its dual according to the rules we have described. However, to avoid confusion, we will adhere to the convention that the primal is a minimization problem, and its dual is a maximization problem. Finally, we will keep referring to the objective function in the dual problem as a "cost" that is being maximized.

A problem and its dual can be stated more compactly, in matrix notation, if a particular form is assumed for the primal. We have, for example, the following pairs of primal and dual problems:

$$\begin{array}{ll}
\text{minimize} & \mathbf{c}'\mathbf{x} \\
\text{subject to} & \mathbf{Ax} = \mathbf{b} \\
& \mathbf{x} \geq \mathbf{0},
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & \mathbf{p}'\mathbf{b} \\
\text{subject to} & \mathbf{p}'\mathbf{A} \leq \mathbf{c}',
\end{array}$$

and

$$\begin{array}{ll}
\text{minimize} & \mathbf{c}'\mathbf{x} \\
\text{subject to} & \mathbf{Ax} \geq \mathbf{b},
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & \mathbf{p}'\mathbf{b} \\
\text{subject to} & \mathbf{p}'\mathbf{A} = \mathbf{c}' \\
& \mathbf{p} \geq \mathbf{0}.
\end{array}$$

Example 4.1 Consider the primal problem shown on the left and its dual shown

We argued for the standard form primal problem that the dual yields a lower bound on its value. This is more generally true when the primal has a mix of equality and inequality constraints, and where some or all of the primal variables may have sign constraints.

Theorem 1 (Weak duality)

If $\underline{x}$ is a feasible solution to the primal problem (i.e. satisfies all constraints), and $\underline{p}$ is a feasible solution to the dual, then

$$\underline{p}^T \underline{b} \leq \underline{c}^T \underline{x}$$

(It is assumed here that primal is a minimisation problem & the dual a maximisation problem.)

Proof: For any $\underline{x} \in \mathbb{R}^n$ & $\underline{p} \in \mathbb{R}^m$, define

$$u_i = p_i(\underline{a}_i^T \underline{x} - b_i),$$

$$v_j = (c_j - \underline{p}^T \underline{A}_j) x_j,$$

where $\underline{a}_i^T$ is the i^{th} row and $\underline{A}_j$ the j^{th} column of A.

Suppose $\underline{x}$ and $\underline{p}$ are primal & dual feasible respectively. The definition of the dual means p_i & $\underline{a}_i^T \underline{x} - b_i$ have the same sign, as do $c_j - \underline{p}^T A_j$ & x_j. Hence $u_i \geq 0$ & $v_j \geq 0$ $\forall i, j$.

Proof ctd. :

Now, $\Sigma u_i = \underline{p}^T A \underline{x} - \underline{p}^T \underline{b}$,
and $\Sigma v_j = \underline{c}^T \underline{x} - \underline{p}^T A \underline{x}$,

so we obtain that

$$0 \leq \Sigma u_i + \Sigma v_j = \underline{c}^T \underline{x} - \underline{p}^T \underline{b}.$$

This proves the result if the primal & dual are both feasible.

If the primal is infeasible, its value is defined to be $+\infty$, so the theorem holds trivially. Likewise, if the dual is infeasible, its value is defined to be $-\infty$, and the theorem is trivial.

#

Corollary

(a) If the primal is unbounded, i.e. has value $-\infty$, then the dual must be infeasible.

(b) If the dual is unbounded, i.e. has value $+\infty$, the the primal must be infeasible.

Theorem : The dual of the dual is the primal.

The main result of duality theory in LP is that the inequality between the values of the primal & dual problems holds with equality.

<u>Theorem</u> <u>3</u> (Strong duality) :

If an LP problem has an optimal solution, then so does its dual, and the respective optimal costs are equal.

<u>Remark</u> : This is quite a deep result & we won't give a proof. One method of proof relies on the fact that the simplex method is guaranteed to converge to a solution, if one exists. Another method is geometrical, and uses the separating Hyperplane Theorem (and generalises to Convex Programming).

The reader is referred to the book of Bertsimas & Tsitsiklis for further details.

We end with the following result, which provides a "certificate of optimality".

<u>Corollary</u> <u>of</u> <u>Theorem</u> <u>1</u> : If $\underline{x}$ is primal feasible and $\underline{p}$ is dual feasible and

$$\underline{c}^T \underline{x} = \underline{p}^T \underline{b},$$

then $\underline{x}$ is primal optimal & $\underline{p}$ is dual optimal.

Another way to say the same thing is the following.

Theorem 4 (Complementary slackness)

Let $\underline{x}$ and $\underline{p}$ be primal & dual feasible. Then $\underline{x}$ and $\underline{p}$ are optimal if and only if:

$$p_i \left(\underline{a}_i^T \underline{x} - b_i \right) = 0 \quad \forall\, i$$

$$\left(c_j - \underline{p}^T A_j \right) x_j = 0 \quad \forall\, j$$