Mathematical Programming	**Lecture 2**
OR 630 Fall 2005	**August 30, 2005**

Last time we considered

$$\begin{array}{rrcl} \max_y & b^T y & & \\ & A^T y & \leq & c, \end{array} \tag{1}$$

the problem of primary interest, and were led to the related problem

$$\begin{array}{rrcl} \min_x & c^T x & & \\ & Ax & = & b, \\ & x & \geq & 0. \end{array} \tag{2}$$

We call this the *dual* of the first problem, and to be precise, call the first problem the *primal*. What if we had started with the second problem? An equivalent problem (equivalent in the sense that the feasible regions and optimal sets are the same, while the optimal values are of opposite signs) is

$$\begin{array}{rrcl} \max\limits_x & (-c)^T x & & \\ & Ax & \leq & b, \\ & -Ax & \leq & -b, \\ & -x & \leq & 0, \end{array}$$

which is of the first form, and applying our usual rules its dual would be

$$\begin{array}{rrcrcrcl} \min_{s,t,u} & b^T s & - & b^T t & & & & \\ & A^T s & - & A^T t & - & u & = & -c \\ & s \geq 0, & & t \geq 0, & & u \geq 0, & & \end{array}$$

or

$$\begin{array}{rrcrcl} \min_{s,t,u} & -b^T(t-s) & & & & \\ & A^T(t-s) & + & u & = & c \\ & & & s,t,u & \geq & 0. \end{array}$$

Now, even if s and t have nonnegative components, $t-s$ can have components of any sign; conversely, any $y \in \mathbf{R}^m$ can be written as $t-s$ with s and t nonnegative, e.g., by setting $t = y_+ := \max\{y, 0\}$ and $s = (-y)_+ = \max\{-y, 0\}$, with "max" interpreted componentwise. Further, $A^T y + u = c$ for some nonnegative u iff (if and only if) $A^T y \leq c$; so this dual problem can be rewritten as

$$\begin{array}{rrcl} \min\limits_y & -b^T y & & \\ & A^T y & \leq & c, \end{array}$$

and, switching the sign of the objective function yet again, we get the "equivalent" problem (1) again. So, at least modulo these equivalences, the *dual of the dual is the primal.*

More intuitively, suppose we wish to obtain lower bounds for the optimal value of (2). Because we have equality constraints $Ax = b$, we can allow arbitrary signs for the components of y and deduce that

$$(A^T y)^T x = y^T Ax = y^T b = b^T y$$

for any feasible solution x to (2). We might require that $A^T y = c$, but this is overly limiting (especially as A is $m \times n$ with typically $m < n$: there may be no solutions to these equations!). Instead, since feasible x's are nonnegative, we can instead ask just that $A^T y \leq c$: then

$$c^T x \geq (A^T y)^T x = y^T Ax = y^T b = b^T y$$

for any feasible solution x to (2) and any y with $A^T y \leq c$. Hence the best such bound is optained from a solution to (1). Notice that the chain of inequalities above is just the reverse of that used in proving the weak duality result last time. The result is the same, but the interpretation is different.

More generally, suppose we have a minimization problem with some equations and some greater-than-or-equal-to constraints, and some nonnegative and some unrestricted variables:

$$(P) \qquad \begin{array}{lcccccc} \min_{x,w} & c_x^T x & + & c_w^T w & & \\ & A_x x & + & A_w w & = & b, \\ & \bar{A}_x x & + & \bar{A}_w w & \geq & \bar{b} \\ & x \geq 0, & & w \text{ unrestricted.} & & \end{array}$$

(We could also allow less-than-or-equal-to constraints and/or nonpositive variables, but these are easily converted to greater-than-or-equal-to constraints and nonnegative variables by multiplying by -1.)

To get a lower bound on the objective function values of feasible solutions to (P), we can take arbitrary multiples of the equality constraints and nonnegative multiples of the greater-than-or-equal-to constraints to get

$$(A_x^T y + \bar{A}_x^T z)^T x + ((A_w^T y + \bar{A}_w^T z)^T w \geq b^T y + \bar{b}^T z$$

whenever $z \geq 0$. Now $x \geq 0$, so we only need $A_x^T y + \bar{A}_x^T z \leq c_x$, but w is unrestricted, so we need $A_w^T y + \bar{A}_w^T z = c_w$. So, whenever (y, z) satisfies

$$\begin{array}{ccccc} A_x^T y & + & \bar{A}_x^T z & \leq & c_x, \\ A_w^T y & + & \bar{A}_w^T z & = & c_w, \\ & & z & \geq & 0, \end{array}$$

we have

$$c_x^T x + c_w^T w \geq b^T y + \bar{b}^T z$$

for every feasible (x, w) to (P). To obtain the best possible bound, we choose (y, z) to maximize the right-hand side above, and are led to

$$(D) \qquad \begin{array}{lccccc} \max_{y,z} & b^T y & + & \bar{b}^T z & & \\ & A_x^T y & + & \bar{A}_x^T z & \leq & c_x, \\ & A_w^T y & + & \bar{A}_w^T z & = & c_w, \\ & & & z & \geq & 0, \end{array}$$

which we **define** to be the dual of (P). Notice that for a primal minimization problem, we have the following correspondences:

- the dual problem is a maximization problem;
- for every nonnegative primal variable there is a less-than-or-equal-to dual constraint;
- for every unrestricted primal variable there is an equality dual constraint;
- for every equality primal constraint there is an unrestricted dual variable;
- for every greater-than-or-equal-to primal constraint there is an nonnegative dual variable;
- the objective function coefficients of the dual come from the right-hand sides of the primal constraints;
- the right-hand sides of the dual constraints come from the objective function coefficients of the primal; and
- the coefficient matrix of the dual constraints is the transpose of that of the primal constraints.

We have proved above

Theorem 1 *(Weak Duality) For every feasible solution (x, w) for (P) and every feasible solution (y, z) for (D), we have*

$$c_x^T x + c_w^T w \geq b^T y + \bar{b}^T z.$$

□

We could instead start with a maximization problem with some less-than-or-equal-to constraints and some equality constraints, and some unrestricted variables and some nonnegative variables. We would view this as our primal problem, but it would have exactly the form of (D) above. To obtain an upper bound on the objective function values of feasible solutions, we would introduce nonnegative multipliers (x) for the less-than-or-equal-to constraints and unrestricted multipliers (w) for the equality constraints. Proceeding exactly as above, we would be led to constraints on these multipliers, and to obtain the best possible bound we would arrive at a minimization problem, which would be exactly (P) above. So we would then **define** (P) to be the dual of (D). As an immediate consequence of these definitions, we would find

The dual of the dual is the primal.

Note that, when the primal problem is a maximization problem, we have the following correspondences:

- the dual problem is a minimization problem;
- for every unrestricted primal variable there is an equality dual constraint;
- for every nonnegative primal variable there is a greater-than-or-equal-to dual constraint;

- for every less-than-or-equal-to primal constraint there is an nonnegative dual variable;
- for every equality primal constraint there is an unrestricted dual variable;
- the objective function coefficients of the dual come from the right-hand sides of the primal constraints;
- the right-hand sides of the dual constraints come from the objective function coefficients of the primal; and
- the coefficient matrix of the dual constraints is the transpose of that of the primal constraints.

Let us consider an example of such a "mixed" linear programming problem.

Robust regression: In fitting problems, we choose the parameters of a model to best fit some data. Most common is linear least squares (LLS), where w is chosen to minimize $\|Aw - b\|_2$, the Euclidean norm of the residuals: if the model was a perfect fit, the parameters would satisfy $Aw = b$, but because of noise, incorrect model, measurement errors, etc., this is not possible. For the same reason, A is usually $m \times n$ with $m > n$, so this is an overdetermined linear system. The solution to the LLS problem can be written down in closed form, and there are efficient and accurate ways to compute the solution.

Robust regression uses a different objective function to minimize: note that LLS squares the residuals, so is sensitive to outliers, which may be undesirable if there are possible large measurement errors. Here we consider L_1-regression. The L_1-norm (or 1-norm) of a vector is the sum of the absolute values of its components: $\|v\|_1 := \sum_j |v_j|$. Note that this is a piecwise-linear function — it is linear in each orthant, but the derivative jumps when you cross any coordinate hyperplane.

So we want to minimize $\|Aw - b\|_1$. To do this, we use a clever modelling trick of linear programming: the absolute value of any real number σ can be represented as the minimum of $\tau + \upsilon$ over all nonnegative τ and υ with $\tau - \upsilon = \sigma$ Note the similarity to how the unrestricted variable y was represented by the difference of the nonnegative variables t and s above. If we do this trick with every component of the residual $Aw - b$, we write this as $x_+ - x_-$, say, with x_+ and x_- nonnegative vectors (of dimension m), and minimize the sum of all the x variables. Using e to denote a vector of ones of appropriate dimension, here m, we arrive at

$$
\begin{array}{llcccccc}
 & \min_{w,x_+,x_-} & & & e^T x_+ & + & e^T x_- & & \\
(P) & & Aw & - & x_+ & + & x_- & = & b, \\
 & & & & x_+, & & x_- & \geq & 0.
\end{array}
$$

This has the same optimal value and the same optimal w's as the original problem $\min\{\|Aw - b\|_1\}$.

Problem (P) is a minimization problem with equality constraints and both unrestricted and nonnegative variables. Its dual will be a maximization problem with unrestricted variables y

subject to both equality and inequality constraints. Using our rules, we obtain

$$
(D) \qquad \begin{array}{rrcl} \max_y & b^T y & & \\ & A^T y & = & 0, \\ & -y & \leq & e, \\ & y & \leq & e. \end{array}
$$

The feasible region of the dual is the intersection of a subspace $\{y \in \mathbf{R}^m : A^T y = 0\}$ and the set $\{y \in \mathbf{R}^m : -e \leq y \leq e\} = \{y \in \mathbf{R}^m : \|y\|_\infty \leq 1\}$, where the L_∞-norm (or ∞-norm) of a vector is the largest absolute value of its components. This norm is dual to the 1-norm in the sense that

$$
\|v\|_\infty = \max\{u^T v : \|u\|_1 \leq 1\} \text{ and } \|u\|_1 = \max\{u^T v : \|v\|_\infty \leq 1\}.
$$

Note that the dual LP problem to the LP problem corresponding to minimizing the 1-norm involves constraints on the dual norm! Also, the primal tries to express b as a linear combination of the a_j's, while the dual looks for a vector orthogonal to all the a_j's but with a large inner product with b.

Let us look at a particular instance:

Example 1 *Let* $A = \begin{bmatrix} 4 & -2 \\ 1 & -3 \\ 3 & -4 \end{bmatrix}$ *and* $b = \begin{pmatrix} 6 \\ 3 \\ 9 \end{pmatrix}$. *We would like* $w = (w_1; w_2)$ *to lie on three lines, but they don't meet in a point. So we want the best* L_1 *fit.*

The corresponding primal problem is

$$
(P) \qquad \begin{array}{lrcrrrrrrrcl} \min & & & & x_1 & +x_2 & +x_3 & +x_4 & +x_5 & +x_6 & & \\ & 4w_1 & - & 2w_2 & -x_1 & & & +x_4 & & & = & 6, \\ & w_1 & - & 3w_2 & & -x_2 & & & +x_5 & & = & 3, \\ & 3w_1 & - & 4w_2 & & & -x_3 & & & +x_6 & = & 9, \\ & & & & & & & & & x & \geq & 0, \end{array}
$$

with dual

$$
(D) \qquad \begin{array}{lccccccl} \max & 6y_1 & + & 3y_2 & + & 9y_3 & & \\ & 4y_1 & + & y_2 & + & 3y_3 & = & 0, \\ & -2y_1 & - & 3y_2 & - & 4y_3 & = & 0, \\ & -1 \leq y_1 \leq 1, & & -1 \leq y_2 \leq 1, & & -1 \leq y_3 \leq 1. & & \end{array}
$$

Here it is easier to picture the dual problem. The two equality constraints each define planes in 3-space, and their intersection is the line of all y *with* $y_1 : y_2 : y_3 = 1 : 2 : -2$. *The other constraints define a cube of side 2 centered at the origin, so the feasible region is the line segment from* $(-1/2; -1; 1)$ *to* $(1/2; 1; -1)$. *It is then easy to see that* $y = (-1/2; -1; 1)$ *is optimal with dual objective function value 3. To check this, note that* $w = (1; -1)$ *with* $x = (0; 1; 0; 0; 0; 2)$ *is feasible in* (P), *with primal objective function value 3. By our sufficient conditions for optimality (Corollary 1 last time), we know both are optimal.*

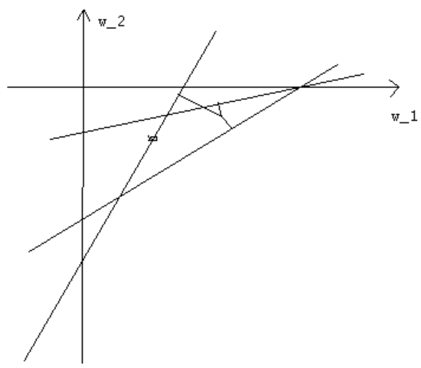

Figure 1: *The three "fitting" lines for Example 1.*

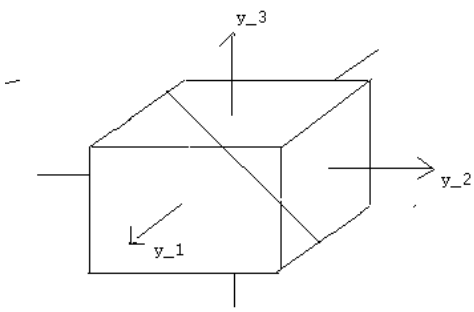

Figure 2: *The feasible region for the dual for Example 1.*

Here are representations, in w- and y-space, of these problems. Note that (D) can be viewed, in w-space, as finding a point where appropriate forces towards the three lines are in equilibrium.

Note that the line of points satisfying the equality constraints of (D) has by coincidence hit two of the bounding hyperplanes simultaneously at each end, so these two solutions are **degenerate**. Also, there are several optimal solutions to (P): as well as $w = (1; -1)$ there are $(6/5; -3/5)$ and $(3/5; -9/5)$ and all points in between.

Here are two special forms of dual problems we will be considering. First, the minimization problem (2) with (all) equality constraints in (all) nonnegative variables is called the LP problem in **standard form**. Second, if we use (all) inequality constraints in (all) nonnegative variables, we get the **symmetric dual problems**

$$\min\{c^T x : Ax \geq b, x \geq 0\} \text{ and } \max\{b^T y : A^T y \leq c, y \geq 0\}.$$