10-725: Optimization **Fall 2013**

Lecture 12: October 3

Lecturer: Ryan Tibshirani *Scribes: Milad Memarzadeh, Xinghai Hu*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

12.1 Duality

Duality for LP came before the general construct for duality. Suppose we are trying to lower bound an LP $B = \min_x f(x)$. For example:

$$\min_{x,y} px + qy$$
$$\text{subject to} \begin{cases} x + y \geq 2 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

Introduce $a, b, c \geq 0$, then we have: $ax + ay + bx + cy \geq 2a$.
Then we can see that $2a$ is the lower bound if x and y are feasible for the LP. So we choose a, b, c such that $a + b = p$ and $a + c = q$ and the lower bound would be $2a$.

So far we have found the lower bound for the above mentioned LP problem. Now we want to find the best lower bound:

$$\min 2a$$
$$\text{subject to} \begin{cases} a + b = p \\ a + c = q \\ a, b, c \geq 0 \end{cases}$$

This problem is called the dual LP for the primal problem that we had. This was an example of finding the dual problem from primal. Now consider the general case as follow,

$$\min_{x \in \mathbb{R}^n} c^T x$$
$$\text{subject to} \begin{cases} ax = b \\ Gx \leq h \end{cases}$$

As you can see we have equality and inequality constraints. Now we want to find the lower bound for this LP as discussed in the above example. for any u, $v \geq 0$, and x primal feasible we have:

$$u^T(Ax - b) + v^T(Gx - h) \leq 0 \text{ i.e.,}$$
$$(-A^T u - G^T v)^T x \geq -b^T u - h^T v$$

Therefore, we see that we get a lower bound on the primal optimal value. Now we would like to find the best lower bound,

$$\min_{u\in\mathbb{R}^m, v\in\mathbb{R}^r} b^T u - h^T v$$
$$\text{subject to} \begin{cases} -A^T u - G^T v = c \\ v \geq 0 \end{cases}$$

Which is the dual LP for the abovementioned general primal LP problem.

12.2 Max Flow and Min Cut

The maximum flow problem was first formulated in 1954 by T. E. Harris as a simplified model of Soviet railway traffic flow. the max-flow min-cut theorem states that in a flow network, the maximum amount of flow passing from the source (start node s) to the sink (end node t) is equal to the minimum capacity that, when removed in a specific way from the network, causes the situation that no flow can pass from the source to the sink.

Given graph $G = (V, E)$, define flow f_{ij}, $(i, j) \in E$ to satisfy:

$$f_{ij} \geq 0, (i, j) \in E$$
$$f_{i,j} \leq c_{ij}, (i, j) \in E \text{ where } c \text{ is capacity}$$
$$\sum_{(i,k)\in E} f_{ik} = \sum_{(k,j)\in E} f_{kj}, k \in V \neq \{s, t\}$$

The last line means that the input flow in any node other than start and end nodes is equal to output flow.

Max flow problem tries to find the flow that maximizes total value of flow from s to t, as an LP:

$$\min_{f\in\mathbb{R}^{|E|}} \sum_{(s,j)\in E} f_{sj}$$
$$\text{subject to} \begin{cases} f_{ij} \geq 0 \\ f_{ij} \leq c_{ij}, (i, j) \in E \\ \sum_{(i,k)\in E} f_{ik} = \sum_{(k,j)\in E} f_{kj}, k \in V \neq \{s, t\} \end{cases}$$

Introducing $a_{ij}, b_{ij} \geq 0, (i, j) \in E$, and $x_k, k \in V \neq \{s, t\}$ we will have the dual LP as follow:

$$\min_{b\in\mathbb{R}^{|E|}, x\in\mathbb{R}^{|V|}} \sum_{(i,j)\in E} b_{ij} c_{ij}$$
$$\text{subject to} \begin{cases} b_{ij} \geq x_i - x_j \\ b_{ij}, x_i, x_j \in \{0, 1\} \end{cases} \text{for all } i, j$$

Which is called the min cut problem. Therefore we know that value of max flow $\leq$ optimal value for dual LP min cut $\leq$ capacity of min cut. Famous result, called the max flow min cut theorem states that value of max flow through a network is exactly the capacity of the min cut. Hence the above inequalities are all equalities where the primal LP and dual LP have exactly the same optimal values which is called strong duality.

12.3 Another perspective on LP duality

The second perspective of LP duality helps define duality in general beyond LP.
A general LP problem is defined as follows.

$$\min_{x\in\mathbb{R}^n} c^T x$$
$$\text{subject to} \begin{cases} Ax = b \\ Gx \le h \end{cases}$$

For any feasible $x \in \mathbb{R}^n$ and $v \ge 0$, $c^T x \ge c^T x + u^T(Ax - b) + v^T(Gx - h) = L(x, u, v)$. Therefore,

$$f^*(x) = \min_{x\in C} c^T x \ge \min_{x\in C} L(x, u, v) \ge \min_{x\in\mathbb{R}^n} L(x, u, v) := g(u, v).$$

Here, $g(u, v) = \begin{cases} b^T u - h^T v, & c = -A^T u - G^T v \\ -\infty, & \text{otherwise} \end{cases}$.
This shows that, for any u and $v \ge 0$, $g(u, v)$ is the lower bound of $f^*(x)$. To find the tightest lower bound among all, we can solve the dual LP, which is defined as follows.

$$\max_{u\in\mathbb{R}^m, v\in\mathbb{R}^r} b^T u - h^T v$$
$$\text{subject to} \begin{cases} -A^T u - G^T v = c \\ v \ge 0 \end{cases}$$

12.4 Lagrangian Duality

The perspective on LP duality can go beyond LP. The general idea is that, we can use Lagrangian to find a lower bound for primal minimum. Then we follow similar steps above, if possible, to find the tightest lower bound.
The general minimization problem is formulated as follows.

$$\min_{x\in\mathbb{R}^n} f(x)$$
$$\text{subject to} \begin{cases} h_i(x) \le 0, & i = 1, 2, .., m \\ l_j(x) = 0, & j = 1, .., r \end{cases}$$

Lagrangian is defined by $L(x, u, v) = f(x) + \sum_{i=1}^{m} u_i h_i(x) + \sum_{j=1}^{r} v_j l_j(x)$, where $u \in \mathbb{R}^m, v \in \mathbb{R}^r, u \ge 0$.
Hence, for any feasible x, $f(x) \ge L(x, u, v)$, because $h_i(x) \le 0, l_j(x) = 0$. This means that, $L(x, u, v)$ is the point-wise lower bound of $f(x)$.
For example, Figure 12.4 shows an one-dimensional optimization problem, with objective being $f(x)$ (solid line) and constraint being $h(x)$(dashed line) ≤ 0. The feasible region of x is approximately $[-0.46, 0.46]$. Each dotted line represents a Lagrangian $L(x, u)$ for different choices of $u \ge 0$. They lie below the solid line within the feasible region.

Let C denote primal feasible set, f^* denote primal optimal value. Minimizing $L(x, u, v)$ over all $x \in \mathbb{R}^n$ gives a lower bound on f^*, because

$$f^*(x) \ge \min_{x\in C} L(x, u, v) \ge \min_{x\in\mathbb{R}^n} L(x, u, v) := g(u, v)$$

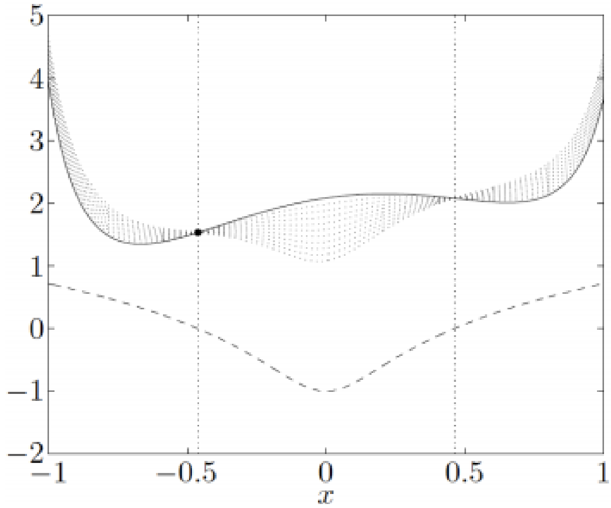

Figure 12.1: Example of Lagrangian Duality

Like LP dual function, $g(u, v)$ is called Lagrangian dual function. The domain of this function is that, $u \geq 0$ and $u \in \mathbb{R}^m, v \in \mathbb{R}^r$.
This relation holds in general for any optimization problem, while the equality may not hold for any pair of (x, u, v). For example, in Figure 12.4, dashed horizontal line is primal optimal f^*, and solid line shows dual function $g(\lambda)$. In this example, $g(\lambda)$ is always smaller than f^*. (Equality will hold for some pair of (x, u, v). Stay tuned for the next lecture.)

12.5 Quadratic Programming

Now we will use a quadratic program problem as an illustration of Lagrangian duality. Consider the following quadratic program.

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x \ (Q \succ 0)$$
$$\text{subject to} \begin{cases} Ax = b \\ x \geq 0 \end{cases}$$

Without any constraints, the optimal value reaches at point $x = -Q^{-1}c$.
Lagrangian of the problem is $L(x, u, v) = \frac{1}{2} x^T Q x + c^T + c^T x - u^T x + v^T (Ax - b)$, where $u \geq 0$. Thus, Lagrangian dual function is $g(u, v) = \min_{\mathbb{R}^n} L(x, u, v) = -\frac{1}{2}(c - u + A^T v)^T Q^{-1} (c - u + A^T v) - b^T v$. Here, $g(u, v)$ is lower bound of f^*.
When $Q \succeq 0$, Lagrangian dual function becomes:

$$g(u, v) = \begin{cases} -\frac{1}{2}(c - u + A^T v)^T Q^+ (c - u + A^T v) - b^T v c - u + A^T v & \text{if} (c - u + A^T v) \perp \text{null}(Q) \\ -\infty & \text{otherwise} \end{cases}$$

Here, Q^+ denotes generalized inverse of Q, and $g(u, v)$ is the non-trivial lower bound of f^* when $(c - u + A^T v) \perp \text{null}(Q)$.

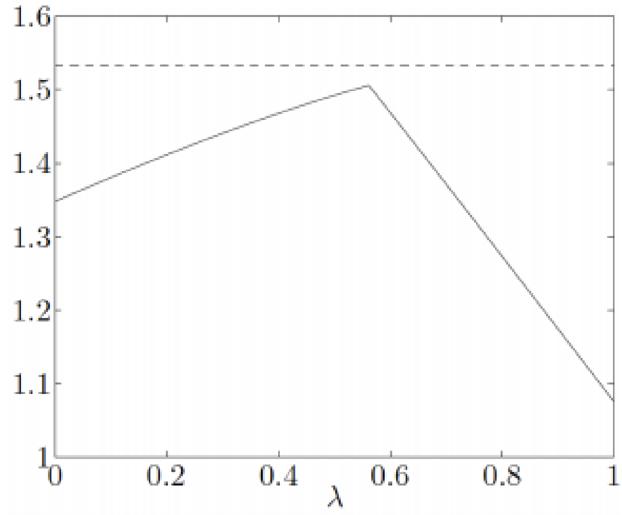

Figure 12.2: Lagrangian dual function.

The relation between primal and dual QP is shown in Figure 12.5. In this example, $f(x)$ is quadratic in 2 variables, with constraints $x \geq 0$. Lagrangian dual function $g(u)$ is also quadratic in 2 variables, subject to $u \geq 0$. As we can see, $g(u)$ provides a lower bound on f^* for any feasible u.

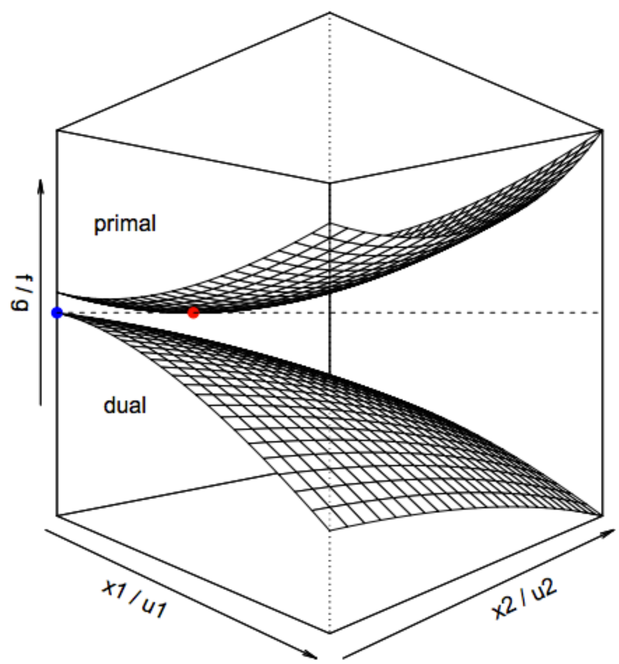

Figure 12.3: Primal and Dual QP in 2D.