EECS 127/227AT **Fall 2020**

Solutions for Discussion 7

1 LP duality in a combinatorial auction with divisible goods

An auctioneer is auctioning K divisible items. Here *divisible* means that each item can be broken up into arbitrary proportions. To be concrete, you can think of the items as being perfumes. The amount of each type of perfume the auctioneer has is 1.

The auctioneer has received B bids, where $B \geq 1$ is an integer. The bids are *combinatorial*. What this means is that each bid is for a subset of the items (i.e. the perfumes) and also states an amount which the bidder is willing to pay for that subset. Thus bid b, $1 \leq b \leq B$, will have the form (S_b, v_b), where $\emptyset \neq S_b \subseteq \{1, \ldots, K\}$, and $v_b \geq 0$.

It is assumed that each bidder will be willing to accept partial satisfaction of their bid. Namely, for any $0 \leq x_b \leq 1$, if the auctioneer gives the bidder b the amount x_b of each perfume in the set S_b, then the bidder will pay $x_b v_b$ for this. However, no bidder will accept any allocation except those that give them the items in the set S_b that they bid for in exactly equal amounts, and give them no other items.

Let A denote the $K \times B$ matrix with entries a_{ib} where

$$a_{ib} := \begin{cases} 1 & \text{if } i \in S_b, \\ 0 & \text{otherwise.} \end{cases}.$$

(a) The auctioneer wants to maximize revenue. Using the matrix A, pose the auctioneer's problem as a linear programming problem.

Solution:
The auctioneer gets to choose the proportions $x_b \in [0, 1]$, $1 \leq b \leq B$ of the individual bids to accept. The auctioneer cannot sell more of any item (i.e. perfume) than is available. These observations lead to the following LP formulation of the auctioneer's problem. Note that this is a maximization problem.

$$\begin{aligned} \max_{x} \quad & \sum_{b=1}^{B} v_b x_b \\ \text{subject to:} \quad & \sum_{b=1}^{B} a_{ib} x_b \leq 1, \; i = 1, \ldots, K, \\ & x_b \geq 0, \; b = 1, \ldots, B. \end{aligned}$$

Here we have observed that, because $S_b \neq \emptyset$, the constraints $x_b \leq 1$ are automatically satisfied by the requirement that the auctioneer cannot sell more than 1 unit of any item, so we have not explicitly introduced these constraints.

(b) Write down the dual LP in terms of dual variables y_i, $1 \leq i \leq K$.

Solution:
Since the primal LP is a maximization problem the dual LP will be a minimization problem. It reads:

$$
\begin{array}{ll}
\min\limits_{y} & \sum_{i=1}^{K} y_i \\
\text{subject to:} & \sum_{i=1}^{K} a_{ib} y_i \geq v_b, \; b = 1, \ldots, B, \\
 & y_i \geq 0, \; i = 1, \ldots, k.
\end{array}
$$

This looks like it has been pulled out of a hat, and indeed it has. You could derive this by going through the details of traditional Lagrange duality on the primal, introducing dual variables y_i, $1 \leq i \leq K$ for the respective inequality constraints $\sum_{b=1}^{B} a_{ib} x_b \leq 1$ in the primal and additional dual variables μ_b, $1 \leq b \leq B$ for the inequality constraints $x_b \geq 0$, which are then replaced, when the μ_b, $1 \leq b \leq B$ are eliminated, by the constraints $\sum_{i=1}^{K} a_{ib} y_i \geq v_b$, $1 \leq b \leq B$. However, LP is so well studied that people know exactly how this detailed calculation is going to work out and therefore just write down the eventual dual without going through this process.

Here is a simple recipe for LP duality. The notation is that of the book of Papadimitriou and Steiglitz on Combinatorial Optimization.

$\min c^T x$	$\max \pi^T b$
$a_i^T x = b_i, i \in M$	π_i unrestricted, $i \in M$
$a_i^T x \geq b_i, i \in \bar{M}$	$\pi_i \geq 0, i \in \bar{M}$
$x_j \geq 0, j \in N$	$\pi^T A_j \leq c_j, j \in N$
x_j unrestricted for $j \in \bar{N}$	$\pi^T A_j = c_j, j \in \bar{N}$

LP duality

Here A_j are the columns of the matrix whose rows are a_i^T. The problem on the left is a minimization problem, which is in duality with the problem on the right (the dual of the dual is the original problem). The variables for the problem on the left are x_j, $j \in N \cup \bar{N}$, and the variables for the problem on the right are π_i, $i \in M \cup \bar{M}$. c is a column vector with $|N \cup \bar{N}|$ coordinates, b is a column vector with $|M \cup \bar{M}|$ coordinates, and the entries of the a_i^T and the A_j are determined by a single matrix of size $|M \cup \bar{M}| \times |N \cup \bar{N}|$.

You can check for yourself that this recipe results in the stated dual from the given primal in the previous part of this question.

(c) Give an intuitive interpretation of the dual problem, based on strong duality and complementary slackness.

Solution:
y_i can be thought of as a price that the auctioneer can attribute to 1 unit of item i after having received the bids (i.e. the auctioneer is using the bids to determine what each item is really worth). For each $1 \leq b \leq B$ the constraint $\sum_{i=1}^{K} a_{ib} y_i \geq v_b$ corresponds to the fact that the bidder b is not going to bid more than the basket of goods S_b is worth. The optimal solution y_i^*, $1 \leq i \leq K$ of the dual problem then intuitively corresponds to the auctioneer determining how to set prices for the individual items (based on the received bids) so as not to overvalue the total stock.

Strong duality tells us that the total value $\sum_{b=1}^{B} x_b^* v_b$ received by the auctioneer at a primal optimal point $(x_b^*, 1 \leq b \leq B)$ is the same as the total value $\sum_{i=1}^{K} y_i^*$ attributed to the stock at a dual optimal point $(y_i^*, 1 \leq i \leq K)$. This may seem a bit strange because it could turn out that the stock is not entirely sold at a primal optimal point. This seeming discrepancy is resolved by complementary slackness.

One way to interpret complementary slackness is that it tells us that if any primal constraint is not active at a given primal optimal point, then the corresponding dual variable at any dual optimal point must equal 0. Here this means that for any item i, any primal optimal point $(x_b^*, 1 \leq b \leq B)$, and any dual optimal point $(y_i^*, 1 \leq i \leq K)$, if $\sum_{b=1}^{B} a_{ib} x_b^* < 1$, then we must have $y_i^* = 0$. Namely, the auctioneer must necessarily attribute zero price to any perfume that is not completely sold.

Another way to interpret complementary slackness (which is logically the same as the first way to interpret it) is that it tells us that if any dual variable is strictly positive at a given dual optimal point, then the corresponding inequality constraints must be active at any primal optimal point. Here this means that for any item i, any primal optimal point $(x_b^*, 1 \leq b \leq B)$, and any dual optimal point $(y_i^*, 1 \leq i \leq K)$, if $y_i^* > 0$, then we must have $\sum_{b=1}^{B} a_{ib} x_b^* = 1$. Namely, any perfume to which the auctioneer attributes a positive price must be completely sold.

Of course, it may turn out that there are perfumes to which the auctioneer attributes zero value at a dual optimal point $(y_i^*, 1 \leq i \leq K)$ which nevertheless get completely sold at a primal optimal point $(x_b^*, 1 \leq b \leq B)$. Namely, we could have items i such that $y_i^* = 0$ and $\sum_{b=1}^{B} a_{ib} x_b^* = 1$. You can check this for yourself in the following example. Suppose there are four perfumes, i.e. $K = 4$, and two bidders, i.e. $B = 2$. The first bidder bids $(\{1, 3, 4\}, 1)$ and the second bidder bids $(\{2, 3, 4\}, 1)$. Then $(x_1^*, x_2^*) = (\frac{1}{2}, \frac{1}{2})$ is a primal optimal point and $(y_1^*, y_2^*, y_3^*, y_4^*) = (0, 0, 1, 0)$ is a dual optimal point. We have $y_4^* = 0$, but nevertheless the constraint corresponding to the fourth perfume is active, i.e. we have $x_1^* + x_2^* = 1$.

It is also worth mentioning that, if the constraint corresponding to perfume i is the only one that is active at given primal optimal point x^*, then the component y_i^* of the dual optimal point y^* can be thought of as what is called the *shadow price* of the constraint i. This interpretation is prevalent in economics. In our example constraint i corresponds to the quantity of perfume i being limited to 1. The interprteation of the shadow price is that it is the implied price to the auctioneer of not having enough of perfume i, in the sense that, heuristically, if there were an infinitesimal amount $\epsilon > 0$ more of perfume i available, then the revenue of the auctioneer would have increased by ϵy_i^*. This interpretation is also appropriate in the case where the constraint corresponding to perfume i is not active, since in that case there is no advantage to having an infinitesimal amount more of the perfume because the auctioneer is not even able to sell the amount that already exists. However, this interpretation would be problematic if there were multiple constraints that were active. For instance, in the example with 4 perfumes and 2 bidders in the preceding paragraph, relaxing the constraint imposed by the limited quantity of perfume 3 does not change the feasible set, because perfume 4 continues to impose the constraint $x_1 + x_2 \leq 1$, so it is not possible to separate out the "shadow price" of the constraint imposed by perfume 3 from that imposed by perfume 4. This corresponds to the fact that both choices $(0, 0, 1, 0)$ and $(0, 0, 0, 1)$ for $(y_1^*, y_2^*, y_3^*, y_4^*)$ are optimal dual points - in the former it looks as if perfume 3 has a shadow price and in the latter it looks as if perfume 4 is the one with the shadow price – and in fact any $(0, 0, y_3^*, y_4^*)$ with nonnegative y_3^* and y_4^* satisfying $y_3^* + y_4^* = 1$ would be an optimal dual point. In spite of such conceptual difficulties, economists continue to treat Lagrange multipliers in all situations (even when there are multiple active constraints) as shadow prices (which is reasonable if the dual optimal point is uniquely defined).

The shadow price concept is most easily understood by considering the general form of our optimization problem with a single inequality constraint, i.e. the problem

$$\begin{array}{rl} \min\limits_{x} & f_0(x) \\ \text{subject to:} & f_1(x) \le c. \end{array}$$

Here $f_0, f_1 : \mathbb{R}^n \to \mathbb{R}$ and we will assume that they are both differentiable, with $\mathrm{dom}(f_0)$ and $\mathrm{dom}(f_1)$ both equal $\mathbb{R}^n$. The constraint is deliberately written as $f_1(x) \le c$ because what we are going to do is to study how the optimal primal value of our optimization problem varies when we vary c.

The Lagrangian is now

$$\mathcal{L}_c(x, \lambda) := f_0(x) + \lambda(f_1(x) - c),$$

where we have introduced a subscript on the Lagrangian to make it clear that we are dealing with different problems for different choices of c. The corresponding dual objective function will also then depend on c and will be given by

$$g_c(\lambda) := \min_{x} \mathcal{L}_c(x, \lambda),$$

and the corresponding dual problem will depend on c and be given by

$$\begin{array}{rl} \max\limits_{\lambda} & g_c(\lambda) \\ \text{subject to:} & \lambda \ge 0. \end{array}$$

For c in an open interval containing 0, assume there is a primal optimal point $x^*(c)$ where the inequality constraint is active and a dual optimal point $\lambda^*(c)$ for the problem with the constraint being defined by c. We also assume that $x^*(c)$ is differentiable in c in this open interval containing 0.

We can compute the derivative of the primal optimal value in c as

$$\frac{d}{dc} p^*(c) = \frac{d}{dc} f_0(x^*(c)) = \nabla_x^T f_0(x^*(c)) \frac{d}{dc} x^*(c),$$

this formula being valid for c in this open interval containing 0.

The Lagrangian stationarity condition at c gives

$$\begin{aligned} 0 &= \nabla_x \mathcal{L}_c(x, \lambda^*(c))|_{x=x^*(c)} \\ &= \nabla_x f_0(x^*(c)) + \lambda^*(c) \nabla_x f_1(x^*(c)). \end{aligned}$$

Since the constraint is active for all c in this interval containing 0, we have $c = f_1(x^*(c))$ throughout this interval. Differentiating this in c gives

$$1 = \frac{d}{dc} f_1(x^*(c)) = \nabla_x^T f_1(x^*(c)) \frac{d}{dc} x^*(c),$$

this formula being valid over the interval of c containing 0.

We therefore get

$$\frac{d}{dc}p^*(c) = \nabla_x^T f_0(x^*(c))\frac{d}{dc}x^*(c) = -\lambda^*(c)\nabla_x^T f_1(x^*(c))\frac{d}{dc}x^*(c) = -\lambda^*(c).$$

In particular

$$\frac{d}{dc}p^*(0) = -\lambda^*(0),$$

which yields the shadow price interpretation, since we are discussing a minimization problem in the general formulation. Increasing c corresponds to relaxing the constraint and results in a better (i.e. lower) optimal primal value and the Lagrange multiplier associated to the constraint has the interpretation of the rate of improvement of the optimal primal value in the parameter c defining the constraint.

2 LP relaxation of a Boolean LP

In a *Boolean linear program* the objective function and the constraint functions are affine, as in a LP, but the variables are constrained to take on the values 0 and 1. Despite the name, a Boolean linear program is *not* a linear program. In general it is not even a convex optimization problem, and is hard to solve. Here is a typical Boolean LP:

$$\begin{array}{ll} \min\limits_{x} & c^T x \\ \text{subject to:} & Ax \leq b, \\ & x_i \in \{0,1\},\ i = 1, \dots, n. \end{array}$$

Here $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$.

(a) Consider the following *LP relaxation* of the Boolean LP. The relaxation is an LP.

$$\begin{array}{ll} \min\limits_{x} & c^T x \\ \text{subject to:} & Ax \leq b, \\ & x_i \in [0,1],\ i = 1, \dots, n. \end{array}$$

Show that the optimal value of the LP relaxation is a lower bound for the optimal value of the original Boolean LP.

Solution:
The feasible set of the LP relaxation contains the feasible set of the Boolean LP. Since the problem is a minimization problem, the claim is obvious.

(b) Find the dual of the LP relaxation of the original Boolean LP. This should be expressed in terms of the dual variables for the constraints $Ax \leq b$ and the dual variables for the constraints $x \leq \mathbb{1}$, where $\mathbb{1}$ denotes the all-ones column vector.

Solution:
We first write a Lagrangian for the LP relaxation, namely

$$\begin{aligned} \mathcal{L}(x, \lambda, \alpha, \beta) &:= c^T x + \lambda^T (Ax - b) - \alpha^T x + \beta^T (x - \mathbb{1}) \\ &= (c + A^T \lambda - \alpha + \beta)^T x - b^T \lambda - \mathbb{1}^T \beta, \end{aligned}$$

where $\lambda \in \mathbb{R}^m$, $\alpha, \beta \in \mathbb{R}^n$ are the Lagrange multipliers corresponding to the three set of inequality constraints and $\mathbb{1}$ denotes the all-ones column vector of the appropriate dimension. The domain of the problem is $\mathcal{D} = \mathbb{R}^n$, and the Lagrangian is defined on $\mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$.

The dual objective function

$$g(\lambda, \alpha, \beta) := \inf_{x \in \mathcal{D}} \mathcal{L}(x, \lambda, \alpha, \beta),$$

can now be computed as

$$g(\lambda, \alpha, \beta) = \begin{cases} -b^T \lambda - \mathbb{1}^T \beta & \text{if } c + A^T \lambda - \alpha + \beta = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

We may therefore elminiate the dual variables α by setting this vector equal to $c + A^T\lambda + \beta$, while introducing the constraint $c+A^T\lambda+\beta \geq 0$ (interpreted coordinatewise). This gives the dual problem as:

$$\begin{aligned} \max_{\lambda,\beta} \quad & -b^T\lambda - \mathbb{1}^T\beta \\ \text{subject to:} \quad & \lambda \geq 0, \\ & \beta \geq 0, \\ & c + A^T\lambda + \beta \geq 0. \end{aligned}$$

Remark: In the solution to part (b) of the preceding question a recipe was provided for how to write down the dual LP without having to go through the detailed calculation via the Lagrangian as we have just done. You can check for yourself that this would have resulted in the same dual LP.

(c) Simplify the dual LP found in the preceding part of this question by eliminating the dual variables corresponding to the constraints $x \leq \mathbb{1}$.

Solution:
It is possible to eliminate β from the LP. For any fixed λ, the constraints on β are coordinate by coordinate and read

$$-\beta_i \leq \min(0, (c + A^T\lambda)_i), \ i = 1, \ldots, n.$$

Hence, with λ fixed, to maximize $-\mathbb{1}^T\beta$ we would simply set β_i to equal $-\min(0, (c + A^T\lambda)_i)$ for each $1 \leq i \leq n$. This gives the equivalent optimization problem:

$$\begin{aligned} \max_{\lambda} \quad & -b^T\lambda + \mathbb{1}^T\min(0, c + A^T\lambda) \\ \text{subject to:} \quad & \lambda \geq 0, \end{aligned}$$

where in the objective function the minimum is interpreted coordinatewise. As written, this is a convex optimization problem, but since the objective function is not affine this is not an LP (even though it is equivalent to an LP).

3 Lagrangian relaxation of a Boolean LP

We return to the Boolean LP considered in the preceding question, namely:

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to:} \quad & Ax \le b, \\ & x_i \in \{0,1\},\ i = 1, \dots, n, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$.

This can be rewritten in the equivalent form:

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to:} \quad & Ax \le b, \\ & x_i(1 - x_i) = 0,\ i = 1, \dots, n. \end{aligned}$$

Note that the Boolean constraint has now been expressed as a quadratic equality constraint. However this is *not* a convex optimization problem. The feasibility set has not changed, and is not convex in general. In the terminology of the textbook of Boyd and Vandenberghe this problem would not be called a QCQP, while in the terminology of the book of Calafiore and El Ghaoui this problem would be called a QCQP but one that is not a convex optimization problem.

(a) Nevertheless, we know how to subject even non-convex problems to traditional Lagrangian duality. Determine the dual of the Boolean LP expressed in the second form.

Remark: By virtue of weak duality, the value of this dual problem will give a lower bound to the value of the original Boolean LP. This method of finding a lower bound to the primal value of a given primal problem is called *Lagrangian relaxation*.

Solution:

We introduce the Lagrangian

$$\begin{aligned} \mathcal{L}(x, \lambda, \nu) &:= c^T x + \lambda^T (Ax - b) + \sum_{i=1}^n \nu_i x_i (1 - x_i). \\ &= (c + A^T \lambda + \nu)^T x - \lambda^T b - x^T \mathrm{diag}(\nu) x. \end{aligned}$$

Here $\lambda \in \mathbb{R}^m$ and $\nu \in \mathbb{R}^n$ are dual variables, and $\mathrm{diag}(\nu)$ denotes the $n \times n$ diagonal matrix with diagonal entries the ν_i in order. The domain of the problem is $\mathcal{D} = \mathbb{R}^n$, so the Lagrangian is defined on $\mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^n$.

We then need to find the dual objective function

$$g(\lambda, \nu) := \min_{x \in \mathcal{D}} \mathcal{L}(x, \lambda, \nu).$$

For this, we first observe that if $\nu_i > 0$ for any $1 \le i \le n$, then the quadratic term in $\mathcal{L}(x, \lambda, \nu)$ dominates (recall that λ and ν are fixed and we are considering $\mathcal{L}(x, \lambda, \nu)$ as a function of x for $x \in \mathcal{D} = \mathbb{R}^n$). This means $g(\lambda, \nu) = -\infty$. Hence, we may assume that $\nu_i \le 0$ for all $1 \le i \le n$.

We next observe that if $(c + A^T\lambda)_i \neq 0$ for any $1 \leq i \leq n$ such that $\nu_i = 0$, then again we get $g(\lambda, \nu) = -\infty$. Thus we may assume that

$$\nu_i = 0 \Longrightarrow (c + A^T\lambda)_i = 0, \; i = 1, \ldots n.$$

Computing

$$\nabla_x \mathcal{L}(x, \lambda, \nu) = c + A^T\lambda + \nu - 2\text{diag}(\nu)x,$$

we see that under these assumptions (i.e. the assumptions that $\nu_i \leq 0$ and $\nu_i = 0 \Longrightarrow (c+A^T\lambda)_i = 0$ for all $1 \leq i \leq n$) the optimal choice, in computing $g(\lambda, \nu)$, of x_i when $\nu_i < 0$ is $x_i = \frac{1}{2\nu_i}(c + A^T\lambda + \nu)_i$, and the choice does not matter when $\nu_i = 0$ and $(c + A^T\lambda)_i = 0$.

This finally gives

$$g(\lambda, \nu) = \begin{cases} -\lambda^T b + \frac{1}{4}\sum_{i=1,\nu_i<0}^n \frac{(c+A^T\lambda+\nu)_i^2}{\nu_i} & \text{if } \nu_i \leq 0, \;\; \forall i \text{ and } \nu_i = 0 \Longrightarrow (c + A^T\lambda)_i = 0, \;\; \forall i \,, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual optimization problem is therefore

$$\begin{aligned} \max_{\lambda,\nu} \quad & -\lambda^T b + \frac{1}{4}\phi(\lambda, \nu) \\ \text{subject to:} \quad & \lambda \geq 0 \\ & \nu \leq 0, \end{aligned}$$

where

$$\phi(\lambda, \nu) := \begin{cases} -\infty & \text{if } \nu_i > 0 \text{ for some } i, \\ -\infty & \text{if } \nu_i = 0 \text{ and } (c + A^T\lambda)_i \neq 0 \text{ for some } i, \\ \sum_{i=1,\nu_i<0}^n \frac{(c+A^T\lambda+\nu)_i^2}{\nu_i} & \text{otherwise.} \end{cases}$$

It can be checked that $\phi(\lambda, \nu)$ is a convex function on the convex set $\{\lambda \geq 0, \nu \leq 0\}$. A simple way to do this is to recognize that

$$\{\lambda \geq 0, \nu \leq 0 : \forall i \text{ we have } \nu_i = 0 \Longrightarrow (c + A^T\lambda)_i = 0\}$$

is a convex set and $\phi(\lambda, \nu)$ is finite and convex on this convex set.

Therefore, this is a convex optimization problem.

(b) Show that the dual optimization problem derived in part (a) of this question is equivalent to the dual of the LP relaxation of the Boolean LP, which was derived in part (b) of the previous question.

Hint: In part (c) of the preceding question we derived another optimization problem equivalent to the dual LP in part (b) of the preceding question.

Solution:
We notice that we can eliminate ν from the dual optimization problem formulated in the preceding part of this question, coordinate by coordinate. To see this, fix λ. For $1 \leq i \leq n$, if $(c + A^T\lambda)_i \geq 0$ then the maximum of $\frac{(c+A^T\lambda+\nu)_i^2}{\nu_i}$ over $\nu_i \leq 0$ occurs at $\nu_i = -(c + A^T\lambda)_i$ and equals 0. If

$(c + A^T\lambda)_i < 0$ then the maximum of $\frac{(c+A^T\lambda+\nu)_i^2}{\nu_i}$ over $\nu_i \leq 0$ occurs at $\nu_i = (c + A^T\lambda)_i$ and equals $4(c + A^T\lambda)_i$. Hence we see that for fixed λ the maximum of $\frac{1}{4}\phi(\lambda, \nu)$ over $\nu \leq 0$ equals $\sum_{i=1}^n \min(0, (c + A^T\lambda)_i)$. This tells us that the dual optimization problem of the preceding part of this question can be equivalently expressed as:

$$\begin{aligned} \max_{\lambda} \qquad & -b^T\lambda + \mathbb{1}^T \min(0, c + A^T\lambda) \\ \text{subject to:} \qquad & \lambda \geq 0, \end{aligned}$$

where in the objective function the minimum is interpreted coordinatewise. This is the same as the problem which we saw was equivalent to the dual of the LP relaxation of the original Boolean LP.