DS 574 Algorithmic Mechanism Design	**Lecture #12**
Fall 2022	Prof. Kira Goldner

Lagrangian Duality

We start with the basics surrounding formulating a partial Lagrangian primal and taking its dual, and understanding the properties of duality.

We begin with a standard maximization problem subject to constraints, which we call the full primal. The set $\mathcal{P}$ here denotes feasibility constraints, while x represents whatever our primal variables are.

Full primal:

$$\begin{aligned} \max \quad & f(x) \\ \text{s.t.} \quad & Ax \leq b \qquad\qquad\qquad \text{(dual variable } \lambda) \\ & x \in \mathcal{P} \end{aligned}$$

We denote the optimal solution to the full primal as x^*; that is, $x^* \in \text{argmax}_{Ax \leq b, x \in \mathcal{P}} f(x)$.

Partial Lagrangian Primal: We now form the partial Lagrangian primal by using the Lagrangian multiplier λ_i for each constraint of the form $(Ax)_i \leq b_i$ and moving it into the objective, where we now minimize over the multipliers λ. We leave all of the feasibility constraints as is, and define $\mathcal{L}(x; \lambda)$ as this new objective.

$$\max_{x \in \mathcal{P}} \min_{\lambda \geq 0} \quad \mathcal{L}(x; \lambda) \quad = \quad \max_{x \in \mathcal{P}} \min_{\lambda \geq 0} \quad f(x) + \lambda^T (b - Ax)$$

First, we observe that the (partial) Lagrangian Primal is indeed a *relaxation* of the full primal. For any feasible x, λ—that is, $Ax \leq b, x \in \mathcal{P}$, and $\lambda \geq 0$—then $f(x) \leq \mathcal{L}(x; \lambda)$.

Partial Lagrangian Dual: By reversing the order of the max and the min, we obtain the dual minimization problem. We notate this dual problem as $D(\lambda)$.

$$\min_{\lambda \geq 0} \quad D(\lambda) \quad = \quad \min_{\lambda \geq 0} \max_{x \in \mathcal{P}} \quad f(x) + \lambda^T (b - Ax)$$

We denote the optimal dual solution as $\lambda^* \in \text{argmin}_{\lambda \geq 0} D(\lambda)$.

Complementary Slackness: We say that x, λ satisfy complementary slackness if $\lambda_i > 0 \implies b_i - (Ax)_i = 0$.

Weak Duality. The value of the full primal is always upper-bounded by the value of the dual problem. Specifically, the value of the full primal is at most $f(x^*)$ by definition, and any feasible dual solution must satisfy $\lambda \geq 0$, so the dual objective is larger: $f(x^*) \leq D(\lambda)$.

Proof.

$$\begin{aligned}
f(x^*) &\leq f(x^*) + \lambda^T(b - Ax^*) && \lambda \geq 0, Ax^* \leq b \\
&\leq \max_{x \in \mathcal{P}} f(x) + \lambda^T(b - Ax) && x^* \in \mathcal{P} \\
&= D(\lambda)
\end{aligned}$$

□

Strong Duality. Strong duality implies that the value of the full primal is equal to the value of the Lagrangian primal, and this is equal to the value of the Lagrangrian dual, when they are all at their optimal solutions. However, strong duality is not a given. We see below that if strong duality holds, there must exist a pair of primal, dual solutions that are optimal. Further, if there exist an optimal pair, then strong duality must hold. Either condition is sufficient to show the other exists.

An Optimal Pair implies Strong Duality. For any choice of dual variables $\hat{\lambda}$, if there exists $\hat{x}$ that forms an optimal pair with $\hat{\lambda}$, that is, $\hat{x}$ such that:

1. $\hat{x} \in \text{argmax}_{x \in \mathcal{P}} \mathcal{L}(x; \hat{\lambda})$ ($\hat{x}$ is optimal)
2. $A\hat{x} \leq b$ ($\hat{x}$ satisfies the Lagrangified constraints)
3. $\hat{x}, \hat{\lambda}$ satisfy complementary slackness

then strong duality holds, that is, $D(\hat{\lambda}) = f(x^*)$.

Proof.

$$\begin{aligned}
D(\hat{\lambda}) &= \max_{x \in \mathcal{P}} \mathcal{L}(x, \hat{\lambda}) \\
&= f(\hat{x}) + \hat{\lambda}^*(b - A\hat{x}) && \text{by (1)} \\
&= f(\hat{x}) && \text{by (3)} \\
&\leq f(x^*) && \text{by (2), } x \in \mathcal{P}
\end{aligned}$$

□

Strong Duality implies an Optimal Pair. If strong duality holds, that is, $\min_{\lambda \geq 0} D(\lambda) = f(x^*)$, then there exists $\hat{x}$ such that

1. $\hat{x} \in \text{argmax}_x \mathcal{L}(x; \lambda^*)$
2. $A\hat{x} \leq b$
3. $\hat{x}, \lambda^*$ satisfy complementary slackness
4. $f(\hat{x}) = f(x^*)$.

Proof. From weak duality, we know that

$$\min_{\lambda \geq 0} D(\lambda) \quad = \quad D(\lambda^*) \quad \geq \quad \mathcal{L}(x^*, \lambda^*) \quad \geq \quad f(x^*).$$

These inequalities must all hold with equality for the premise to hold. The first inequality's tightness implies condition (1), and the second inequality's tightness implies condition (3). Condition (2) is true by the definition of x^*. $\square$

For further background on Lagrangian duality, see [2].

Maximizing Revenue

Now we will use this theory of duality to formulate the general Lagrangian linear program for revenue maximization. The following theory is due to Yang Cai, Nikhil Devanur, and Matt Weinberg [CDW '16]. We will let v_i be a vector which can be indexed for each item j. Similarly for allocation x at bidder i and item j. V_i represents the type space, or the support of the distribution F_i—the possible valuations that v_i can take.

We use $\varnothing$ to denote the type of not participating in the auction. Let $V_i^+ = V_i \cup \{\varnothing\}$. We use $\mathcal{P}$ to denote the polytope of feasible allocation rules.

Decision variables interim allocations $x_{ij}(v_i)$ and payments $p_i(v_i)$.

$$\begin{aligned}
\max \quad & \sum_{i=1}^{n} \sum_{v_i \in V_i} f_i(v_i) \cdot p_i(v_i) \\
\text{s.t.} \quad & x_i(v_i) \cdot v_i - p_i(v_i) \geq x_i(v_i') \cdot v_i - p_i(v_i') \qquad \forall i, v_i \in V_i, v_i' \in V_i^+ (\text{dual variable } \lambda_i(v_i, v_i')) \\
& x \in \mathcal{P}
\end{aligned}$$

Partial Lagrangian Primal:

$$\max_{x \in \mathcal{P}, p} \min_{\lambda \geq 0} \quad \mathcal{L}(\lambda; x, p)$$

where

$$\begin{aligned}
\mathcal{L}(\lambda; x, p) &= \sum_{i=1}^{n} \left(\sum_{v_i \in V_i} f_i(v_i) \cdot p_i(v_i) + \sum_{v_i \in V_i} \sum_{v_i' \in V_i^+} \lambda_i(v_i, v_i') \cdot \big(v_i \cdot \big(x(v_i) - x(v_i')\big) - \big(p_i(v_i) - p_i(v_i')\big)\big) \right) \\
&= \sum_{i=1}^{n} \sum_{v_i \in V_i} p_i(v_i) \left(f_i(v_i) + \sum_{v_i' \in V_i} \lambda_i(v_i', v_i) - \sum_{v_i' \in V_i^+} \lambda_i(v_i, v_i') \right) + \\
&\qquad \sum_{i=1}^{n} \sum_{v_i \in V_i} x_i(v_i) \left(\sum_{v_i' \in V_i^+} v_i \cdot \lambda_i(v_i, v_i') - \sum_{v_i' \in V_i} v_i' \cdot \lambda_i(v_i', v_i) \right)
\end{aligned}$$

Partial Lagrangian Dual:

$$\min_{\lambda \geq 0} \max_{x \in \mathcal{P}, p} \quad \mathcal{L}(\lambda; x, p)$$

For the dual to provide a useful (finite) upper bound we need $\max_{x \in \mathcal{P}, p} \mathcal{L}(\lambda; x, p) < \infty$. For this to be true, we must have the coefficient of $p_i(v_i)$ equal to 0, that is:

$$f_i(v_i) + \sum_{v_i' \in V_i} \lambda_i(v_i', v_i) = \sum_{v_i' \in V_i^+} \lambda_i(v_i, v_i').$$

We think of this as a "flow conservation" constraint in the following set-up. A dual solution λ is useful if and only if for each bidder i, λ_i forms a valid flow, i.e., if and only if the following satisfies flow conservation at all nodes except the source and the sink:

- Nodes: A super source s and a super sink $\varnothing$, along with a node v_i for every type $v_i \in V_i$.
- Flow from s to v_i of weight $f_i(v_i)$ for all $v_i \in V_i$.
- Flow from v to v' of weight $\lambda_i(v, v')$ for all $v \in V$ and $v' \in V^+$ (including the sink $\varnothing$).

Then

$$\sum_{i=1}^{n} \sum_{v_i \in V_i} f_i(v_i) p_i(v_i) \leq \sum_{i=1}^{n} \sum_{v_i \in V_i} f_i(v_i) \cdot x_i(v_i) \cdot \Phi_i^{\lambda}(v_i)$$

for

$$\Phi_i^{\lambda}(v_i) = v_i - \frac{1}{f_i(v_i)} \sum_{v_i' \in V_i} \lambda_i(v_i', v_i)(v_i' - v_i).$$

and this holds with equality if and only if x, p, λ are optimal solutions to the primal and dual respectively.

The Canonical Flow

The way that we will set the dual variables, which is in fact optimal in the single-dimensional setting, is as follows: $\lambda_i(v, v+1) = 1 - F_i(v+1) = \Pr_{v_i}[v_i > v+1]$. All other $\lambda_i(v, v') = 0$ except $\lambda_i(0, \varnothing) = 1$. Then

$$\Phi_i^{\lambda}(v_i) = \varphi_i(v_i)$$

is Myerson's virtual value.

Ironing

For a non-monotone interval $[L, H]$ in which $\Phi_i^{\lambda}(L) > \cdots > \Phi_i^{\lambda}(H)$, we augment the following dual variables until $\Phi_i^{\lambda}(L) = \cdots = \Phi_i^{\lambda}(H)$ by increasing $\lambda_i(v+1, v)$ and $\lambda_i(v, v+1)$ by ε for $v \in [L, H-1]$.

References

[1] Yang Cai, Nikhil R. Devanur, and S. Matthew Weinberg. A Duality Based Unified Approach to Bayesian Mechanism Design. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 926–939, New York, NY, USA, 2016. ACM.

[2] R Tyrrell Rockafellar. *Conjugate duality and optimization*, volume 16. Siam, 1974.

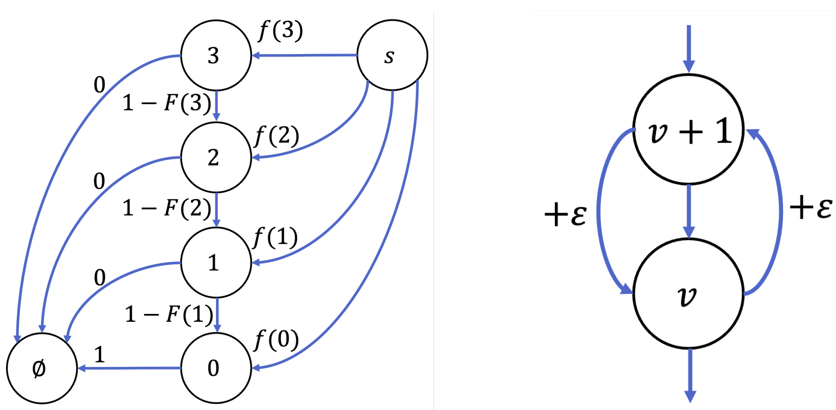

Figure 1: Left: The single-dimensional canonical dual resulting in Myersonian virtual values. Right: The process for ironing.