Lecture 4:
Lagrangian duality, part III: Discrete optimization

Ann-Brith Strömberg

2018-01-23

A reminder of nice properties in the (strictly) convex case

- Example I (explicit dual) ($\mathbf{x}^* = (2,2), \mu^* = 4, f^* = 8$)

$$
\begin{array}{rl}
f^* := \text{minimum} & f(\mathbf{x}) := x_1^2 + x_2^2 \\
\text{subject to} & g(\mathbf{x}) = -x_1 - x_2 + 4 \leq 0 \\
& x_1, x_2 \geq 0
\end{array}
$$

- Let $X := \left\{ \mathbf{x} \in \mathbb{R}^2 \,\middle|\, x_1,\ x_2 \geq 0 \right\} = \mathbb{R}^2_+$
- $L(\mathbf{x}, \mu) = x_1^2 + x_2^2 + \mu \cdot (-x_1 - x_2 + 4)$

... nice properties in the convex case

$$
\begin{aligned}
q(\mu) &:= \underset{\mathbf{x}\in X}{\text{minimum}} \left\{x_1^2 + x_2^2 + \mu \cdot (-x_1 - x_2 + 4)\right\} \\
&= 4\mu + \underset{\mathbf{x}\in X}{\text{minimum}} \left\{x_1^2 + x_2^2 - \mu x_1 - \mu x_2\right\} \\
&= 4\mu + \underset{x_1\geq 0}{\text{minimum}} \left\{x_1^2 - \mu x_1\right\} + \underset{x_2\geq 0}{\text{minimum}} \left\{x_2^2 - \mu x_2\right\}
\end{aligned}
$$

- For a fixed value of $\mu \geq 0$, the minimum of $L(\mathbf{x}, \mu)$ over $\mathbf{x} \in X$ is attained at $x_1(\mu) = \frac{\mu}{2}$, $x_2(\mu) = \frac{\mu}{2}$

$\Rightarrow$ $q(\mu) = L(\mathbf{x}(\mu), \mu) = \ldots = 4\mu - \frac{\mu^2}{2}$ for all $\mu \geq 0$

- The dual function q is concave and differentiable
- $f^* = f(\mathbf{x}^*) = 8 = q^*$
- $\mu^* = 4$, $\mathbf{x}(\mu^*) = \mathbf{x}^* = (2, 2)$

Weak duality! Strong duality?

- The primal optimal solution is obtained from the Lagrangian dual optimal solution under convexity and CQ
  (For a non-strictly convex function f we have to deal with the non-coordinability, though)
- What applies otherwise?

$$f(\mathbf{x}) \quad [\mathbf{x} \in X, \ \mathbf{g}(\mathbf{x}) \leq \mathbf{0}^m]$$

$$f^* - q^* = 0? \qquad \begin{array}{l} f^* = f(\mathbf{x}^*) \\ q^* = q(\boldsymbol{\mu}^*) \end{array}$$

$$q(\boldsymbol{\mu}) \quad [\boldsymbol{\mu} \geq \mathbf{0}^m]$$

- How do we generate optimal primal solutions in the case of a positive duality gap?

A first example with non-zero duality gap

- Example II ($\mathbf{x}^* = (0, 1, 1),\ f^* = 17$)

$$\begin{aligned} f^* := \text{minimum} \quad & f(\mathbf{x}) := 3x_1 + 7x_2 + 10x_3 \\ \text{subject to} \quad & x_1 + 3x_2 + 5x_3 \geq 7 \\ & x_j \in \{0, 1\}, \quad j = 1, 2, 3 \end{aligned}$$

- Let $X := \left\{ \mathbf{x} \in \mathbb{R}^3 \,\middle|\, x_j \in \{0, 1\},\ j = 1, 2, 3 \right\} = \mathbb{B}^3$
- Let $g(\mathbf{x}) := 7 - x_1 - 3x_2 - 5x_3$
- $L(\mathbf{x}, \mu) := 3x_1 + 7x_2 + 10x_3 + \mu \cdot (7 - x_1 - 3x_2 - 5x_3)$

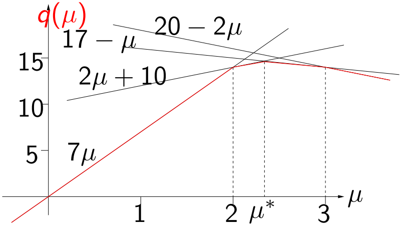
The dual function is computed according to

$$
\begin{aligned}
q(\mu) &:= 7\mu + \operatorname*{minimum}_{\mathbf{x} \in X} \left\{ (3-\mu)x_1 + (7-3\mu)x_2 + (10-5\mu)x_3 \right\} \\
&= 7\mu + \operatorname*{minimum}_{x_1 \in \{0,1\}} \left\{ (3-\mu)x_1 \right\} + \operatorname*{minimum}_{x_2 \in \{0,1\}} \left\{ (7-3\mu)x_2 \right\} \\
&\quad + \operatorname*{minimum}_{x_3 \in \{0,1\}} \left\{ (10-5\mu)x_3 \right\}
\end{aligned}
$$

- $X(\mu)$ is obtained by setting
  $x_j(\mu) = \begin{cases} 1 \\ 0 \end{cases}$ when the objective coefficient is $\begin{cases} \leq 0 \\ \geq 0 \end{cases}$

Subproblem solutions and the dual function

$\mu \in$	$x_1(\mu)$	$x_2(\mu)$	$x_3(\mu)$	$g(\mathbf{x}(\mu))$	$q(\mu)$
$[-\infty, 2]$	0	0	0	7	7μ
$[2, \frac{7}{3}]$	0	0	1	2	$2\mu + 10$
$[\frac{7}{3}, 3]$	0	1	1	-1	$-\mu + 17$
$[3, \infty]$	1	1	1	-2	$-2\mu + 20$



Properties of the dual function

- q is concave; it is non-differentiable at break points $\mu \in \{2, \frac{7}{3}, 3\}$

	$\mu < \mu^*$	$\mu = \mu^*$	$\mu > \mu^*$	
slope of q at μ	> 0	non-diff.	< 0	
$\mathbf{x}(\mu)$	infeasible	either	feasible	(w.r.t. "$g(\mathbf{x}) \leq 0$")

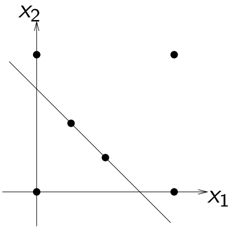
- Check that the slope equals the value of the constraint function
- The one-variable function q has a "derivative" which is non-increasing

  This is a property of every concave function of one variable
- $\mu^* = \frac{7}{3}$, $q^* = q(\mu^*) = \frac{44}{3} = 14\frac{2}{3}$
- Recall: $\mathbf{x}^* = (0, 1, 1)$, $f^* = 17$
- A positive duality gap!
- $X(\mu^*) = \big\{(0, 0, 1), (0, 1, 1)\big\} \ni \mathbf{x}^*$
- Not all solutions in $X(\mu^*)$ are feasible

Another example with non-zero duality gap

- Example III ($\mathbf{x}^* = (2,1),\ f^* = -3$)

$$\begin{aligned} f^* := \min \quad & f(\mathbf{x}) = -2x_1 + x_2 \\ \text{s.t.} \quad & x_1 + x_2 - 3 = 0, \\ & \mathbf{x} \in X = \{(0,0),(0,4),(4,4),(4,0),(1,2),(2,1)\} \end{aligned}$$



- $L(\mathbf{x}, \mu) = -3\mu + (-2+\mu)x_1 + (1+\mu)x_2$

Another example with non-zero duality gap, cont.

- Observe: $\mu \in \mathbb{R}$ (since the relaxed constraint is an equality: $g(\mathbf{x}) = 0$)
- $X(\mu) = \operatorname{argmin} \Big\{ (-2+\mu)x_1 + (1+\mu)x_2 \;\Big|\; \mathbf{x} \in \big\{(0,0),(0,4),(4,4),(4,0),(1,2),(2,1)\big\} \Big\}$

$$X(\mu) = \begin{cases} \{(4,4)\}, & \mu < -1 \\ \{(4,4),(4,0)\}, & \mu = -1 \\ \{(4,0)\}, & \mu \in (-1,2) \\ \{(4,0),(0,0)\}, & \mu = 2 \\ \{(0,0)\}, & \mu > 2 \end{cases} \; ; \; q(\mu) = \begin{cases} -4+5\mu, & \mu \le -1 \\ -8+\mu, & \mu \in [-1,2] \\ -3\mu, & \mu \ge 2 \end{cases}$$

- $\mu^* = 2$; $q^* = q(\mu^*) = -6 < f^* = -3$, $\mathbf{x}^* = (2,1) \notin X(\mu^*)$
- The set $X(\mu^*)$ does not even contain a feasible solution!

Strong duality—repetition

The following three statements are equivalent

(a) $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ is a saddle point to $L(\mathbf{x}, \boldsymbol{\mu})$

(b)
   i. $f(\mathbf{x}^*) + (\boldsymbol{\mu}^*)^{\mathrm{T}}\mathbf{g}(\mathbf{x}^*) = \min_{\mathbf{x}\in X} \left\{ f(\mathbf{x}) + (\boldsymbol{\mu}^*)^{\mathrm{T}}\mathbf{g}(\mathbf{x}) \right\}$
   $\left[\iff \mathbf{x}^* \in X(\boldsymbol{\mu}^*) \right]$

   ii. $(\boldsymbol{\mu}^*)^{\mathrm{T}}\mathbf{g}(\mathbf{x}^*) = 0$

   iii. $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}^m$

(c) $f^* = f(\mathbf{x}^*) = q(\boldsymbol{\mu}^*) = q^*$

A method for finding an optimal solution

1) Solve the Lagrangian dual problem $\Longrightarrow \boldsymbol{\mu}^*$

2) Find a vector $\mathbf{x}^* \in X$ which satisfies (b)

- When does this work? What if it doesn't?
- First the convex case (with zero duality gap)
- Even with a zero duality gap, it is not always trivial to find an optimal primal solution in this way.
- The latter since the set $X(\boldsymbol{\mu}^*)$ is normally *not explicitly available*—we normally compute only *one* element of the set $X(\boldsymbol{\mu})$

Method for finding an optimal solution, cont.

Example II from Lecture 3 (a 2-variable LP problem)

- Imagine using the simplex method for solving each LP subproblem. Then, we only get extreme points of X.
- Here, $\mathbf{x}^*$ is an extreme point of $X \cap \{\, \mathbf{x} \in \mathbb{R}^2 \mid g(\mathbf{x}) \leq 0 \,\}$ (since it is an LP) but *not* an extreme point of X

There are several ways to escape from this *non-coordinability*

The Dantzig–Wolfe (DW) decomposition method

Remember all the points $\mathbf{x}(\boldsymbol{\mu}^k) \in X(\boldsymbol{\mu}^k)$ visited (methodologically, this is "called column generation").
At the end, solve an LP to find the best point in their convex hull and which also feasible in the original problem.
Finite convergence.

Method for finding an optimal solution, cont.

Weighted averages of Lagrangean subproblem solutions

Construct a primal sequence as a *convex combination* of the points $\mathbf{x}(\boldsymbol{\mu}^k) \in X(\boldsymbol{\mu}^k)$ visited. We must not solve any extra optimization problems, and virtually no extra memory is needed (compare with DW). Convergence in the limit.

- This is elaborated in the papers by Gustavsson, Larsson, Patriksson, and Strömberg, and Önnheim (1999, 2014, and 2016; see the course homepage)
- It is also the topic of the next Lecture

Augmented Lagrangian methods

Introduce non-linear price functions for the constraints, instead of the linear one given by Lagrangian relaxation (not covered in this course)

Linear integer optimization: The strength of the Lagrangian relaxation

- Compare with a continuous (LP) relaxation:

$$
\begin{aligned}
v_{LP} := \min \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\
\text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\
& \mathbf{D}\mathbf{x} \leq \mathbf{d} \\
& \mathbf{x} \in \mathbb{R}^n_+
\end{aligned}
\qquad \leq \qquad
\begin{aligned}
v^* := \min \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\
\text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\
& \mathbf{D}\mathbf{x} \leq \mathbf{d} \\
& \mathbf{x} \in \mathbb{Z}^n_+
\end{aligned}
$$

- Let $X =: \left\{ \mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^K \right\}$ denotee the set of points in $X = \left\{ \mathbf{x} \in \mathbb{Z}^n_+ \;\middle|\; \mathbf{A}\mathbf{x} \leq \mathbf{b} \right\}$

The strength of the Lagrangian relaxation, cont.

- Lagrangian relax the constraints $\mathbf{Dx} \leq \mathbf{d}$:

$$v_L := \max_{\boldsymbol{\mu} \geq \mathbf{0}} \left(\min_{\mathbf{x} \in X} \left[\mathbf{c}^{\mathrm{T}} \mathbf{x} + \boldsymbol{\mu}^{\mathrm{T}} (\mathbf{Dx} - \mathbf{d}) \right] \right)$$
$$= \max_{\boldsymbol{\mu} \geq \mathbf{0}} \left(\min_{k=1,\dots,K} \left[\mathbf{c}^{\mathrm{T}} \mathbf{x}^k + \boldsymbol{\mu}^{\mathrm{T}} (\mathbf{Dx}^k - \mathbf{d}) \right] \right)$$

[Picture of a piece-wise linear function of $\boldsymbol{\mu}$!]

$$= \max_{\boldsymbol{\mu} \geq \mathbf{0}, \theta \in \mathbb{R}} \left\{ \theta \;\middle|\; \theta - (\mathbf{Dx}^k - \mathbf{d})^{\mathrm{T}} \boldsymbol{\mu} \leq \mathbf{c}^{\mathrm{T}} \mathbf{x}^k, \quad k = 1, \dots, K \right\}$$

[Picture including θ!]

- Introduce (LP) dual variables y_k. Continuing ...,

The strength of the Lagrangian relaxation, cont.

$$v_L := \max_{\boldsymbol{\mu} \geq \mathbf{0}, \theta \in \mathbb{R}} \left\{ \theta \mid \theta - (\mathbf{D}\mathbf{x}^k - \mathbf{d})^{\mathrm{T}} \boldsymbol{\mu} \leq \mathbf{c}^{\mathrm{T}} \mathbf{x}^k, \quad k = 1, \ldots, K \right\}$$

LP dual:

$$v_L = \min \sum_{k=1}^{K} \left(\mathbf{c}^{\mathrm{T}} \mathbf{x}^k\right) y_k = \mathbf{c}^{\mathrm{T}} \underbrace{\sum_{k=1}^{K} \mathbf{x}^k y_k}_{\in \operatorname{conv} X}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} y_k = 1$$

$$\sum_{k=1}^{K} \left(\mathbf{D}\mathbf{x}^k - \mathbf{d}\right) y_k \leq \mathbf{0} \iff \mathbf{D} \underbrace{\sum_{k=1}^{K} \mathbf{x}^k y_k}_{\in \operatorname{conv} X} \leq \mathbf{d} \underbrace{\sum_{k=1}^{K} y_k}_{=1}$$

$$y_k \geq 0, \qquad k = 1, \ldots, K$$

$$= \min \left\{ \mathbf{c}^{\mathrm{T}} \mathbf{x} \mid \mathbf{D}\mathbf{x} \leq \mathbf{d}, \mathbf{x} \in \operatorname{conv} X \right\} =: v_C$$

The strength of the Lagrangian relaxation, cont.

- Hence, Lagrangian relaxation is a convexification!
- Generating primal solutions through, e.g., Dantzig–Wolfe decomposition, or an ergodic sequence of subproblem solutions (Larsson, Patriksson, and Strömberg, 1999), yields a solution to a primal LP problem equivalent to the original IP problem where, however, X is replaced by $\operatorname{conv} X$
- Using the Relaxation Theorem we conclude that $v^* \geq v_C = v_L \geq v_{LP}$

  C: convexification of X

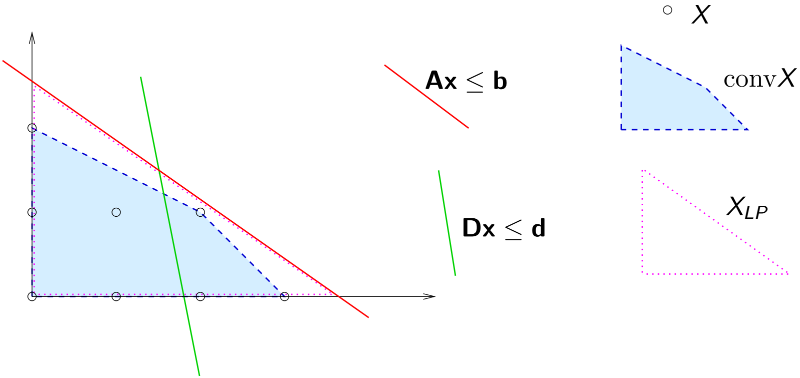
  L: Lagrangian dual

  LP: linear programming relaxation

The strength of a Lagrangian dual problem

Lagrangian and LP bounds

Since $X \subseteq \mathrm{conv} X \subseteq X_{LP} = \left\{ \mathbf{x} \in \mathbb{R}^n_+ \mid \mathbf{Ax} \leq \mathbf{b} \right\}$ it follows that $v^* \geq v_L \geq v_{LP}$



Integrality property

- If $\min_{\mathbf{x} \in X_{LP}} \mathbf{p}^{\mathrm{T}}\mathbf{x} = \min_{\mathbf{x} \in \operatorname{conv} X} \mathbf{p}^{\mathrm{T}}\mathbf{x}$ holds for all $\mathbf{p} \in \mathbb{R}^n$, that is, if the Lagrangian subproblem has the *integrality property*, then it holds that $v_L = v_{LP}$
- Otherwise, v_L is a *better* bound on v^* than is v_{LP} $[v_L \geq v_{LP}]$
- Integrality property $\underbrace{\iff}_{\text{often}}$ easy problem
- Easy subproblem "$\Longrightarrow$" Bad bounds
- Difficult subproblem "$\Longrightarrow$" Better bounds

$\Longrightarrow$ The subproblem should *not* be *too easy* to solve!

The strength of the Lagrangian relaxation: An example

- Consider the *generalized assignment problem* (GAP):

$$
\begin{aligned}
v^* := \min \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} & & \\
\text{s.t.} \quad & \sum_{i=1}^{m} x_{ij} = 1, & j = 1, \dots, n & \quad (1) \\
& \sum_{j=1}^{n} a_{ij} x_{ij} \leq b_i, & i = 1, \dots, m & \quad (2) \\
& x_{ij} \in \{0, 1\}, & \forall i, j & \quad (3)
\end{aligned}
$$

[Draw a bipartite graph!]

The strength of the Lagrangian relaxation, cont.

(1) Every job j must be performed on exactly one machine

(2) The total work done on machine i must not exceed the capacity of the machine.

- Lagrangian relax (1) $\Longrightarrow$ binary knapsack problem (difficult, no integrality property) $\Longrightarrow v_L^1 \leq v^*$
- Lagrangian relax (2) $\Longrightarrow$ Semi-assignment problem (easy, integrality property) $\Longrightarrow v_L^2 \leq v_L^1 \leq v^*$
- Relax (3) to $x_{ij} \in [0, 1] \Longrightarrow$ Linear program (easy, polynomially solvable) $\Longrightarrow v_{LP} = v_L^2$
- Hence, $v_{LP} = v_L^2 \leq v_L^1 \leq v^*$
- We prefer the Lagrangian relaxation of (1), since this gives much better bounds from the Lagrangian dual problem, and knapsack problems are relatively easy to solve (as far as NP-complete problems go . . .)