

# Lectures 5: Lagrangian duality for discrete optimization

Michael Patriksson

2010-09-10

# A reminder of nice properties in the convex case, I

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

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

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

## A reminder of nice properties in the convex case, II

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

- ▶ For a fixed value of  $\mu \geq 0$ , the minimum of  $L(\mathbf{x}, \mu)$  over  $\mathbf{x} \in X$  is attained at  $x_1(\mu) = \frac{\mu}{2}$ ,  $x_2(\mu) = \frac{\mu}{2}$
- $\Rightarrow q(\mu) = L(\mathbf{x}(\mu), \mu) = \dots = 4\mu - \frac{\mu^2}{2}$  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 non-strictly convex  $f$  we have to deal with the non-coordinability, though)
- ▶ What happens otherwise?
- ▶ Recall the relationships between feasible values of  $f$  and  $q$  and the respective optimal values
- ▶ How do we generate optimal primal solutions in the case of a positive duality gap?

# A first example where the duality gap is non-zero, I

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

# A first example where the duality gap is non-zero, II

- ▶ The dual function is computed according to

$$\begin{aligned}q(\mu) &= 7\mu + \underset{\mathbf{x} \in X}{\text{minimum}} \{(3 - \mu)x_1 + (7 - 3\mu)x_2 + (10 - 5\mu)x_3\} \\ &= 7\mu + \underset{x_1 \in \{0,1\}}{\text{minimum}} \{(3 - \mu)x_1\} + \underset{x_2 \in \{0,1\}}{\text{minimum}} \{(7 - 3\mu)x_2\} \\ &\quad + \underset{x_3 \in \{0,1\}}{\text{minimum}} \{(10 - 5\mu)x_3\}\end{aligned}$$

- ▶  $X(\mu)$  is obtained by setting

$$x_j(\mu) = \begin{cases} 1 \\ 0 \end{cases} \quad \text{when the objective coefficient is } \begin{cases} \leq 0 \\ \geq 0 \end{cases}$$

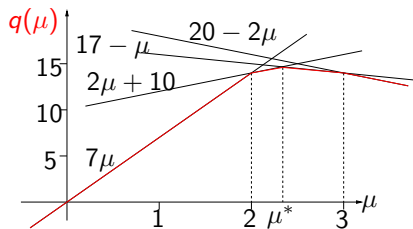
# Subproblem solutions and the dual function, I

- ▶ The subproblem solution:

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

# Subproblem solutions and the dual function, II

- ▶ The dual function:





# Subproblem solutions and the dual function, III

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

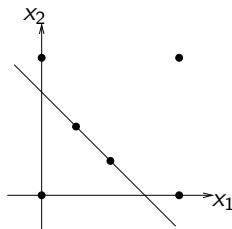
|                   |               |               |
|-------------------|---------------|---------------|
|                   | $\mu < \mu^*$ | $\mu > \mu^*$ |
| slope of $q(\mu)$ | $> 0$         | $< 0$         |
| $\mathbf{x}(\mu)$ | infeasible    | feasible      |

- ▶ Check that the slope equals the value of the constraint function!
- ▶ In particular, the slope of  $q$  is negative for objective pieces corresponding to feasible solutions to the original problem ( $\mathbf{g}(\mathbf{x}(\mu)) \leq \mathbf{0}^m$ )
- ▶ 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!
- ▶  $X(\mu^*) = \{(0, 0, 1), (0, 1, 1)\} \ni \mathbf{x}^*$

# Another example with non-zero duality gap, I

- ▶ 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, II

- ▶ Observe:  $\mu \in \mathbb{R}$  (the relaxed constraint is  $g(\mathbf{x}) = 0$ )
- ▶  $X(\mu) = \operatorname{argmin}_{\mathbf{x} \in \{(0,0), (0,4), (4,4), (4,0), (1,2), (2,1)\}} \{(-2 + \mu)x_1 + (1 + \mu)x_2\}$

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

- ▶  $\mu^* = 2$ ;  $q^* = q(\mu^*) = -6 < f^* = -3$ ,  $\mathbf{x}^* = (2, 1) \notin X(\mu^*)$
- ▶ 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.  $f(\mathbf{x}^*) + (\boldsymbol{\mu}^*)^\top \mathbf{g}(\mathbf{x}^*) = \min_{\mathbf{x} \in X} \{f(\mathbf{x}) + (\boldsymbol{\mu}^*)^\top \mathbf{g}(\mathbf{x})\}$   
[ $\iff \mathbf{x}^* \in X(\boldsymbol{\mu}^*)$ ]
  - ii.  $(\boldsymbol{\mu}^*)^\top \mathbf{g}(\mathbf{x}^*) = 0$
  - iii.  $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}^m$
- (c)  $f^* = f(\mathbf{x}^*) = q(\boldsymbol{\mu}^*) = q^*$

# Method for finding an optimal solution, I

- 1) Solve the Lagrangian dual problem  $\implies \mu^*$ ;
- 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, since the set  $X(\mu^*)$  is normally *not explicitly available*—we normally get *one* element of the set  $X(\mu)$

# Method for finding an optimal solution, II

- ▶ A good example was given in Lecture 3—Example II (the 2-variable LP problem)
- ▶ Imagine using the simplex method for solving each LP subproblem. Then, we only get extreme points of  $X$ , and  $\mathbf{x}^*$  was, in this case, 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  $X$ !

## Method for finding an optimal solution, III

- ▶ Several ways out from this *non-coordinability*:
  - (1) Remember all the points  $\mathbf{x}(\boldsymbol{\mu}^k) \in X(\boldsymbol{\mu}^k)$  visited  
At the end, solve an LP to find the best point in their convex hull—also feasible in the original problem.  
(*The Dantzig–Wolfe (DW) decomposition method*)
  - (2) 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 to DW)  
On the other hand, DW converges finitely for LP problems, which this technique does not. Read the paper by Larsson, Patriksson, and Strömberg (1999)
  - (3) Introduce non-linear price functions for the constraints, instead of the linear one given by Lagrangian relaxation  
(*Augmented Lagrangian methods*)

# Linear integer optimization: The strength of the Lagrangian relaxation, I

- ▶ Compare with a continuous (LP) relaxation:

$$\begin{array}{ll} v_{LP} = \min & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} & \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{Dx} \leq \mathbf{d} \\ & \mathbf{x} \in \mathbb{R}_+^n \end{array} \leq v^* = \min \begin{array}{ll} \mathbf{c}^T \mathbf{x} \\ \text{s.t.} & \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{Dx} \leq \mathbf{d} \\ & \mathbf{x} \in \mathbb{Z}_+^n \end{array}$$

- ▶ Let  $X = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^K\}$  be the set of points in  $X = \{\mathbf{x} \in \mathbb{Z}_+^n \mid \mathbf{Ax} \leq \mathbf{b}\}$



# Linear integer optimization: The strength of the Lagrangian relaxation, II

- ▶ Lagrangian relax the constraints  $\mathbf{D}\mathbf{x} \leq \mathbf{d}$ :

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

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

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

[Picture including  $\theta$ !]

# Linear integer optimization: The strength of the Lagrangian relaxation, III

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

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

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

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

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

$$y_k \geq 0, \quad k = 1, \dots, K$$

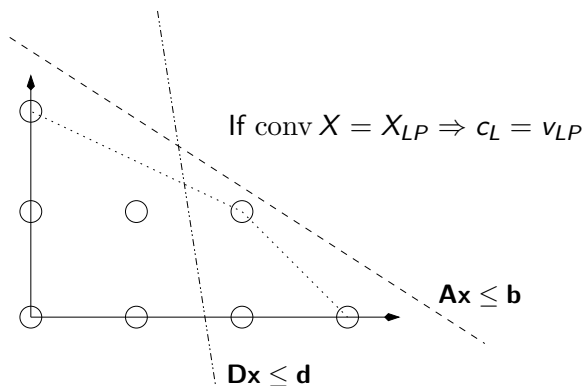
$$= v_C := \min \left\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{D}\mathbf{x} \leq \mathbf{d}, \mathbf{x} \in \text{conv} X \right\}$$

# Linear integer optimization: The strength of the Lagrangian relaxation, IV

- ▶ Hence, Lagrangian relaxation is a convexification!
- ▶ Generating primal solutions through, for example, Dantzig–Wolfe decomposition, or the ergodic sequence method (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  $\text{conv } X$ .
- ▶ We conclude:  $v^* \geq v_C = v_L \geq v_{LP}$ .  
C: convexification of  $X$   
L: Lagrangian dual  
LP: Linear programming relaxation

# The strength of a Lagrangian dual problem

Since  $X \subseteq \text{conv} X \subseteq X_{LP} = \{\mathbf{x} \in \mathbb{R}_+^n \mid \mathbf{Ax} \leq \mathbf{b}\}$  we have that  $v^* \geq v_L \geq v_{LP}$



# Integrality property

- ▶ If  $\min_{\mathbf{x} \in X_{LP}} \mathbf{p}^T \mathbf{x} = \min_{\mathbf{x} \in \text{conv} X} \mathbf{p}^T \mathbf{x}$ , for all  $\mathbf{p} \in \mathbb{R}^n$ , that is, if the Lagrangian subproblem has the *integrality property*, then  $v_L = v_{LP}$ .
  - ▶ Otherwise,  $v_L$  is a *better* bound on  $v^*$  than is  $v_{LP}$  [ $v_L \geq v_{LP}$ ]
  - ▶ Integrality property  $\iff$  easy problem  
often
  - ▶ Easy subproblem “ $\implies$ ” Bad bounds
  - ▶ Difficult subproblem “ $\implies$ ” Better bounds
- $\implies$  The subproblem should *not* be *too easy* to solve!

# The strength of the Lagrangian relaxation

## An example, I

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

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

[Draw a bipartite graph!]

# The strength of the Lagrangian relaxation

## An example, II

- (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)  $\implies$  binary knapsack problem! (Difficult)  
 $\implies v_L^1$
  - ▶ Lagrangian relax (2)  $\implies$  Semi-assignment problem! (Easy!)  
 $\implies v_L^2 \leq v_L^1$
  - ▶ We prefer the Lagrangian relaxation of (1), since this gives much better bounds from the Lagrangian dual problem, and knapsack problems are relatively easy (as far as NP-complete problems go ...)