Operations Research - LINMA 2491

Course 2 - Lagrangian duality for MIP

Mehdi Madani

UCLouvain & N-SIDE

February 10, 2023

Table of contents

Lagrangian relaxations and the Lagrangian Dual

The notes mainly follow (Conforti et al., 2014, [4]) and (Geoffrion, 1974, reprint 2010, [5]).
Let us consider the Mixed Integer Program (MIP)

$$z_I = \max c^T x \tag{1}$$

subject to :

$$
\begin{aligned}
Ax &\leq b & [\pi \geq 0] \quad & (2)\\
Cx &\leq d & [\mu \geq 0] \quad & (3)\\
x_j &\in \mathbb{Z} & \mathit{for}\ j = 1, ..., p \quad & (4)
\end{aligned}
$$

and let $Q := \{x \mid (3) - (4)\}$.

For $\pi \geq 0$,

$$L(\pi) := \max_{x \in Q} c^T x + \pi^T (b - Ax) \tag{5}$$

is called a *Lagrangian relaxation* of the primal (1)-(4) (the choice of the partition of the constraints into $Ax \leq b$ and $C\dot{x}$ depends on the problem structure to exploit)

Lagrangian relaxations and the Lagrangian Dual

For $\pi \geq 0$, $L(\pi)$ is a relaxation since

$$L(\pi) \geq z_I \tag{6}$$

Proof.

Exercise. □

The problem of finding the best upper bound is called the *Lagrangian Dual*:

$$\min_{\pi \geq 0} L(\pi) = \min_{\pi \geq 0} \max_{x \in Q} c^T x + \pi^T (b - Ax) \tag{7}$$

Motivations for solving Lagrangian Duals in MIP

- Find good dual bounds (e.g. upper bounds for maximization problems), that can be leveraged in solution algorithms (Lagrangian-relaxation-based Branch-and-bounds) or to compute optimality gaps
- We first focus here on an application in electricity markets: Convex Hull Pricing (companion slides on the application in scope in Course 2)
- We will then move to general properties of the *LD* and solution methods (next slides)

Convex Hull Pricing

Consider a welfare maximization problem of the form

$$\max_{(u,x)} \sum_c (U_c u_c + W_c x_c) := \sum_c V_c(u_c, x_c) \tag{8}$$

s.t.

$$\sum_c Q_c x_c = 0 \qquad [\pi] \tag{9}$$

$$u_c \in \{0,1\} \qquad \forall c \in C \tag{10}$$

$$(u_c, x_c) \in X_c \qquad \forall c \in C \tag{11}$$

In applications, one can safely assume that the X_c are compact.

The context and notation here are the simplest possible, but can be adapted to consider multiple periods, locations and transmission constraints.

Convex Hull Pricing: lost profits and uplifts

Consider an optimal solution (u^*, x^*) of the previous primal problem, and some market price π.

The lost profit of the market participant c facing market prices π and having to follow the decisions (u_c^*, x_c^*) is:

$$LostProfit_{(u_c^*, x_c^*)}(\pi) := \max_{(u_c, x_c) \in X_c} [V_c(u_c, x_c) - \pi Q_c x_c] - (V_c(u_c^*, x_c^*) - \pi Q_c x_c^*) \quad (12)$$

The lost profit is the gap between the maximum surplus participant c could extract facing the market price π by choosing the best option regarding only its own technical constraints, and the surplus obtained with this same market price and the welfare maximizing solution. The idea with Convex Hull Pricing is to compensate participants for these lost profits, and find prices minimizing the compensations paid.

Minimizing lost profits via a Lagrangian Dual

Theorem

Let π^* solve the Lagrangian dual of (8)-(11) where the balance constraint(s) (9) have been dualized:

$$\min_{\pi} \left[\max_{(u_c, x_c) \in X_c, c \in C} \left[\sum_c V_c(u_c, x_c) - \pi \sum_c Q_c x_c \right] \right] \quad (13)$$

Then, π^* solves:

$$\min_{\pi} \sum_c LostProfit_{(u_c^*, x_c^*)}(\pi) \quad (14)$$

where $LostProfit_{(u_c^*, x_c^*)}(\pi)$ is given by (12) above.

Proof.

$$\sum_c LostProfit_{(u_c^*, x_c^*)}(\pi) \tag{15}$$

$$= \sum_c \left(\max_{(u_c, x_c) \in X_c} [V_c(u_c, x_c) - \pi Q_c x_c] - (V_c(u_c^*, x_c^*) - \pi Q_c x_c^*) \right) \tag{16}$$

$$= \max_{(u_c, x_c) \in X_c} \left(\sum_c [V_c(u_c, x_c) - \pi Q_c x_c] \right) - \sum_c V_c(u_c^*, x_c^*) + \pi \sum_c Q_c x_c^* \tag{17}$$

$$= \max_{(u_c, x_c) \in X_c} \left(\sum_c [V_c(u_c, x_c) - \pi Q_c x_c] \right) - \sum_c V_c(u_c^*, x_c^*) \tag{18}$$

where the last equality follows from the fact that $\sum_c Q_c x_c^* = 0$ (see (9)).

The first term in (18) is a Lagrangian relaxation (where the balance condition (9) has been dualized), and the right-hand side of (18) is the gap between the dual bound of the relaxation with some price π, and the optimal primal objective value. The next slide concludes with the fact that the Lagragian Dual seeks to minimize the gap.

It follows that

$$\min_\pi \sum_c LostProfit_{(u_c^*, x_c^*)}(\pi) =$$
$$\left(\min_\pi \max_{(u_c, x_c) \in X_c} \sum_c [V_c(u_c, x_c) - \pi Q_c x_c] \right) - \sum_c V_c(u_c^*, x_c^*) \tag{19}$$

□

The above developments show that the minimized sum of the lost profits is equal to the duality gap.

If the LD is not solved to optimality, the identity in the previous slide still holds, and the gap (not minimized) is still equal to the sum of the lost profits, which is higher in that case.

Primal reformulation

Let us consider the Lagrangian Dual:

$$z_{LD} = \min_{\pi \geq 0} L(\pi) = \min_{\pi \geq 0} \max_{x \in Q} c^T x + \pi^T (b - Ax) \tag{20}$$

with $Q = \{Cx \leq D, \quad x_j \in \mathbb{Z} \text{ for } j = 1, ..., p\}$.

Theorem

Assume that $\{x \mid Ax \leq b, x \in conv(Q)\}$ is non-empty. Then

(a)

$$z_{LD} = \{\max c^T x \mid Ax \leq b, x \in conv(Q)\} \tag{21}$$

(b) *If Q is compact, the optimal solution in (21) is finite, and the optimal values π^* associated to $Ax \leq b$ in the (classic LP) dual solve the Lagrangian Dual problem.*

This is another motivation to find descriptions of $conv(Q)$ for a Mixed Integer Set Q. Extended formulations (see next course) is a way to describe $conv(Q)$ as the projection of a higher-dimensional polyhedron.

Primal reformulation: proof of (a)

Proof of (a) following [5]. (Proof of (b) is left as an exercise.)
Consider the Lagrangian Dual

$$z_{LD} = \min_{\pi \geq 0} L(\pi) = \min_{\pi \geq 0} \max_{x \in Q} c^T x + \pi^T (b - Ax) \tag{22}$$

1. Since the optimal value of a *linear* function over a set $Q \subseteq \mathbb{R}^n$ is not changed if the set is replaced by its convex hull $conv(Q)$ (see Lemma 1 in the next Section for a more detailed statement + proof):

$$z_{LD} = \min_{\pi \geq 0} \max_{x \in conv(Q)} c^T x + \pi^T (b - Ax) \tag{23}$$

2. By Meyer's theorem (if C, d have rational entries), $conv(Q)$ can be described by some finite set of linear inequalities $Ex \leq F$, and:

$$z_{LD} = \min_{\pi \geq 0} \max_{x \mid Ex \leq F} c^T x + \pi^T (b - Ax) \tag{24}$$

3. Finally, by Lagrangian duality for LPs (see Course 1)
$z_{LD} = \max\{c^T x \mid Ax \leq b, \ Ex \leq F\} = \max\{c^T x \mid Ax \leq b, \ conv(Q)\}$ □

Results used in the Primal reformulation

Lemma (Optimizing a linear function over Q or over $conv(Q)$)

Let $Q \subset \mathbb{R}^n$ and $c \in \mathbb{R}^n$.

(a) $\sup\{c^T x \mid x \in Q\} = \sup\{c^T x \mid x \in conv(Q)\}$

(b) *there exists x^* such that $c^T x^* = \sup\{c^T x \mid x \in Q\}$ if and only if there exists y^* such that $c^T y^* = \sup\{c^T x \mid x \in conv(Q)\}$*

(In other words, the sup is achieved and is a max in the Q case if and only if it is achieved in the $conv(Q)$ case.)

Proof.

See proof of Lemma 1.3 in [4]. □

Theorem (Meyer)

If A, G are matrices with rational entries, and b is a vector with rational entries, there exists A', G', b' with rational entries such that

$$conv(\{(x, y) \mid Ax + Gy \leq b, y \in \mathbb{Z}\}) = \{(x, y) \mid A'x + G'y \leq b'\} \quad (25)$$

Proof.

See [4], p. 159. □

Properties of the Lagrangian Relaxation

Consider the primal $\max\{c^T x \mid Ax \leq b, Cx \leq d, x_j \in \mathbb{Z} \text{ for } j = 1, ..., p\}$

Theorem

The Lagrangian function:

$$L : \pi \geq 0 \rightarrow L(\pi) = \max_{x \in Q = \{x \mid Cx \leq d, x_j \in \mathbb{Z} \text{ for } j=1,\dots,p\}} c^T x + \pi^T (b - Ax) \quad (26)$$

is a piecewise linear convex function over its domain.

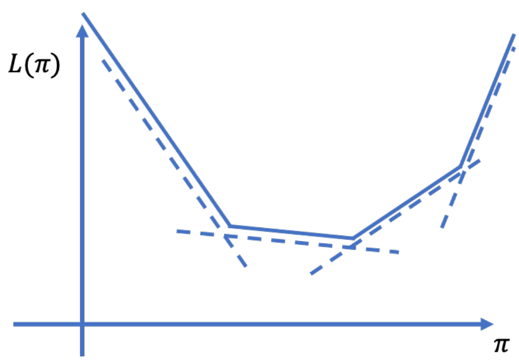

Figure: The Lagrangian function is a piecewise linear convex function

Proof.

We focus on the case where Q is compact and hence $conv(Q)$ is a polytope. Let $v^k, k \in K = \{1, ..., N\}$ denote the vertices of the polytope $conv(Q)$.

$$\begin{aligned} L(\pi) &= \max_{x \in Q = \{x | Cx \leq d, x_j \in \mathbb{Z} \text{ for } j=1,\dots,p\}} c^T x + \pi^T (b - Ax) && (27) \\ &= \max_{x \in conv(Q)} c^T x + \pi^T (b - Ax) && (28) \\ &= \max_{k \in K} c^T v^k + \pi^T (b - Av^k) && (29) \end{aligned}$$

This shows that $L(\pi)$ is the maximum of a finite number of affine functions (one per vertex of $conv(Q)$), and is hence piecewise linear and convex.

$\square$

N.B. When $conv(Q)$ is an unbounded polyhedron, $L(\pi) < +\infty$ if and only if $(c^T r^h - \pi^T A\, r^h) \leq 0$ for each extreme ray r^h of $conv(Q)$, so the domain of $L(\pi)$ is the polyhedron $P = \{\pi \mid c^T r^h - \pi^T A\, r^h) \leq 0, \ \text{for } h \in H\}$ and L is convex over its (convex) domain, hence it is convex.

Dantzig-Wolfe reformulation

Consider a LD reformulation $z_{LD} = \{\max c^T x \mid Ax \leq b, x \in conv(Q)\}$.

Let $v_k, k \in K, r_h, h \in H$ be the vertices and extreme rays of $conv(Q)$.

If a description of $conv(Q)$ is not readily available, one can reformulate the problem using, implicitly, the vertices and extreme rays of $conv(Q)$ as below, and tackle the problem via column generation: the vertices and extreme rays and their associated variables λ, μ will be generated on a need basis thanks to a pricing problem using Q, see LINMA2450.

$$z_{LD} = \max \sum_{k \in K} (c^T v^k) \lambda_k + \sum_{h \in H} (c^T r^h) \mu_h \qquad (30)$$

subject to

$$\begin{aligned} &\sum_{k \in K} (Av^k) \lambda_k + \sum_{h \in H} (Ar^h) \mu_h \leq b && (31) \\ &\sum_{k \in K} \lambda_k = 1 && (32) \\ &\lambda_k, \mu_h \geq 0 \qquad \forall k \in K, h \in H && (33) \end{aligned}$$

Subgradient method

As $L : \pi \to L(\pi)$ is a (non-smooth) convex function, one can minimize it using various methods applicable to this class of problems. (We again restrict ourselves to the case where $conv(Q)$ is compact, so $dom(L) = \mathbb{R}^n_+$.)

Definition

Let f be a convex function $g : \mathbb{R}^n \to \mathbb{R}$. The vector g is a subgradient of f at x if

$$f(y) - f(x) \geq g^T(y - x) \quad \forall y$$

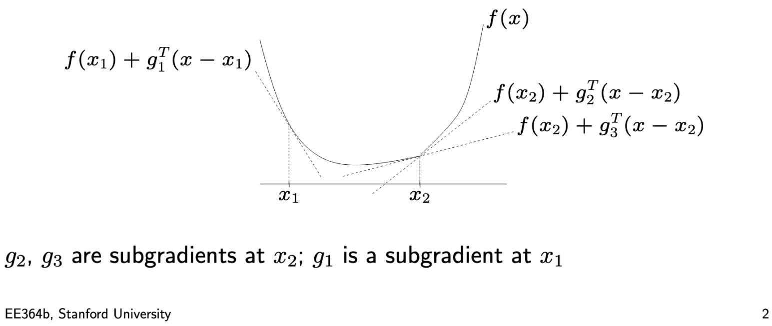

Figure: Illustration of subgradients, taken from [2]

Subgradient method

One can show that the set of subgradients at a given point x is a convex set, called the subdifferential ∂f at x.

Lemma

A point x^* is a minimizer of a convex function f if and only if $0 \in \partial f$

Proof.

Suppose $0 \in \partial f$. Then, by definition of a subgradient at x^*, for any y, $f(y) - f(x^*) \geq 0(y - x^*)$, hence $f(y)) \geq f(x^*)$ and x^* is a minimize.

Conversly, if x^* is a minimizer, for any y, $f(y)) \geq f(x^*)$, so $f(y) - f(x^*) \geq 0(y - x^*)$ and $0 \ \partial f$

□

Subgradient method to solve LD

To solve $\min_{\pi \geq 0} L(\pi) = \min_{\pi \geq 0} \max_{x \in Q} c^T x + \pi^T (b - Ax)$,
we will need to be able to compute a subgradient of L at any point π.

Fortunately:

Lemma

Consider $\pi^* \geq 0$, and x^* a solution of $L(\pi^*) = \max_{x \in Q} c^T x + (\pi^*)^T (b - Ax)$.
Then $g = (b - Ax^*) \in \partial L(\pi^*)$, i.e. $(b - Ax^*)$ is a subgradient of L at π^*

Proof.

$$
\begin{aligned}
L(\pi) &= \max_{x \in Q} c^T x + \pi^T (b - Ax) && (34)\\
&\geq c^T x^* + \pi^T (b - Ax^*) && (35)\\
&= c^T x^* + (\pi^*)^T (b - Ax^*) + (\pi - \pi^*)^T (b - Ax^*) && (36)\\
&= L(\pi^*) + (\pi - \pi^*)^T (b - Ax^*) && (37)
\end{aligned}
$$

which rearranged shows that $L(\pi) - L(\pi^*) \geq (b - Ax^*)^T (\pi - \pi^*)$ □

Subgradient method to solve LD

Consider the subgradient method:

Step 1 Initialization: choose some initial point $\pi^0 \geq 0$, set $L_{best} = L(\pi_0)$, and the counter t:=1.

Step 2 Compute a subgradient of L at π^t as $(b - Ax_t^*)$ where x_t^* solves $\max_{x \in Q} c^T x + (\pi^t)^T (b - Ax^*)$

Step 3 Updates of the current iterate + check of the stopping criterion

- Update the current point: $\pi^{t+1} := \pi^t - \lambda_t (b - Ax_t^*)$.
  N.B. If $\pi^{t+1} < 0$, replace it by 0 = projection of π^{t+1} on the positive orthant = $dom(L)$.
- Update the current best value of $L_{best}^{t+1} = min[L_{best}^t, L(\pi^{t+1})]$
- $t \rightarrow t + 1$. If some predefined stopping criterion is met (e.g. max number of iterations), stop. See [3] for relevant stopping criteria (and a broader coverage of subgradient methods)

The sequence λ_t is called the steplength sequence. If appropriately chosen, the method will converge to an optimal solution π^* of L.

Subgradient method to solve LD

Theorem

(Poljak, 1967, [6], Theorem 8.11 in [4])

Assume that

- *the Lagrangian dual* $\min_{\pi \geq 0} L(\pi)$ *admits an optimal solution* L^*
- *the length of all subgradients of L is bounded by a constant* $S > 0$
- *the steplength sequence* λ_t *satisfies* $\lim_{t \to +\infty} \lambda_t = 0, \sum_{t=1}^{+\infty} \lambda_t = +\infty$

Then the sequence L^t_{best} *converges to* L^* *as* $t \to +\infty$*.*

For example, the steplength sequence $\lambda_t = \frac{1}{t}$ satisfies the conditions for the steplength.

LP duals & recovery of primal solutions

Consider the primal LP

$$\max c^T x \tag{38}$$

subject to :

$$Ax \leq b \qquad [\pi \geq 0] \tag{39}$$

$$Cx \leq d \qquad [\mu \geq 0] \tag{40}$$

and the Lagrangian Dual:

$$\min_{\pi \geq 0} L(\pi) = \min_{\pi \geq 0} \max_{x | Cx \leq d} c^T x + \pi^T (b - Ax) \tag{41}$$

Suppose that we solve it via the subgradient method. Though L^t_{best} converges to L^* (and the corresponding π^t to π^*), the "primal iterates" x^*_t in general won't converge to a primal optimal solution. However, under certain conditions, a sequence $\overline{x^*_t}$ converging to a primal optimal solution can be built. For example, if the steplength sequence is $\lambda_t = \frac{1}{t}$, any accumulation point of the sequence $\overline{x^*_T} = \frac{1}{T} \sum_{t=1}^{T} x^*_t$ is a primal optimal solution, see [1].

References I

[1] K. M. Anstreicher and L. A. Wolsey.
Two "well-known" properties of subgradient optimization.
Mathematical Programming, 120(1):213–220, 2009.

[2] S. Boyd.
Subgradient methods.
EE364b, Stanford University.
Available at `https://web.stanford.edu/class/ee364b/lectures/subgradients_slides.pdf`.

[3] S. Boyd and J. Park.
Subgradient methods.
lecture notes of EE364b, Stanford University, Spring 2013–14, 2024.
Available at `https://web.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf`.

[4] M. Conforti, G. Cornuéjols, G. Zambelli, et al.
Integer programming, volume 271.
Springer, 2014.

References II

[5] A. M. Geoffrion.
Lagrangian relaxation for integer programming.
In *50 Years of Integer Programming 1958-2008*, pages 243–281. Springer, 2010.

[6] B. T. Polyak.
A general method for solving extremal problems.
In *Doklady Akademii Nauk*, volume 174, pages 33–36. Russian Academy of Sciences, 1967.