B6.3 Integer Programming

Raphael Hauser

Mathematical Institute
University Of Oxford

Michaelmas Term 2018

Lagrangian Relaxation

Many IPs have a structure

$$
\begin{aligned}
\text{(IP)} \quad z = \max\, & c^{\mathsf{T}} x \\
\text{s.t. } & Ax \leq a \\
& Dx \leq d \\
& x \geq 0, \ x \in \mathbb{Z}^{n},
\end{aligned}
$$

such that relaxing the constraints $Dx \leq d$ yields a substantially more tractable problem where $Ax \leq a$ is a *benign* set of constraints (e.g., totally unimodular) in the sense that

$$
\begin{aligned}
\max\, & c^{\mathsf{T}} x \\
\text{s.t. } & Ax \leq a \\
& x \geq 0, \ x \in \mathbb{Z}^{n}.
\end{aligned}
$$

Thus, we may interpret $Ax \leq a$ as benign and $Dx \leq d$ as malicous constraints that render the problem (IP) hard to solve.

Note that what is benign or malicous is in the eye of the beholder, as it may be that

$$
\begin{aligned}
\max\, & c^{\mathsf{T}} x \\
\text{s.t. } & Dx \leq d \\
& x \geq 0, \ x \in \mathbb{Z}^{n}
\end{aligned}
$$

is also an easy problem, but it is the *combination* of the two constraint sets $Ax \leq a$ and $Dx \leq d$ that renders the problem hard.

Example (Uncapacitated facility location (UFL))

Consider the uncapacitated facility location problem from Lecture 1,

$$\begin{aligned}
\text{(IP)} \quad z = \max & \sum_{i \in M} \sum_{j \in N} c_{ij} x_{ij} - \sum_{j \in N} f_j y_j \\
\text{s.t.} & \sum_{j \in N} x_{ij} = 1 \quad (i \in M) \\
& x_{ij} - y_j \leq 0 \quad (i \in M, j \in N) \\
& x \in \mathbb{R}_+^{|M| \times |N|}, \; y \in \{0,1\}^{|N|},
\end{aligned}$$

where

- M is the set of customer locations,
- N is the set of potential facility locations,
- f_j are the fixed costs for opening facility j,
- we replaced the original servicing costs c_{ij} with $-c_{ij}$ to turn the problem into a maximisation problem.

One may take the viewpoint that it is the demand constraints

$$(1) \qquad \sum_{j \in N} x_{ij} = 1, \quad (i \in M)$$

that render the problem hard, because these constraints introduce a functional dependence between the decisions pertaining to different facility locations.

Example (UFL continued)

Instead of imposing these constraints, let us add a multiple u_i of each residual

$$1 - \sum_{j \in N} x_{ij}$$

to the objective function. The objective function is now

$$\sum_{i \in M} \sum_{j \in N} c_{ij} x_{ij} - \sum_{j \in N} f_j y_j + \sum_{i \in M} u_i (1 - \sum_{j \in N} x_{ij}),$$

and we say that the constraints (1) have been *dualised*.

The new problem is called a *Lagrangian relaxation*,

$$
\begin{aligned}
(\mathrm{IP}(u)) \quad z(u) = \max & \sum_{i \in M} \sum_{j \in N} (c_{ij} - u_i) x_{ij} - \sum_{j \in N} f_j y_j + \sum_{i \in M} u_i \\
\text{s.t. } & x_{ij} - y_j \leq 0 \quad (i \in M, j \in N) \\
& x \in \mathbb{R}_+^{|M| \times |N|}, \; y \in \{0,1\}^{|N|}.
\end{aligned}
$$

Example (UFL continued)

Note that because the constraint that linked the different facility locations to one another have been subsumed in the objective function, (IP(u)) decouples,

$$z(u) = \sum_{j \in N} z_j(u) + \sum_{i \in M} u_i,$$

where $z_j(u)$ is the optimal solution of the following problem,

$$\begin{aligned}(\mathrm{IP}_j(u)) \quad z_j(u) = \max \sum_{i \in M} &(c_{ij} - u_i)x_{ij} - f_j y_j \\ \text{s.t. } x_{ij} - y_j &\leq 0 \quad (i \in M) \\ x_{ij} \geq 0 \ (i \in M),&\ \ y_j \in \{0,1\}^{|N|}.\end{aligned}$$

Furthermore, (IP$_j(u)$) is easily solved by inspection:

- If $y_j = 0$, then $x_{ij} = 0$ for all i, and the objective value is 0.
- If $y_j = 1$, then all clients i for which $c_{ij} - u_i > 0$ will be served, and the objective value is $\sum_{i \in M} \max(0, c_{ij} - u_i)$.

Therefore, $z_j(u) = \max\left(0, \sum_{i \in M} \max(0, c_{ij} - u_i) - f_j\right)$.

Definition (Relaxation)

A relaxation of an integer programming problem (IP) $z = \max\{f(x) : x \in \mathscr{F}\}$ is any optimisation problem (R) $w = \max\{g(x) : x \in \mathscr{R}\}$ with feasible set $\mathscr{R} \supseteq \mathscr{F}$ and an objective function $g(x)$ that satisfies $g(x) \geq f(x)$ for all $x \in \mathscr{F}$.

Lemma (Dual bounds by relaxation)

If (R) is a relaxation of (IP), then $w \geq z$.

Proof. See Problem Sheet 4.

Corollary (Optimality by relaxation)

Let $x^* \in \arg\max\{g(x) : x \in \mathscr{R}\}$. If $x^* \in \mathscr{F}$ and $g(x^*) = f(x^*)$, then x^* is an optimal solution of (IP).

Proof. By Lemma (Dual bounds by relaxation) then the following inequality holds for all $x \in \mathscr{F}$,

$$\mathbf{c}^{\mathbf{T}} x \leq z \leq w = g(x^*) = \mathbf{c}^{\mathbf{T}} x^*.$$

Example (UFL continued)

Problem (IP(u)) constructed in Example (UFL) is indeed a relaxation:

- Giving up on the requirement $\sum_{j\in N} x_{ij} = 1$ constitutes an enlargement of the feasible set.
- The restriction of the new objective

$$\max_{x,y} g(x,y) = \sum_{i\in M}\sum_{j\in N} c_{ij}x_{ij} - \sum_{j\in N} f_j y_j + \sum_{i\in M} u_i(1 - \sum_{j\in N} x_{ij})$$

to the feasible set of the UFL coincides with the objective of the latter,

$$\max_{x,y} \sum_{i\in M}\sum_{j\in N} c_{ij}x_{ij} - \sum_{j\in N} f_j y_j,$$

as any (UFL)-feasible solution (x, y) satisfies the demand constraints $\sum_{j\in N} x_{ij} = 1$, which implies

$$\sum_{i\in M} u_i(1 - \sum_{j\in N} x_{ij}) = 0.$$

Generalisation

Let us now consider an (IP) in the slightly more general form

$$
\begin{aligned}
\text{(IP)} \quad z = \max\, & c^{\mathrm{T}} x \\
\text{s.t. } & D_1 x \le d_1, \\
& D_2 x = d_2, \\
& x \in \mathscr{X} = \{x \in \mathbb{R}^n : Ax \le a,\ x \ge 0,\ x \in \mathbb{Z}^n\}
\end{aligned}
$$

where $\mathscr{X}$ is a feasible set of "benign" type.

We write $D = [D_1^{\mathrm{T}}, D_2^{\mathrm{T}}]^{\mathrm{T}}$ and $d = [d_1^{\mathrm{T}}, d_2^{\mathrm{T}}]^{\mathrm{T}}$ in block form and denote the set of row indices of D that correspond to inequality constraints by $\mathcal{I}$ and indices corresponding to equality constraints by $\mathcal{E}$.

Definition (Lagrangian relaxation)

A *Lagrangian relaxation* of (IP) is a problem of the form

$$\text{(IP}(u)) \quad z(u) = \max\{c^{\mathrm{T}} x + u^{\mathrm{T}}(d - Dx) : x \in \mathscr{X}\}$$

where $u \in \mathbb{R}^m$ is a fixed vector *Lagrange multipliers* chosen so that $u_i \ge 0$ for $i \in \mathcal{I}$.

Proposition (Lagrangian relaxations)

Problem (IP(u)) *is a relaxation of problem* (*IP*).

Proof. The feasible region of (IP(u)) contains that of (IP), since

$$\mathscr{X} \supseteq \mathscr{F} = \{x \in \mathscr{X} : D_1 x \leq d_1,\ D_2 x = d_2\}.$$

For all (IP)-feasible x, the objective function of (IP(u)) is at least as large as that of (IP),

$$c^T x + u^T (d - Dx) = c^T x + \sum_{i \in \mathcal{I}} u_i (d_i - D_{i,:} x) \geq c^T x.$$

In the context of Lagrangian relaxations, Corollary (Optimality by relaxation) can be recast in terms of a complementarity condition:

Proposition (Optimality by Lagrangian relaxation)

Let $x(u)$ be an optimal solution of the Lagrangian relaxation

$$(\mathrm{IP}(u)) \quad x(u) \in \arg\max_{x} \{c^{\mathsf{T}} x + u^{\mathsf{T}}(d - Dx) : x \in \mathscr{X}\}.$$

If $x(u)$ is (IP)-feasible, that is $D_i x \leq d_i$ for all $(i \in \mathcal{I})$ and $D_i x = d_i$ for all $(i \in \mathcal{E})$, and if the complementarity conditions

$$u_i(d_i - [Dx(u)]_i) = 0 \quad \forall\, i \in \mathcal{I}$$

are satisfied, then $x(u)$ is an optimal solution of (IP).

Proof. By complementarity, $z \leq z(u) = c^{\mathsf{T}} x(u) + u^{\mathsf{T}}(d - Dx(u)) = c^{\mathsf{T}} x(u) \leq z$, hence $c^{\mathsf{T}} x(u) = z$.

Example (Lagrangian relaxation of STSP)

Recall our formulation of the symmetric travelling salesman problem from Lecture 1,

$$\begin{aligned}
\text{(IP)}\quad z = \min & \sum_{e\in E} c_e x_e \\
\text{s.t.} & \sum_{e\in\delta(i)} x_e = 2 \quad (i \in V) \\
& \sum_{e\in E(S)} x_e \le |S| - 1 \quad (S \subset V \text{ s.t. } 2 \le |S| \le |V| - 1) \\
& x \in \{0,1\}^{|E|},
\end{aligned}$$

Lemma (Redundant subtour elimination constraints)

Half the subtour elimination constraints $\sum_{e\in E(S)} x_e \le |S| - 1$ *are redundant.*

Proof. For any x feasible for the LP relaxation of (IP) we have

$$|S| - \sum_{e\in E(S)} x_e = \frac{1}{2}\sum_{i\in S}\sum_{e\in\delta(i)} x_e - \sum_{e\in E(S)} x_e = \frac{1}{2}\sum_{e\in\delta(S,S^c)} x_e,$$

where $\delta(S, S^c)$ is the set of edges in E that are incident to one node from S and one from $S^c := V \setminus S$. Since $\delta(S, S^c) = \delta(S^c, S)$, we now have

$$|S| - \sum_{e\in E(S)} x_e = \frac{1}{2}\sum_{e\in\delta(S,S^c)} x_e = |S^c| - \sum_{e\in E(S^c)} x_e,$$

and hence, $\sum_{e\in E(S)} x_e \le |S| - 1 \Leftrightarrow \sum_{e\in E(S^c)} x_e \le |S^c| - 1$.

Example (Lagrangian relaxation of STSP continued)

- Introduce a new (redundant) constraint $\sum_{e \in E} x_e = n$, obtained by summing all degree constraints.
- Eliminate all subtour elimination constraints corresponding to sets S that contain node 1, which are redundant by Lemma (Redundant subtour elimination constraints).
- Dualise the degree constraints $\sum_{e \in \delta(i)} x_e = 2$, $(i \neq 1)$.

This yields the following Lagrangian relaxation of (STSP),

$$
\begin{aligned}
(\mathrm{IP}(u)) \quad z(u) = \min & \sum_{e=(ij) \in E} (c_e - u_i - u_j) x_e + 2 \sum_{i \in V} u_i \\
& \sum_{e \in \delta(1)} x_e = 2 \\
& \sum_{e \in E(S)} x_e \leq |S| - 1, \quad \forall S \subset V \text{ s.t. } 2 \leq |S| \leq |V| - 1,\ 1 \notin S \\
& \sum_{e \in E} x_e = n \\
& x \in \{0, 1\}^{|E|}.
\end{aligned}
$$

For notational convenience we included a term $u_1 (2 - \sum_{e \in \delta(1)} x_e) = 0$ in the objective.

Lemma (1-Tree Characterisation)

A binary vector $x \in \{0,1\}^{|E|}$ *is* $(\mathrm{IP}(u))$*-feasible if and only if its support* $E(x) := \{e \in E : x_e = 1\})$ *is a* 1-tree *in* $G = (V, E)$.

Proof. $\sum_{e \in \delta(1)} x_e = 2$ guarantees that in the subgraph $G_x := (V, E(x))$ exactly two edges are incident to node 1 .

Constraints $\sum_{e \in E(S)} x_e \leq |S| - 1$ guarantee that when node 1 is removed, then there is no cycle left in $E(x) \setminus \delta(1)$.

$\sum_{e \in E} x_e = n$ guarantees that $|E(x) \setminus \delta(1)| = n - 2$ is a cycle free subgraph on $|V \setminus \{1\}| = n - 1$ nodes, which is only possible if $E(x) \setminus \delta(1)$ is a spanning tree on $V \setminus \{1\}$.

Conversely, if $E(x)$ is a 1-tree, then $\sum_{e \in \delta(1)} x_e = 2$ and $\sum_{e \in E} x_e = n$ are clearly satisfied, and since $E(x) \setminus \delta(1)$ a tree, the subtour elimination constraints are satisfied.

Example (Lagrangian relaxation of STSP continued)

Let us now look at a numerical example and consider the STSP on 5 nodes with edge cost matrix

$$[c_e] = \begin{bmatrix} - & 30 & 26 & 50 & 40 \\ 30 & - & 24 & 40 & 50 \\ 26 & 24 & - & 24 & 26 \\ 50 & 40 & 24 & - & 30 \\ 40 & 50 & 26 & 30 & - \end{bmatrix}.$$

Note that the Lagrange multipliers u are unrestricted, since dualised constraints are *equality constraints*. Therefore,

$$u = \begin{bmatrix} 0 & 0 & -15 & 0 & 0 & 0 \end{bmatrix}$$

is a legitimate choice.

Writing $\bar{c}_{ij} := c_{ij} - u_i - u_j$, we obtain the revised edge cost matrix

$$[\bar{c}_e] = \begin{bmatrix} - & 30 & 41 & 50 & 40 \\ 30 & - & 39 & 40 & 50 \\ 41 & 39 & - & 39 & 41 \\ 50 & 40 & 39 & - & 30 \\ 40 & 50 & 41 & 30 & - \end{bmatrix}.$$

Using the greedy algorithm, we find that $\{(1,2),(1,5),(4,5),(2,3),(3,4)\}$ is a minimum weight 1-tree for the revised edge costs. Since this is a Hamiltonian tour, it must be an optimal solution of the STSP by Proposition (Optimality by Lagrangian relaxation).

The Lagrangian Dual Problem

In summary so far, it follows from Proposition (Lagrangian relaxations) and Lemma (Dual bounds by relaxation) that for all $u \in \mathbb{R}^m$ with $u_i \geq 0$ for $i \in \mathcal{I}$,

$$z(u) := \max\{c^{\mathsf{T}} x + u^{\mathsf{T}}(d - Dx) : \ x \in \mathscr{X}\}$$

is a dual bound on the optimal objective value of problem (IP),

$$z = \max\{c^{\mathsf{T}} x : \ x \in \mathscr{X}, \ D_1 x \leq d_1, \ D_2 x = d_2\}.$$

Definition (Lagrangian Dual)

The problem of finding the best upper bound obtainable in this fashion can now be cast as an optimisation problem over the Lagrange multipliers u as decision variables,

$$\text{(LD)} \quad w_{LD} = \min\{z(u) : \ u \in \mathbb{R}^m, \ u_i \geq 0 \ (i \in \mathcal{I})\}.$$

This is called the *Lagrangian Dual* of problem (IP),

Theorem (Characterisation of Lagrangian dual bound)

The Lagrangian dual bound is characterised as follows,

$$(\mathrm{LD}') \quad w_{LD} = \max_{x} c^{\mathsf{T}} x \quad \text{s.t. } D_1 x \leq d_1, \quad D_2 x = d_2, \quad x \in \operatorname{conv}(\mathscr{X}).$$

Proof. We give the proof in the special case where $\mathscr{X} = \{x^{[1]}, \dots, x^{[T]}\}$ is a finite set. Then

$$\begin{aligned} w_{LD} &= \min_{u_i \geq 0,\, i \in \mathcal{I}} z(u) = \min_{u_i \geq 0,\, i \in \mathcal{I}} \left\{ \max\{c^{\mathsf{T}} x^{[t]} + u^{\mathsf{T}}(d - Dx^{[t]}) : t = 1, \dots, T\} \right\} \\ &= \min_{(\eta, u) \in \mathbb{R}^{m+1}} \left\{ \eta : \eta \geq c^{\mathsf{T}} x^{[t]} + u^{\mathsf{T}}(d - Dx^{[t]}),\ (t = 1, \dots, T),\ u_i \geq 0,\ i \in \mathcal{I} \right\} \\ &= \min_{(\eta, u) \in \mathbb{R}^{m+1}} \left\{ \eta + 0^{\mathsf{T}} u : \eta + (Dx^{[t]} - d)^{\mathsf{T}} u \geq c^{\mathsf{T}} x^{[t]},\ (t = 1, \dots, T),\ u_i \geq 0,\ i \in \mathcal{I} \right\}. \end{aligned}$$

Taking the dual of the latter LP, strong LP duality implies

$$\begin{aligned} w_{LD} &= \max_{\mu \in \mathbb{R}^T} \Bigg\{ \sum_{t=1}^{T} \mu_t (c^{\mathsf{T}} x^{[t]}) : \sum_{t=1}^{T} \mu_t (Dx^{[t]} - d)_i \leq 0,\ (i \in \mathcal{I}),\ \sum_{t=1}^{T} \mu_t (Dx^{[t]} - d)_i = 0, \\ &\qquad\qquad (i \in \mathcal{E}),\ \sum_{t=1}^{T} \mu_t = 1,\ \mu \geq 0 \Bigg\} \\ &= \max_{\mu \in \mathbb{R}^T} \left\{ c^{\mathsf{T}} x : D_1 x - d_1 \leq 0,\ D_2 x - d_2 = 0,\ x \in \operatorname{conv}(\{x^{[1]}, \dots, x^{[T]}\}) \right\}. \end{aligned}$$