More on Lagrangian relaxation

(Wolsey: Chapter 10)

Consider an integer optimization problem of form:

\[
(P) \quad \begin{align*}
  z &= \max c^T x \\
  D x &\leq d \\
  x &\in X
\end{align*}
\]

"complicating constraints"

\[
\begin{align*}
  \text{and include the } \\
  \text{integerity constraints}
\end{align*}
\]

Let \( D \) be \( m \times n \) matrix,

\( d \in \mathbb{R}^m \)

\( x \in \mathbb{R}^n \)

For any \( \mu = (\mu_1, \ldots, \mu_m) > 0 \), consider the Lagrangian relaxation:

\[
(P_\mu) \quad \begin{align*}
  z(\mu) &= \max c^T x + \mu (d_1 - D x) \\
  x &\in X
\end{align*}
\]
Proposition: \((P_u)\) is a relaxation of \((P)\) for all \(u \geq 0\).

Proof:

(i) \(\forall x : Dx \leq d, x \in X \forall y \in X\)

the feasible region of \((P_u)\) is at least as large as \((P)\).

(ii) \(c x + u(d - Dx) \geq c x \quad \forall x \in X\)

\[\min \quad \geq 0 \quad \geq 0\]

\(\forall x \in X\) the objective value \(u_i (P_u)\) is at least as great as \(u_i (P)\).

Therefore: \(z(\mu) \geq z \quad \forall \mu \geq 0\)

(maximization problem)
To find the best (= smallest) upper bound \( z(u) \) for \( u > 0 \), we solve the Lagrangian Dual Problem:

\[
\begin{align*}
& (LD) \quad \omega_{LD} = \min_{u > 0} z(u) \\
& \text{Solving a Lagrangian relaxation may sometimes lead to solve (P):} \\
& \text{let } u > 0: \\
& \text{Proposition:} \quad \text{If:} \\
& (i) \quad x(u) \text{ is an optimal solution to } (P_u) \\
& (ii) \quad D x(u) \leq d \\
& (iii) \quad \text{If } u_i > 0 \text{ then } (D x(u))_i = d_i \quad \text{(complementarity)} \\
& \text{then } x(u) \text{ is optimal for (P).} \\
& \text{Proof:} \\
& \text{From (i):} \\
& \omega_{LD} \leq z(u) = c x(u) + u (d - D x(u)) \]
From (iii):
\[ c x(w) + u(d - D x(w)) = c x(w) \]

\[ \Rightarrow 0 \]

From (ii):
\[ c x(w) \leq \varepsilon \quad \text{since } x(w) \text{ is feasible for } (P) \]

It follows that:
\[ w_{2D} \leq c x(w) \leq \varepsilon \]

Since \( w_{2D} \geq \varepsilon \) (it provides an upper bound), we get \( w_{2D} = c x(w) = \varepsilon \).

So, \( x(w) \) is optimal for \((P)\).

\[ \text{Obs 1: in this case } w \text{ is optimal for } (2D) \]

\[ \text{Obs 2: if the constraints "dualized" (i.e. } D x = d \text{) are equality constraints (i.e. } D x = d \text{) then (iii) is always satisfied.} \]
In this case, an optimal solution \( x(u) \) to \((P_u)\) is optimal for \((P)\) if it is feasible for \((P)\).

**Example**

\[\text{(UFL)}\]

consider the strong formulation \((UFL_2)\)

\[
\max_{\mu} \quad \sum_{i \in H} \sum_{j \in E} c_{ij} x_{ij} + \sum_{j \in N} g_j y_j
\]

\[
\text{subject to:} \\
\sum_{j \in N} x_{ij} = 1 \quad \forall i \in H
\]

\[
x_{ij} \leq y_j \quad \forall i \in H, j \in N
\]

\[
x_{ij} \geq 0 \quad \forall i \in H, j \in N
\]

\[
y_j \in [0, +\infty) \quad \forall j \in N
\]

\[
\varepsilon(u) = \min_{\mu} \sum_{i \in H} \sum_{j \in E} (c_{ij} - \mu_i) x_{ij} + \sum_{j \in N} g_j y_j + \sum_{i \in H} \mu_i
\]

\[
x_{ij} \leq y_j \quad \forall i \in H, j \in N
\]

\[
x_{ij} \geq 0 \quad \forall i \in H, j \in N
\]

\[
y_j \in [0, +\infty) \quad \forall j \in N
\]
\((P_{\mu})\) decomposes into \(n\) subproblems, one for each \(\mu\):

\[
z(\mu) = \sum_{\gamma \in N} z_\gamma(\mu) + \sum_{i \in H} \mu_i.
\]

where

\((P^0_{\mu})\)

\[
z_\gamma(\mu) = \min \sum_{\gamma \in H} (c_{i,\gamma} - \mu_i) x_{i,\gamma} + b_{\gamma} y_{\gamma}.
\]

\[
x_{i,\gamma} \leq y_{\gamma} \quad \forall i \in H
\]

\[
x_{i,\gamma} \geq 0 \quad \forall i \in H
\]

\[
y_{\gamma} \in \mathbb{Q}, \forall \gamma
\]

\((P^0_{\mu})\) can be solved by inspection:

1) if \(y_{\gamma} = 0\) then \(x_{\gamma} = 0 \quad \forall i \in H\), so

\[
z^1_{\gamma}(\mu) = 0.
\]

2) if \(y_{\gamma} = 1\):

\[
x_{\gamma} = \frac{1}{2} \quad \text{if} \quad c_{i,\gamma} - \mu_i < 0 \quad \forall i \in H
\]

\[
0 \quad \text{otherwise}
\]

so

\[
z^2_{\gamma}(\mu) = \sum_{i \in H} \min \left\{ c_{i,\gamma} - \mu_i, 0 \right\} y_{\gamma} + b_{\gamma}.
\]
Therefore:

\[ z_0^*(u) \quad z_1^*(u) \]

\[ z_2(u) = \min \{ 0, \sum_{i} c_{ij} - u_i, 0 \} + \frac{f}{g} \]

\[ u_i = 0 \quad \alpha = 1 \]

< easily solvable! >

Example

\[ m = 6 \quad n = 5 \]

\[ f = (2, 4, 5, 3, 3) \]

\[ (c_{ij}) = \begin{bmatrix} 6 & 2 & 1 & 3 & 5 \\ 4 & 10 & 2 & 6 & 3 \\ 3 & 2 & 4 & 1 & 3 \\ 2 & 0 & 4 & 1 & 4 \\ 1 & 8 & 6 & 2 & 5 \\ 3 & 2 & 4 & 8 & 1 \end{bmatrix} \]

If we choose:

\[ u = (5 \ 6 \ 3 \ 2 \ 5 \ 4) \]
Modified costs:

\[
\begin{bmatrix}
1 & -3 & -4 & -2 & 0 \\
-2 & 4 & -4 & 0 & -5 \\
0 & -1 & 1 & -2 & 0 \\
0 & -2 & 2 & -1 & 2 \\
-4 & 3 & 1 & -3 & 0 \\
-1 & -2 & 0 & 4 & -3
\end{bmatrix}
\]

\[
\sum_{i \in H} u_i = 25
\]

For \( j = 2 \):

- \( y_{2j} = 0 \) then \( x_{12} = 1 \)
- \( y_{22} = 1 \) then \( x_{22} = 1 \)
- \( x_{32} = 1 \)
- \( x_{42} = 1 \)
- \( x_{62} = 1 \)

\[
\bar{z}_2^2(u) = -3 - 1 - 2
\]

\[
-2 + \frac{8}{2} = -4
\]

So it is optimal to set

\( y_{22} = 1 \) giving \( \bar{z}_2(u) = -4 \)
We have to perform a similar calculation for each location to find $z(u)$:

$$z(u) = z_1(u) + z_2(u) + z_3(u) + z_4(u) + z_5(u) + \sum_{i \in \mathcal{N}} u_i$$

For other $u$?

- How can we find the best Lagrangian bound?

- How good is such an upper bound, i.e. $w_u$?
Strength of the Lograngian Dual

Let \((P)\) \[ z = \max c x \]
\[ \text{subject to } D x \leq d \]
\[ x \in X \]

\((P_{\mu})\) \[ z(\mu) = \max c x + \mu(d - D x) \]
\[ x \in X \]
\(\mu \geq 0\)

\((\leq D)\) \[ w_{\leq D} = \min \{ z(\mu) : \mu \geq 0 \} \]

Suppose for simplicity that \(X = \{x_1, \ldots, x_T\}\), i.e. finite set of points:

\[ w_{\leq D} = \min \{ \max \{ c x + \mu(d - D x) \} : \mu \geq 0 \} \]
\[ x \in X \]

\[ = \min \{ \max \{ c x_\ell + \mu(d - D x_\ell) \} : \ell = 1, \ldots, T \} \]
\[ \mu \geq 0 \]

\[ = \min \{ \max \{ c x_\ell + \mu(d - D x_\ell) \} : \ell = 1, \ldots, T \}
\]

maximum of T values
Let us introduce an auxiliary variable \( \eta \) to estimate the maximum

\[
\eta_d = \min \eta \quad \text{constant}
\]

\[
(\lambda_t) \quad \eta \geq c x_t + w (d - D x_t) \quad t = 1, \ldots, T
\]

\[
w > 0
\]

The latter is a linear programming problem. By the previous assumption on \( x \) it has finite optimum value. Therefore, from strong duality we can replace it by its dual:

\[
\eta_d = \max \sum_{t=1}^{T} \lambda_t (c x_t)
\]

\[
\sum_{t=1}^{T} \lambda_t (D x_t - d) \leq 0
\]

\[
\sum_{t=1}^{T} \lambda_t = 1
\]

\[
\lambda_t \geq 0 \quad t = 1, \ldots, T
\]

Setting

\[
x = \sum_{t=1}^{T} \lambda_t x_t \quad \text{s.t.} \quad \sum_{t=1}^{T} \lambda_t = 1, \lambda_t \geq 0 \quad t = 1, \ldots, T
\]
we get
\[ \omega_{\leq D} = \max c \cdot \sum_{\ell=1}^{T} \lambda_{\ell} x_{\ell} \]

\[ \sum_{\ell=1}^{T} \sum_{x=\ell}^{\infty} D \cdot \sum_{\ell=1}^{T} \lambda_{\ell} x_{\ell} - d \cdot \sum_{\ell=1}^{T} \lambda_{\ell} \leq 0 \]

\[ x = \sum_{\ell=1}^{T} \lambda_{\ell} x_{\ell} \]

\[ \lambda_{\ell} \geq 0 \quad \ell = 1, \ldots, T \]

That is,

\[ \omega_{\leq D} = \max c \cdot x \]

\[ D x \leq d \]

\[ x \in \text{conv}(X) \]

called "convexified relaxation"

So:

**Theorem:** \( \omega_{\leq D} = \max \{ c \cdot x : D x \leq d, x \in \text{conv}(X) \} \)

More generally, this holds true for any \( X = \{ x \in \mathbb{Z}^{n} : A x \leq b \} \).
This theorem gives information on the strength of the Lagrangian dual. In certain cases, it is no stronger than the LP bound:

**Corollary 1:** If \( X = \frac{1}{y} x \in \mathbb{Z}^n : A x \leq b y \) and \( \text{conv}(X) = \{ x \in \mathbb{R}^n : A x \leq b y \} \), then \( \omega_{LD} = \max \{ cx : D x = d, A x \leq b y \} \), which is the LP bound (LP relaxation).

**Obs 1:** write

\[
\begin{align*}
(P_u) \quad z(u) &= \max \{ cx + u(d - Dx) : \\
A x &\leq b \\
x &\in \mathbb{Z}^n \}
\end{align*}
\]

then if \((P_u)\) satisfies the integrality property, i.e., \( \text{conv}(\{ x \in \mathbb{Z}^n : A x \leq b y \}) = \{ x \in \mathbb{R}^n : A x \leq b y \}\), then \( \omega_{LD} \) is equal to and so no stronger than the LP bound.

**Obs 2:** the property is true also for minimization and mixed integer problems.
example (UFL2)

Recall that (P_u) has always an integer optimum solution; therefore \( w_{LD} \) corresponding to the considered Lagrangian relaxation is equal to the LP bound corresponding to (UFL2).

**Corollary 2:** If (P) is a Linear Programming problem, then \( w_{LD} = z \).

The proof of the theorem suggests how to compute \( w_{LD} \), i.e., how to solve the Lagrangian Dual.

**Corollary 3:** \( w_{LD} \leq z_{LP} \), where \( z_{LP} \) is the linear programming bound (for a maximization problem).

**Proof:**
\[
\text{conv}(X) \subseteq \{ x \in \mathbb{R}^n : A x \leq b \}
\]
if \( X = \{ x \in \mathbb{Z}^n : A x \leq b \} \).