Lagrangian Relaxation Notes

John E. Mitchell

Department of Mathematical Sciences
RPI, Troy, NY 12180 USA

April 2019
Outline

1. Assignment problem with budget constraint
   - Relax the budget constraint
   - Relax the assignment constraints

2. Solving the Lagrangian dual problem
We consider the assignment problem with a budget constraint

\[
\max_{x \in \mathbb{R}^{n^2}} \quad \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}
\]

subject to

\[
\sum_{i=1}^{n} x_{ij} = 1 \quad \forall j
\]

\[
\sum_{j=1}^{n} x_{ij} = 1 \quad \forall i
\]

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} t_{ij} x_{ij} \leq b
\]

\[
x \quad \text{binary}
\]

Several different Lagrangian relaxations are possible.
Outline

1. **Assignment problem with budget constraint**
   - Relax the budget constraint
   - Relax the assignment constraints

2. **Solving the Lagrangian dual problem**
Relax the budget constraint

If we relax the final constraint, the Lagrangian relaxation is the assignment problem

\[
\max_{x \in \mathbb{R}^{n^2}} \quad b\lambda + \sum_{i=1}^{n} \sum_{j=1}^{n} (c_{ij} - \lambda t_{ij}) x_{ij}
\]

subject to

\[
\sum_{i=1}^{n} x_{ij} = 1 \quad \forall j
\]

\[
\sum_{j=1}^{n} x_{ij} = 1 \quad \forall i
\]

\[x \text{ binary}\]

Since the Lagrangian relaxation is an assignment problem, it can be solved by solving its LP relaxation.

Thus, for this approach, the optimal value of the Lagrangian dual is only as good as the value of the LP relaxation of the original problem.
Lagrangian dual:
\[ Z_{LD} = \max \sum_i \sum_j c_{ij} x_{ij} \]
\[ s.t. \sum_i \sum_j c_{ij} x_{ij} \leq b \]
\[ x \in \text{conv}(Q) \]

= \max \sum_i \sum_j c_{ij} x_{ij}
\[ s.t. \sum_i \sum_j c_{ij} x_{ij} \leq b \]
\[ \sum_j x_{ij} = 1 \ \forall j \]
\[ \sum_i x_{ij} = 1 \ \forall i \]
\[ x_{ij} \geq 0 \]

So \( Z_{LD} = Z_{LP} \), the value of the LP relaxation.

No integrality restriction.
Outline

1. Assignment problem with budget constraint
   - Relax the budget constraint
   - Relax the assignment constraints

2. Solving the Lagrangian dual problem
Relax the assignment constraints

If we relax the assignment constraints, we introduce two sets of Lagrangian multipliers: $\lambda$ for the first set of constraints and $\mu$ for the second set.

The Lagrangian relaxation is a *knapsack problem*:

$$\max_{x \in \mathbb{R}^{n^2}} \sum_{j=1}^{n} \lambda_j + \sum_{i=1}^{n} \mu_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \left( c_{ij} - \lambda_j - \mu_i \right) x_{ij}$$

subject to

$$\sum_{i=1}^{n} \sum_{j=1}^{n} t_{ij} x_{ij} \leq b$$

$$x \text{ binary}$$

Since the Lagrangian relaxation cannot be solved just by solving its LP relaxation, we should obtain a better bound, so the optimal value of the Lagrangian dual should be larger than the optimal value of the LP relaxation of the original problem.

The drawback is that it is harder to solve the LP relaxation.
Other relaxations are possible.
The trade off is in the ease of solving the relaxation versus the strength of the bound.
Outline

1. Assignment problem with budget constraint
   - Relax the budget constraint
   - Relax the assignment constraints

2. Solving the Lagrangian dual problem
An integer program

We let $z_{IP}$ denote the optimal value of the integer program

$$z_{IP} := \max c^T x$$

subject to

$$A^1 x \leq b^1$$  \hspace{0.5cm} \text{(complicating constraints)}
$$A^2 x \leq b^2$$  \hspace{0.5cm} \text{(nice constraints)}
$$x \in \mathbb{Z}_+^n$$

The matrices $A^1 \in \mathbb{R}^{m_1 \times n}$ and $A^2 \in \mathbb{R}^{m_2 \times n}$, and all vectors are dimensioned appropriately.
The Lagrangian relaxation

For any $\lambda \in \mathbb{R}_+^{m_1}$, we obtain the **Lagrangian relaxation**

$$z_{LR}(\lambda) := \max \quad c^T x + \lambda^T (b^1 - A^1 x)$$
subject to

$$A^2 x \leq b^2$$

$$x \in \mathbb{Z}_+^n$$
The Lagrangian dual problem

The optimal value of the Lagrangian relaxation problem gives an upper bound on the optimal value of the integer program for any $\lambda \geq 0$. In the Lagrangian dual problem, this upper bound is minimized:

$$z_{LD} := \min \, z_{LR}(\lambda) \quad \text{subject to} \quad \lambda \geq 0.$$
**Convexity**

**Theorem**

The function \( z_{LR}(\lambda) \) is convex.

**Proof.**

Let \( Q \) denote the feasible region for the Lagrangian relaxation. Note that this is the same for all \( \lambda \). For simplicity, we assume \( Q \) is bounded, so it contains a finite number of points \( \{x^1, \ldots, x^q\} \). For any \( \lambda \), we have

\[
z_{LR}(\lambda) = \max_{x \in Q} \{ c^T x + \lambda^T (b^1 - A^1 x) \} = \lambda^T b^1 + \max_{i=1,\ldots,q} \left\{ \left( c^T - \lambda^T A^1 \right) x^i \right\}
\]

which is the maximum of a finite number of linear functions, which is itself a piecewise linear convex function.
Convexity

Theorem

The function $z_{LR}(\lambda)$ is convex.

Proof.

Let $Q$ denote the feasible region for the Lagrangian relaxation. Note that this is the same for all $\lambda$. For simplicity, we assume $Q$ is bounded, so it contains a finite number of points $\{x^1, \ldots, x^q\}$. For any $\lambda$, we have

$$z_{LR}(\lambda) = \max_{x \in Q} \left\{ c^T x + \lambda^T (b^1 - A^1 x) \right\} = \lambda^T b^1 + \max_{i=1,\ldots,q} \left\{ \left( c^T - \lambda^T A^1 \right) x^i \right\}$$

which is the maximum of a finite number of linear functions, which is itself a piecewise linear convex function.
**Alternative proof:**

\[
\max f(x)
\]

s.t. \( g_i(x) \leq 0 \quad i = 1, \ldots, m \quad x \in X \)

\[
Z_{LR}(\lambda) = \max_{x \in X} f(x) - \sum_{i} \lambda_i g_i(x)
\]

**Show:** \( Z_{LR}(\lambda) \leq \mu Z_{LR}(\lambda') + (1-\mu) Z_{LR}(\lambda^2) \)

**Assume:** \( Z_{LR}(\lambda) \) is achieved by \( \bar{x} \)

So \( Z_{LR}(\lambda) = f(\bar{x}) - \sum_{i} \bar{x}_i g_i(\bar{x}) \)

\[
= \mu \left[ f(\bar{x}) - \sum_{i} \lambda'_i g_i(\bar{x}) \right] + (1-\mu) \left[ f(\bar{x}) - \sum_{i} \lambda^2_i g_i(\bar{x}) \right]
\]

from definition of \( \bar{x} \)

\[
\leq \mu Z_{LR}(\lambda') + (1-\mu) Z_{LR}(\lambda^2) \quad \text{from def of } Z_{LR}
\]

as required
Subgradients

Theorem

Let $\hat{x}$ solve (3) for a given value $\lambda$. Then

$$z_{LR}(\lambda) \geq z_{LR}(\hat{\lambda}) + \left(\lambda - \hat{\lambda}\right)^T \left(b^1 - A^1 \hat{x}\right),$$

so $b^1 - A^1 \hat{x}$ is a subgradient of the convex function $z_{LR}(\lambda)$ at $\lambda = \hat{\lambda}$. 
The proof

Proof.

This follows directly from the equality in the previous theorem. In particular, we have

\[ z_{LR}(\lambda) \geq c^T \hat{x} + \lambda^T (b^1 - A^1 \hat{x}) \quad \text{since } \hat{x} \in Q \]

\[ = c^T \hat{x} + \hat{\lambda}^T (b^1 - A^1 \hat{x}) + (\lambda - \hat{\lambda})^T (b^1 - A^1 \hat{x}) \]

\[ = z_{LR}(\hat{\lambda}) + (\lambda - \hat{\lambda})^T (b^1 - A^1 \hat{x}) \]

as required.
Algorithms

Since we have a subgradient for the convex function $z_{LR}(\lambda)$, we can minimize it using various different convex optimization algorithms. For example, we can use subgradient descent methods.
Cutting plane methods

Alternatively, we can use cutting plane methods, using the inequality from Theorem 2.

Each time we solve the Lagrangian relaxation, we get a solution $x^p$. We thus can approximate the Lagrangian dual function using the linear program:

\[
\begin{align*}
\min_{\theta, \lambda} & \quad \theta \\
\text{subject to} & \quad \theta \geq c^T x^p + \lambda^T (b^1 - A^1 x^p) \quad \text{for } p = 1, 2, \ldots \\
& \quad \lambda \geq 0
\end{align*}
\]
Example

For example, returning to the problem from the previous lecture, perhaps initially we have two points $x^1 = (0, 3)$ and $x^2 = (3, 2)$, giving us the following piecewise linear approximation:
Solve a subproblem

This approximation is minimized by $\hat{\lambda} = \frac{3}{7}$. The value of the approximation is $\hat{\theta} = 2\frac{1}{7}$. We then get

$$z_{LR} \left( \frac{3}{7} \right) = \max_{x \in \mathbb{R}^2} x_1 + \frac{3}{7} (2 - 2x_1 + x_2)$$
subject to

$$\begin{align*}
x_1 &+ 2x_2 \leq 8 \\
x_1 &\leq 3 \\
x_2 &\leq 3 \\
x &\geq 0, \text{ integer}
\end{align*}$$

The solution is $x = (2, 3)$ with value $z_{LR}(\frac{3}{7}) = 2\frac{3}{7} > \hat{\theta}$.

The constraint added to the Lagrangian dual approximation is

$$\theta \geq 2\frac{3}{7} + \left( \lambda - \frac{3}{7} \right) (2 - 4 + 3) = 2\frac{3}{7} + \left( \lambda - \frac{3}{7} \right) = 2 + \lambda.$$
Updated approximation to Lagrangian dual function

\[
\begin{align*}
\theta &= 5\lambda \\
\theta &= 2 + \lambda \\
\theta &= 3 - 2\lambda
\end{align*}
\]

\((\frac{1}{3}, 2\frac{1}{3}) (\frac{3}{7}, 2\frac{3}{7})\)
Solve updated approximation

The updated value of $\hat{\lambda} = \frac{1}{3}$ with the value of the approximation equal to $\hat{\theta} = 2\frac{1}{3}$.

It can be verified that $z_{LR}(\hat{\lambda}) = \hat{\theta}$, so the algorithm can terminate at this point with the optimal solution to the Lagrangian dual.
Separable subproblems

When the Lagrangian relaxation is separable, it may be advantageous computationally to break up the scalar variable $\theta$ into a sum of variables $\theta_i$, each corresponding to one of the separated subproblems.