Lagrange Relaxation on TSP

Prasangsha Ganguly

December 28, 2020

1 Background

Consider the following problem: **IP:**

$$Z = Max\ C^T X \tag{1}$$
$$S.T.\ AX \leq b \tag{2}$$
$$DX \leq d,\ d \in \mathbb{Z}^m \tag{3}$$
$$X \in \mathbb{Z}^n_+ \tag{4}$$

That is, we divide the full set of constraints into two subsets. Assume, $AX \leq b, X \in \mathbb{Z}^n_+$ are nice in the sense that, solving IP with only these constraints is easy. Let, $\chi = \{X : AX \leq b, X \in \mathbb{Z}^n_+\}$. Now, let $\mathbf{u} \geq 0$ be defined as $(u_1, ..., u_m)$ where, each u_i corresponds to each one of the m constraints of $DX \leq d$.

Now, consider the transformation, $\mathbf{IP(u)}$ i.e., the problem parameterized by u:

$$Z(u) = Max\ C^T X + \mathbf{u}^T(d - DX) \tag{5}$$
$$S.T.\ X \in \chi \tag{6}$$

Here, $(d - DX)$ indicates the violation of the constraint $DX \leq d$. If the constraint is violated, $(d - DX)$ is going to be negative, lowering the objective value of a maximization problem which is not desirable. We say that $\mathbf{IP(u)}$ is a relaxation of **IP** because, every feasible solution of **IP** is feasible to $\mathbf{IP(u)}$ because $\mathbf{IP(u)}$ has less number of constraints. Also, objective of $\mathbf{IP(u)}$ is at least as good as objective of **IP**. Hence, $\mathbf{IP(u)}$ is a relaxation of **IP** known as Lagrange Relaxation.

In reality, $Z(u)$ give us a lower bound parameterized by u. But to obtain the best bound we have to identify the best solution over all u. Hence we end up by having a **min-max problem** corresponding to the original max problem:

$$\omega_{LR} = Min_u\ Z(u) \tag{7}$$
$$Z(u) = Max\ C^T X + \mathbf{u}^T(d - DX) \tag{8}$$
$$S.T.\ X \in \chi \tag{9}$$

Similarly, if the the original problem is a min problem, we will have $\omega_{LR} = Max\ Z(u)$ and $Z(u)$ is a min problem.

Now, assume that, $\{X^1, X^2, ..., X^k\}$are all the feasible solutions of χ. Now, the Lagrange Dual can be represented as,

$$\omega_{LD} = Min_{u \geq 0} \ \{Max_{l=1,...,k} \ \{C^T X^l + \mathbf{u}^T(d - DX^l)\}\}$$

Hence, it can also be represented as, **LD**:

$$\omega_{LD} = Min_{\mathbf{u}} \ \eta \tag{10}$$

$$\eta \geq C^T X^l + \mathbf{u}^T(d - DX^l) \ \ \forall l = 1, ..., k \tag{11}$$

$$\mathbf{u} \geq 0 \tag{12}$$

Here, the decision variables are, $\mathbf{u}$ and η. As we know X^l, as the feasible solution of the easy part of the problem, it is not a variable. Every constraint of (11), is of the form of a straight line and there are k such constraints. Hence we have a piecewise linear convex region as intersection of all these constraints. Thus, we conclude that, for any linear optimization problem, even if the original problem is non-convex, the resulting Lagrange relaxation is convex. The optimal solution of Lagrange relaxation may not be the optimal solution of the non-convex problem but it provides a bound. The $\mathbf{u}$ are called Lagrange dual or Lagrange multipliers.

Essentially, if we take the dual of the **LD**, then we end up with a problem as,

$$Max \ C^T X \tag{13}$$

$$S.T. \ DX \leq d \tag{14}$$

$$X \in Conv(\chi) \tag{15}$$

We knew, χ was easy to solve. So, we are actually convexifying the easy part of the problem. By this way, we move close towards the convex hull of the original problem.

2 Applying LR to TSP

When LR is applied to TSP, then the constraint that is going to be dualized is,

$$\sum_{e \in \delta(i)} X_e = 2 \ \ \forall i \in V \setminus \{1\}$$

. i.e., the constraint $DX \leq d$ corresponds to this constraint which says, for every vertex there are two edges adjacent to it. We remove this constraint to form χ. We say, for TSP, χ is the set of all $\mathbf{1 - trees}$. We have convexified that region. $1 - tree$ is a graph such that, it is a spanning tree for the vertices $2, ..., n$ and there are 2 edges incident on node 1. Finding the minimum cost $1 - tree$ is an easy problem. Essentially, in a minimum $1 - tree$, we first remove the vertex 1 from the graph and find a MST using Prim's algorithm (say). Then we select two minimum cost edges adjacent to vertex 1 and add them to the MST. It results in a $1 - tree$.

3 Solving LR using Sub-gradient Algorithm

As we have to optimize over $\mathbf{u}$, we use sub-gradient algorithm. Let us assume that, the current solution at time t is given by $\mathbf{u_t}$. Let, δ be the direction towards which we move to improve our solution and μ is the step size that we move each time. So,

$$\mathbf{u_{t+1}} = \mathbf{u_t} + \mu_t \delta_t$$

Hence, we are going to start with an initial solution and then, we are going to find a direction and step length at that instant and then that is going to led us to a new solution.

3.1 Identifying a good δ

We note that the dual of the TSP is a maximization problem. As we do in the case of a maximization problem, we move towards the gradient to maximize. As our region has piece wise linear functions, we will use the best gradient.

We have

$$Max\{Z(\mathbf{u}) = Min\{C^T X + \mathbf{u}^T(d - DX)\}\}$$

So, we take the best gradient as,

$$\frac{\partial}{\partial u_i}[C^T X^*(u) + \mathbf{u}^T(d - DX^*(u))]$$

where $X^*(u)$ is the optimal for $Z(u)$ for the inner minimization problem. Thus, we get, required $\delta = (d - DX^*(u))$

3.2 Identifying a good μ

We pick a sequence of μ_t such that it gradually decreases time after time and and at $t = \infty$ it becomes 0 and if we sum all of them, we get ∞. Satisfying these two conditions, we pick μ as

$$\mu_t = M\rho^t$$

where $M > 0$ and $0 < \rho < 1$. Typically, $M = 10$ and $\rho = 0.8$.

3.3 The Algorithm

Finally the sub-gradient algorithm with the different stopping criteria is given as follows.

1. Solve $Z(u_t) = L_t$. Here, the

   $$L_t = Min \sum_{e \in E}(c_e - u_i - u_j)X_e + 2\sum_{i \in V} u_i$$

   Let, $X(u_t)$ be the corresponding solution of Lagrange dual.

2. Find

   $$\delta_t = [2 - \sum_{e \in \delta(i)} X_e(u_t)]_{i=1}^n$$

3. If $\delta_t = 0$, stop. The solution is optimal.

4. If $|L_t - L_{t+1}| < \epsilon$, stop. No further improvement is possible.

5. If $t > T$, stop. Too many iterations. Else go to step 4.

6. $t = t + 1$, $\mu_t = M\rho^t$ and $u_{t+1} = u_t + \delta\mu$. Go to step 1.