Travelling Salesman Problem

Basic algorithm

**Definition 1** A Hamiltonian cycle in a graph is a simple circuit that passes through each vertex exactly once.

**Problem.** Let $G(V,E)$ be a complete undirected graph, and $c : E \mapsto \mathbb{R}^+$ be a cost function defined on the edges. Find a minimum cost Hamiltonian cycle in $G$.

This problem is called a travelling salesman problem. To illustrate the motivation, imagine that a salesman has a list of towns it should visit (and at the end to return to his home city), and a map, where there is a certain distance between any two towns. The salesman wants to minimize the total distance traveled.

**Definition 2** We say that the cost function defined on the edges satisfies a triangle inequality if for any three edges $\{u,v\}$, $\{u,w\}$ and $\{v,w\}$ in $E$ it holds:

$$c(\{u,v\}) + c(\{v,w\}) \geq c(\{u,w\}) .$$

In what follows we assume that the cost function $c$ satisfies the triangle inequality.

**The main idea of the algorithm:**

Observe that removing an edge from the optimal travelling salesman tour leaves a path through all vertices, which contains a spanning tree.

We have the following relations:

$$\text{cost(TSP)} \geq \text{cost(Hamiltonian path without an edge)} \geq \text{cost(MST)} .$$

---

1 Additional reading: Section 3 in the book of V.V. Vazirani “Approximation Algorithms”.
Next, we describe how to build a travelling salesman path. Find a minimum spanning tree in the graph. Then by using the edges of the tree in both directions will result in a (not necessarily simple) circuit that passes through all vertices, possibly some vertices are visited more than once. This path has a total cost equal to $2 \cdot \text{cost}(\text{MST})$.

If some town is visited more than once, we can modify a path such that will skip the town that was already visited. For example, if Tartu was already visited, we replace the part of the path Tallinn → Tartu → Riga by a direct link Tallinn → Riga:

Due to a triangle inequality,

$$c(e_1) + c(e_2) \geq c(e_3),$$

this procedure can only reduce the total cost of the path. Therefore,

$$\text{cost(solution)} \leq 2 \cdot \text{cost}(\text{MST}) \leq 2 \cdot \text{cost(OPT)}.$$ 

We obtain an approximation algorithm with approximation factor 2.

Improved algorithm

Another view on the previous solution.

1. Find an MST;
2. Double every edge of MST to obtain an Eulerian graph (the degree of every vertex is even);
3. Find an Eulerian path in that graph (the path that uses every edge once);
4. Output the tour that visits vertices of $G$ in the order of their appearance.
Reminder: An undirected graph $G(V, E)$ has an Eulerian cycle if and only if all of its vertices have even degrees.

Algorithm

1. Find $T$, an MST in $G$;
2. Compute a minimum cost perfect matching $M$ in a subgraph induced by the odd-degree vertices in $T$;
3. Find an Eulerian cycle in $T \cup M$;
4. Output the tour that visits the vertices of $G$ in the order of their appearance in the Eulerian cycle.

Lemma 1 Let $V' \subseteq V$ such that $|V'|$ is even. Let $M$ be a minimum cost perfect matching of the subgraph induced by the vertices of $V'$. Then $\text{cost}(M) \leq \frac{\text{OPT}}{2}$.

Proof. Consider an optimal travelling salesman tour of $G$, $\tau$. Let $\tau'$ be a tour of a subgraph induced by $V'$, obtained by short-cutting $\tau$. By the triangle inequality,

$$\text{cost}(\tau') \leq \text{cost}(\tau).$$

Observe that $\tau'$ is the union of two perfect matchings of $V'$, each matching consists of alternate edges of $\tau'$. Thus, one of these two matching, which has the lower cost, has cost

$$\text{cost}(M) \leq \frac{\text{cost}(\tau')}{2} \leq \frac{\text{OPT}}{2}.$$

Theorem 2 The approximation factor of the proposed algorithm is $3/2$.

Proof.

$$\text{cost(solution)} = \text{cost}(T) + \text{cost}(M) \leq \text{OPT} + \frac{\text{OPT}}{2} = \frac{3}{2} \cdot \text{OPT}.$$

TSP in directed graphs

Let $G(V, E)$ be a simple complete directed graph. Assume the triangle inequality

$$c((u, v)) + c((v, w)) \geq c((u, w)).$$

for any $u, v, w \in V$.

Goal: find a cyclic path that visits each vertex once. The previous approach does not work because the graph is not symmetric anymore.

Definition 3 A vertex-disjoint cycle cover is a collection of simple circuits such that every vertex in $V$ participates in exactly one such circuit (simple circuits of length 2 are allowed). Note: a vertex-disjoint cycle cover can be found in a polynomial time.
A cycle-shrinking algorithm

**Input:** $\mathcal{G}(V, E), c : E \mapsto \mathbb{R}^+$.  

**Output:** Hamiltonian cycle in $\mathcal{G}$ with the minimum cost.

1. Find a minimum cost vertex-disjoint cycle cover;  
2. Pick a representing vertex for each cycle;  
3. Recursively solve the problem on representatives of each cycle;  
4. Extend the Hamiltonian cycle using smaller cycles.  

**Example:**

![Graph Example](image)

**Theorem 3** Let $\{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_k\}$ be a collection of vertex-disjoint cycle covers, where $\mathcal{C}_i$ is the $i$-th cycle cover found by the algorithm. Let $c(\mathcal{C}_i)$ be the cost of $\mathcal{C}_i$. Then, the following holds:

**Claim 1:** $c(\mathcal{C}_i) \leq \text{OPT}$;  

**Claim 2:** $k \leq \log_2 n$.  

**Proof:**

**Claim 1:** Let $\mathcal{V}_i$ be a set of representatives in the level $i$ of recursion, and $\mathcal{G}_i$ be a graph induced from $\mathcal{G}$ by $\mathcal{V}_i$. Then,

$$\text{OPT}(\mathcal{G}_i) \leq \text{OPT}(\mathcal{G}),$$

where $\text{OPT}(\mathcal{G}_i)$ and $\text{OPT}(\mathcal{G})$ denote the cost of the optimal Hamiltonian cycle in $\mathcal{G}_i$ and $\mathcal{G}$, respectively. This can be shown by short-cutting the Hamiltonian walk on $\mathcal{G}$ so it passes through the vertices in $\mathcal{V}_i$ only.
Let $W_i$ be the optimal Hamiltonian walk in $G$ short-cutted to $G_i$. Then, $c(C_i) \leq c(W_i) = \text{OPT}(G_i)$, because Hamiltonian walk is a cycle cover. Then,

$$c(C_i) \leq c(W_i) \leq \text{OPT}(G) ,$$

which proves Claim 1.

**Claim 2:** at most half of the vertices survive till the next stage. Thus,

$$|\mathcal{V}_{i+1}| \leq \frac{1}{2} \cdot |\mathcal{V}_i| .$$

The claim follows. □