

Lecture 19, Nov 20, 2023

Minimum Spanning Trees (MSTs)

- Recall that a *tree* is any undirected graph that contains no cycles
- Note some tree facts:
 - A tree with n nodes has $n - 1$ edges
 - Adding any edge to a tree creates a cycle containing that edge
 - Removing any edge from that cycle gives you a tree

Definition

Given a connected graph G , a *spanning tree* of the graph is any tree T that connects (i.e. spans) all the nodes of G .

If G is a weighted graph, then a *minimum spanning tree* (MST) of G is any tree T that has the minimum possible weight, defined as the total weight of all its edges.

- The spanning tree will be a subgraph of G , and is not unique; the MST is also non-unique if there are repeated edge weights
- The MST construction problem is the problem of constructing an MST T of G , given a weighted undirected connected graph $G = (V, E)$
- A *spanning forest* of G is a “fragment” of a spanning tree of G (i.e. the spanning tree with some edges removed)
 - It is a set of trees $T_1 = (V_1, E_1), \dots, T_k = (V_k, E_k)$, such that $\{V_1, V_2, \dots, V_k\}$ is a partition of V and $E_i \subseteq E$

Theorem

MST Construction Theorem: Suppose some MST T of G contains the spanning forest T_1, \dots, T_k of G . Let (u, v) be an edge G of minimum weight between T_i and the other trees of the spanning forest (i.e. $u \in V_i, v \in V \setminus V_i$), then there must be an MST T^* that contains T_1, \dots, T_k and the edge (u, v) .

- This means starting from the trivial forest (where there are no edges), we can repeatedly apply this algorithm to build up the MST edge by edge, and stop when the number of edges is exactly $n - 1$
- Proof:
 - Suppose some MST T of G contains the spanning forest T_1, \dots, T_k of G
 - Let (u, v) be an edge of G of minimum weight that connects T_i and any other tree in the forest
 - If the edge is a part of T , then $T^* = T$ and we're done
 - Otherwise, let $T' = T + (u, v)$ (i.e. add the new edge to T); this will create a cycle containing (u, v) , so T' is no longer a tree
 - The existence of this cycle means there is another edge (u', v') such that $u' \in V_i$ and $v' \in V \setminus V_i$ (i.e. (u', v') goes back from the rest of the world to V_i)
 - Since (u, v) has minimum weight, the weight of (u, v) must be less than or equal to the weight of (u', v')
 - Let $T^* = T' - (u', v') = T + (u, v) - (u', v')$
 - * Note that T^* must be a spanning tree, since removing an edge from the cycle gives back a tree
 - * T^* contains T_1, \dots, T_k and (u, v)
 - * Also, the weight of T^* is less than or equal to T ; but T is an MST, so T^* must have the same weight and therefore it is an MST