10 The Minimum-Cost Flow Problem

The remaining lectures will be concerned with optimization problems on networks, in particular with flow problems.

10.1 Networks

A directed *graph*, or *network*, $G = (V, E)$ consists of a set V of *vertices* and a set $E \subseteq V \times V$ of *edges*. When the relation E is symmetric, G is called an undirected graph, and we can write edges as unordered pairs $\{u, v\} \in E$ for $u, v \in V$. The *degree* of vertex $u \in V$ in graph G is the number $|\{v \in V : (u, v) \in E \text{ or } (v, u) \in E\}|$ of other vertices connected to it by an edge. A *walk* from $u \in V$ to $w \in V$ is a sequence of vertices $v_1, \dots, v_k \in V$ such that $v_1 = u$, $v_k = w$, and $(v_i, v_{i+1}) \in E$ for $i = 1, \dots, k-1$. In a directed graph, we can also consider an undirected walk where $(v_i, v_{i+1}) \in E$ or $(v_{i+1}, v_i) \in E$ for $i = 1, \dots, k-1$. A walk is a *path* if $v_1, \dots, v_k$ are pairwise distinct, and a *cycle* if $v_1, \dots, v_{k-1}$ are pairwise distinct and $v_k = v_1$. A graph that does not contain any cycles is called *acyclic*. A graph is called *connected* if for every pair of vertices $u, v \in V$ there is an undirected path from u to v. A *tree* is a graph that is connected and acyclic. A graph $G' = (V', E')$ is a subgraph of graph $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. In the special case where G' is a tree and $V' = V$, it is called a *spanning tree* of G.

10.2 Minimum-Cost Flows

Consider a network $G = (V, E)$ with $|V| = n$, and let $b \in \mathbb{R}^n$. Here, b_i denotes the amount of flow that enters or leaves the network at vertex $i \in V$. If $b_i > 0$, we say that i is a *source* supplying b_i units of flow. If $b_i < 0$, we say that i is a *sink* with a demand of $|b_i|$ units of flow. Further let $C, \underline{M}, \overline{M} \in \mathbb{R}^{n \times n}$, where c_{ij} denotes the cost associated with one unit of flow on edge $(i, j) \in E$, and $\underline{m}_{ij}$ and $\overline{m}_{ij}$ respectively denote lower and upper bounds on the flow across this edge. The minimum-cost flow problem then asks for flows x_{ij} that conserve the flow at each vertex, respect the upper and lower bounds, and minimize the overall cost. Formally, $x \in \mathbb{R}^{n \times n}$ is a *minimum-cost flow* of G if it is an optimal solution of the following optimization problem:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{(i,j) \in E} c_{ij} x_{ij} \\
\text{subject to} \quad & b_i + \sum_{j:(j,i) \in E} x_{ji} = \sum_{j:(i,j) \in E} x_{ij} && \text{for all } i \in V, \\
& \underline{m}_{ij} \leqslant x_{ij} \leqslant \overline{m}_{ij} && \text{for all } (i, j) \in E.
\end{aligned}
$$

The minimum-cost flow problem is a linear programming problem, with constraints of the form $Ax = b$ where

$$a_{ik} = \begin{cases} 1 & \text{kth edge starts at vertex } i, \\ -1 & \text{kth edge ends at vertex } i, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\sum_{i \in V} b_i = 0$ is required for feasibility, and that a problem satisfying this condition can be transformed into an equivalent problem where $b_i = 0$ for all i by introducing an additional vertex, and new edges from each sink to the new vertex and from the new vertex to each of the sources with upper and lower bounds equal to the flow that should enter the sources and leave the sinks. The latter problem is known as a *circulation problem*, because flow does not enter or leave the network but merely circulates. We can further assume without loss of generality that the network G is connected. Otherwise the problem can be decomposed into several smaller problems that can be solved independently.

An important special case is that of *uncapacitated flow problems*, where $\underline{m}_{ij} = 0$ and $\overline{m}_{ij} = \infty$ for all $(i,j) \in E$. Clearly, an uncapacitated flow problem is either unbounded, or has an equivalent problem with finite capacities.

10.3 Sufficient Conditions for Optimality

The Lagrangian of the minimum-cost circulation problem is

$$L(x, \lambda) = \sum_{(i,j) \in E} c_{ij} x_{ij} - \sum_{i \in V} \lambda_i \left(\sum_{j:(i,j) \in E} x_{ij} - \sum_{j:(j,i) \in E} x_{ji} \right) = \sum_{(i,j) \in E} (c_{ij} - \lambda_i + \lambda_j) x_{ij}.$$

If the Lagrangian is minimized subject to the regional constraints $\underline{m}_{ij} \leqslant x_{ij} \leqslant \overline{m}_{ij}$ for $(i,j) \in E$, Theorem 2.1 yields a set of conditions that are sufficient for optimality. It will be instructive to prove this result directly.

THEOREM 10.1. *Consider a feasible flow $x \in \mathbb{R}^{n \times n}$ for a circulation problem, and let $\lambda \in \mathbb{R}^n$ such that*

$$\begin{aligned} c_{ij} - \lambda_i + \lambda_j > 0 &\quad \textit{implies} \quad x_{ij} = \underline{m}_{ij}, \\ c_{ij} - \lambda_i + \lambda_j < 0 &\quad \textit{implies} \quad x_{ij} = \overline{m}_{ij}, \textit{ and} \\ \underline{m}_{ij} < x_{ij} < \overline{m}_{ij} &\quad \textit{implies} \quad c_{ij} - \lambda_i + \lambda_j = 0. \end{aligned}$$

Then x is optimal.

Proof. For $(i,j) \in E$, let $\bar{c}_{ij} = c_{ij} - \lambda_i + \lambda_j$. Then, for every feasible flow x',

$$\begin{aligned}
\sum_{(i,j)\in E} c_{ij} x'_{ij} &= \sum_{(i,j)\in E} c_{ij} x'_{ij} - \sum_{i \in V} \lambda_i \left(\sum_{j:(i,j)\in E} x'_{ij} - \sum_{j:(j,i)\in E} x'_{ji} \right) \\
&= \sum_{(i,j)\in E} \bar{c}_{ij} x'_{ij} \\
&\geqslant \sum_{\substack{(i,j)\in E \\ \bar{c}_{ij}<0}} \bar{c}_{ij} \overline{m}_{ij} + \sum_{\substack{(i,j)\in E \\ \bar{c}_{ij}>0}} \bar{c}_{ij} \underline{m}_{ij} \\
&= \sum_{(i,j)\in E} \bar{c}_{ij} x_{ij} = \sum_{(i,j)\in E} c_{ij} x_{ij}
\end{aligned}$$

□

The Lagrange multiplier λ_i is also referred to as a *node number*, or as a *potential* associated with vertex $i \in V$. Since only the difference between pairs of Lagrange multipliers appears in the optimality conditions, we can set $\lambda_1 = 0$ without loss of generality.

10.4 The Transportation Problem

An important special case of the minimum-cost flow problem is the *transportation problem*, where we are given a set of *suppliers* $i = 1, \ldots, n$ producing s_i units of a good and a set of *consumers* $j = 1, \ldots, m$ with demands d_j such that $\sum_{i=1}^{n} s_i = \sum_{j=1}^{m} d_j$. The cost of transporting one unit of the good from supplier i to consumer j is c_{ij}, and the goal is to match supply and demand while minimizing overall transportation cost. This can be formulated as an uncapacitated minimum-cost flow problem on a *bipartite network*, i.e., a network $G = (S \uplus C, E)$ with $S = \{1, \ldots, n\}$, $C = \{1, \ldots, m\}$, and $E \subseteq S \times C$. As far as optimal solutions are concerned, edges not contained in E are equivalent to edges with a very large cost. We can thus restrict our attention to the case where $E = S \times C$, known as the *Hitchcock transportation problem*:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{m} x_{ij} = s_i \qquad \text{for } i = 1, \ldots, n \\
& \sum_{i=1}^{n} x_{ij} = d_j \qquad \text{for } j = 1, \ldots, m \\
& x \geqslant 0.
\end{aligned}$$

It turns out that the transportation problem already captures the full expressiveness of the minimum-cost flow problem.

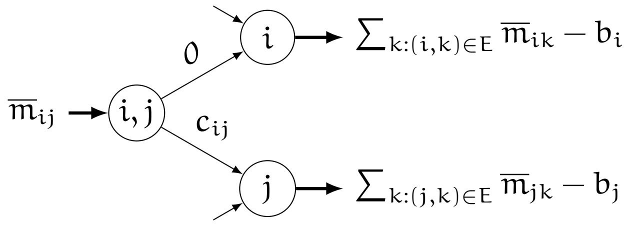

Figure 10.1: Representation of flow conservation constraints by an instance of the transportation problem

THEOREM 10.2. *Every minimum-cost flow problem with finite capacities or non-negative costs has an equivalent transportation problem.*

Proof. Consider a minimum-cost flow problem for a network (V, E) and assume without loss of generality that $\underline{m}_{ij} = 0$ for all $(i, j) \in E$. If this is not the case, we can instead consider the problem obtained by setting $\underline{m}_{ij}$ to zero, $\overline{m}_{ij}$ to $\overline{m}_{ij} - \underline{m}_{ij}$, and replacing b_i by $b_i - \underline{m}_{ij}$ and b_j by $b_j + \underline{m}_{ij}$. A solution with flow x_{ij} for the new problem then corresponds to a solution with flow $x_{ij} + \underline{m}_{ij}$ for the original problem. We can further assume that all capacities are finite: if some edge has infinite capacity but costs are non-negative then setting the capacity of this edge to a large enough number, for example $\sum_{i\in V} |b_i|$, does not affect the optimal solution of the problem.

We now construct an instance of the transportation problem as follows. For every vertex $i \in V$, we add a consumer with demand $\sum_k \overline{m}_{ik} - b_i$. For every edge $(i, j) \in E$, we add a supplier with supply $\overline{m}_{ij}$, an edge to vertex i with cost $c_{ij,j} = 0$, and an edge to vertex j with cost $c_{ij,j} = c_{ij}$. The situation is shown in Figure 10.1.

We now claim that there exists a direct correspondence between feasible flows of the two problems, and that these flows have the same costs. To see this, let the flows on edges (ij, i) and (ij, j) be $\overline{m}_{ij} - x_{ij}$ and x_{ij}, respectively. The total flow into vertex i then is $\sum_{k:(i,k)\in E}(\overline{m}_{ik} - x_{ik}) + \sum_{k:(k,i)\in E} x_{ki}$, which must be equal to $\sum_{k:(i,k)\in E} \overline{m}_{ik} - b_i$. This is the case if and only if $b_i + \sum_{k:(k,i)\in E} x_{ki} - \sum_{k:(i,k)\in E} x_{ik} = 0$, which is the flow conservation constraint for vertex i in the original problem. $\square$