An Improved Method for Sensitivity Analysis in Minimum Cost Flow Problem

Mahdalena, Herman Mawengkang
Graduate School of Mathematics, University of Sumatera Utara
Campus USU, Medan, Indonesia - 20155

Abstract - The minimum cost flow problem is a combinatorial optimization model which can be modelled as a linear programming problem. This paper proposed a sensitivity analysis method for minimum cost flow problem, by exploring the bound constraint structure, such that maintains the structure of the optimal solution. We improved a sensitivity analysis method which is formerly applicable to a tree solution. the proposed method preserves solution of upper bounds at upper bounds and those of lower bounds at lower bounds.

Keywords: Sensitivity analysis, Network, Shortest path, Integer programming, Bound constraints

1. INTRODUCTION

Given a connected and directed graph G = (V,A), where \( V = (v_1, v_2, ..., v_n) \) is a set of nodes, and A is a set of directed arcs, the cost flow problem can be defined as follows. Let \( c_{ij} \) be the cost per unit flow for each arc \((i, j) \in A\). We assume that the flow cost varies linearly with the amount of flow. We also associate with each arc \((i, j) \in A\) a capacity \(a_{ij}\) that denotes the maximum amount that can flow on the arc and lower bound \(l_{ij}\) that denotes the minimum amount that must flow \(m\) the arc. We associate with each node \(i \in A\) an integer value \(b_i\) representing its supply/demand. If \(b_i > 0\), node \(i\) is a supply node; if \(b_i < 0\) node \(i\) is a demand node with a demand of \(-b_i\); and if \(b_i = 0\) node \(i\) is a transhipment node. The decision variables \(m\) the minimum cost flow problem are arc flow which is represented by \(x_{ij}\) on \((i, j) \in A\). The model of the minimum cost flow problem (MCFP) can be expressed as follow.

\[
\begin{align*}
\text{min} \quad & \sum_{(i,j) \in A} c_{ij} x_{ij} \\
\text{subject to} \quad & \sum_{j:(i,j) \in A} x_{ij} - \sum_{j:(j,i) \in A} x_{ji} = b_i, \forall i \in V \\
& l_{ij} \leq x_{ij} \leq a_{ij}, \forall (i,j) \in A \\
& \sum_{j=1}^{n} b_i = 0
\end{align*}
\]

All network problems are special cases of the minimum cost flow problem. Like the maximum flow problem, it considers flows in networks with capacities. Like the shortest path problem, it considers a cost for flow through an arc. Like the transportation problem, it allows multiple sources and destinations. Therefore, all of these problems can be seen as special cases of the minimum cost flow problem.

As we can see that the model \((P)\) is in the form of linear programming. The parameter values and assumption in the model are subject to change and error. Therefore we would like to consider sensitivity analysis with respect to changes in the cost coefficients in problem \((P)\). Sensitivity analysis (SA), broadly defined, is the investigation of the potential changes and errors and their impacts on conclusions to be drawn from the model [1]. The main idea of sensitivity analysis is to determine changes in the optimal solution of the problem resulting from changes in the data (supply/demand, capacity or cost of any arc) [10, 5, 2, 7, 9, 6, 4]. Traditionally, researchers have induced this sensitivity analysis using primal and dual simplex method [2]. There is, however a conceptual drawback to this approach. The simplex based approach maintains a basis free at every iteration and conduct sensitivity analysis by determining changes in the basis free precipitated by changes in the data. The basis in the simplex method is often degenerate and consequently changes in the basis tree are not necessarily translated into the changes in the optimal solution. The other thing is the relative inefficiency and complexity of the simplex methods (primal, dual, and other variations) for network models. Therefore, the simplex based approach does not give information about the changes in the solution as the data changes; instead it tells us about the changes in the basis free [8, 3, 11, 2].

The main idea for avoiding the shortcomings is to explore the bound structure given to the decision variables \(x_{ij}\). Let \(x_{ij}^*\) be an optimal solution of the minimum cost flow problem. Regarding to the bound constraint \(l_{ij} \leq x_{ij}^* \leq a_{ij}\) the arc set \(A\) is partitioned in to three subsets as follows. Subset \(P\) contains \((i, j) \in A\), such that,
l_{ij} < x^*_{ij} < a_{ij}, subset Q contains (i, j) \in A, such that, x^*_j = l_{ij}, and subset R contains (i, j) \in A, such that, x^*_i = a_{ij}

We call the triple \((P, Q, R)\) as the optional solution structure.

**DEFINITION 1.1.** Given an optimum solution \(x^*_{ij}\) of the minimum cost flow problem, SA is to determined changes in the data that the optimal solution structure remains unchanged.

2. CHANGING THE COST COEFFICIENTS

Here we consider SA with respect to changes in the cost coefficients. Suppose first that an optimal solution \(x^*_{ij}\) of the \(MCFP\) is given. If the cost \(c_{ij}\) of an arc \((i, j)\) changes to \(\hat{c}_{ij}\), then its corresponding optional solution \(x^*_{ij}\) or the optional solution structure \((P, Q, R)\) may also change. Hence, cost sensitivity analysis should determine changes in the cost of any arc such that the given optional solution structure \((P, Q, R)\) is unchanged. First, given a feasible solution \(x_0\) of the \(MCFP\), \(x_0\) is an optimum solution of the \(MCFP\) if and only if for some set of node potentials \(w\), the reduced costs and flow values satisfy the following complementary slackness optimality condition for every arc \((i, j)\) in \(A\) [8, 5, 2]

\[
\begin{align*}
\text{If } x_{ij}^* = l_{ij}, & \text{ the } \hat{c}_{ij} - w_i^* - w_j^* \geq 0, \text{the } x_{ij}^* = l_{ij}; \\
\text{If } x_{ij}^* = u_{ij}, & \text{ the } \hat{c}_{ij} - w_i^* - w_j^* < 0, \text{the } x_{ij}^* = u_{ij}; \\
\text{If } x_{ij}^* = l_{ij} - w_i^* + w_j^* = 0, & \text{the } x_{ij}^* \leq l_{ij} \leq u_{ij}.
\end{align*}
\]

Thus, given an optimum solution \(x^*_{ij}\) of the \(MCFP\), SA with respect to changes in the cost coefficient in equivalent to determining cost range \(c_{ij}\) satisfying the following conditions:

\[
\begin{align*}
\text{If } x_{ij}^* = l_{ij}, & \text{ the } \hat{c}_{ij} - w_i^* - w_j^* \geq 0 \quad (8) \\
\text{If } x_{ij}^* = u_{ij}, & \text{ the } \hat{c}_{ij} - w_i^* - w_j^* < 0 \quad (9) \\
\text{If } l_{ij} \leq x_{ij}^*, & \text{the } \hat{c}_{ij} - w_i^* - w_j^* = 0 \quad (10)
\end{align*}
\]

Now, we consider a concept for the method of calculating SA for cost coefficient. Given an optimal solution \(x^*_{ij}\) of the \(MCFP\), suppose that the cost \(c_{ij}\) of an arc \((i, j)\) is changed to \(\hat{c}_{ij}\). If there is a path \(P(i, j)\) that can reroute the flow \(x^*_{ij}\) from node \(i\) to node \(j\) with less cost than \(c_{ij}\) without violating any of the optimality conditions, we should reroute the flow \(x^*_{ij}\) along \(P(i, j)\). This rerouting changes the optimal solution structure \((P, Q, R)\) due to a change in arc flow. Thus range of cost coefficients that maintains the optimal structure \((P, Q, R)\) is the range right before the point where the flow of arc is changed.

Now suppose that the cost \(c_{ij}\) of an arc \((i, j)\) increases by \(\hat{c}_{ij} = c_{ij} + \theta\). Then we want to know the range of \(\theta\), that is \(0 \leq \theta \leq \theta_0\), where \(\theta_0\) is the amount of maximum increasing (decreasing) flow that preserves the optimal solution structure \((P, Q, R)\). We first consider the following network \(G = (V, A)\) with multiple path \(P(i, j)\) from node \(i\) to node \(j\). Suppose that the cost \(c_{ij}\) of an arc \((i, j)\) is changed to \(\hat{c}_{ij}\). This changes the reduced cost optimality conditions to restore the optimality condition of the arc, we must change the flow on arc \((i, j)\). We can both increase and decrease the flow in an arc \((i, j)\) while knowing the bounds an arcs flows. However, in an arc \((i, h)\) at its lower bound (i.e., \(x_{ih} = l_{ih}\)) we can only increase the flow, and for flow on an arc \((i, m)\) at its upper bound (i.e., \(x_{im} = u_{im}\)) we can only decrease the flow.

Using the three concepts, we can find two paths \(P^1(i, j) = [(i, h), (h, j)]\) and \(P^2(i, j) = [(k, i), (k, j)]\) that can send the flow from node \(j\). In our example, per unit cost to send one unit of flow along the path \(P^1(i, j)\) is \(c_{ih} + c_{hj}\). If \(c_{ih} + c_{hj} < \hat{c}_{ij}\), then the flow is sent along \(P^1(i, j)\). In this case the optimal structure \((P, Q, R)\) may change. But, if the per unit cost to send one unit of flow along \(P^1(i, j)\) is greater than the per unit cost \(\hat{c}_{ij}\), the flow is not sent along \(P^1(i, j)\). In this case since the flow along \(P^1(i, j)\) is not created, we can maintain the given optimal solution structure \((P, Q, R)\) in the interval of \(\theta \leq c_{ij} + e_{ih} + e_{hj}\). Similarly, \(\theta^1_u\) the per unit cost to send one unit of flow along \(P^2(i, j)\), is \(-c_{ih} + c_{hj}\). If \(-c_{ih} + c_{hj} < \hat{c}_{ij}\) the flow will not be sent along \(P^2(i, j)\) and the structure \((P, Q, R)\) can be preserved in the interval of \(\theta \leq c_{ij} + e_{ih} + e_{hj}\). Here, let \(\theta^1_u\) and \(\theta^2_u\) be the upper bound of \(\theta\) with respect to the path \(P^1(i, j)\) and \(P^2(i, j)\) can preserve the given structure \((P, Q, R)\). Then \(\theta^1_u = -e_{ij} + c_{ih} + c_{hj}\) and \(\theta^2_u = -e_{ij} - e_{im} + e_{nj}\). Consequently, \(\theta_u\) that maintains the optimal structure

\[
\theta_u = \min(\theta^1_u, \theta^2_u)
\]
\((P,Q,R)\) is restricted in value of \(\theta_a = \min(\theta_a^1, \theta_a^2)\). In order to do these calculations easily consider a definition of transformed network.

DEFINITION 2.1. The transformed network \(G'=(V,A')\) corresponding to a network \(G=(V,A)\) of the MCFP is defined as follows. We replace each arc \((i,j)\in A\) with flow \(l_{ij} < x_{ij}^* < u_{ij}\) by two arcs \((i,j)\) and \((j,i)\) with the cost \(\overline{c}_{ij}=c_{ij}\) and \(\overline{c}_{ji}=-c_{ji}\) respectively. We also replace each arc \((i,j)\in A\) with the flow \(x_{ij}^* = u_{ij}\) by arc \((j,i)\) with the cost \(\overline{c}_{ij} = -c_{ji}\). Finally, we replace each arc \((i,j)\in A\) with the flow \(x_{ij}^* = l_{ij}\) by arc \((i,j)\) with the cost \(\overline{c}_{ij} = c_{ij}\).

Using the above transformed network \(G'=(V,A')\), we can calculate \(\theta_a\) by summing up the length of the directed path \(P(i,j)\) from node \(i\) to node \(j\) and \(-c_{ij}\) in \(G'=(V,A')\). If there exists multiple directed paths \(P(i,j)\) form node \(i\) to node \(j\), calculate \(\theta_a\) by summing up the minimum length of the directed path from node \(i\) to node \(j\) and \(-c_{ij}\) by summing the minimum length of the directed path among various directed path and \(-c_{ij}\) in \(G'=(V,A')\). Since among the various directed path the length of the directed path associated with the shortest path from node \(i\) to node \(j\) is the minimum, we need to calculate the length of the shortest path for the calculation of \(\theta_a\) [12].

LEMMA 2.1. The transformed network \(G'=(V,A')\) contains a negative cycle.

PROOF. Assume the transformed network \(G'=(V,A')\) contains a negative cycle. This means that the network \(G=(V,A)\) does not satisfy the optimality condition, and we can improve the current optimal solution with respect to this negative cycle. Because a feasible solution which is less than the objective function value of the given optimal solution \(x_{ij}^*\) exists, it contradicts the assumption that the optimum solution in given. Therefore, the transformed network \(G'=(V,A')\) contains no negative cycles.

LEMMA 2.2. If the directed path \(P(i,j)\) from node \(i\) to node \(j\) in \(G'=(V,A')\) does not exist, there is no path \(P(i,j)\) that can send the flow from node \(i\) to node \(j\) in the network \(G=(V,A)\). Therefore, even if the cost \(c_{ij}\) of an arc \((i,j)\) increases infinitely, it is optimal to send the current flow \(x_{ij}^*\) along the arc \((i,j)\). So if the directed path \(P(i,j)\) does not exist in \(G'=(V,A')\), the \(\theta_a = \infty\).

Ultimately, if the shortest path exist in \(\theta_a = \infty\) we obtain a constant as the upper bound value \(\theta_a\). But, if the shortest path does not exist in \(G'=(V,A')\) we obtain an infinite value as the upper bound.

THEOREM 2.1. If \(G'=(V,A')\) contains a shortest path from node \(i\) to node \(j\) and \(l\) is the length of the shortest path, then \(\theta_a = l - c_{ij}\). If \(G'=(V,A')\) contain no shortest path from node \(i\) to node \(j\), then \(\theta_a = \infty\).

PROOF. By lemma 2.1, \(G'=(V,A')\) contains no negative cycle. Therefore if the shortest path exists in \(G'=(V,A')\), we can solve it. If the shortest \(G'=(V,A')\), the directed paths \(P(i,j)\) does not exist and by lemma 2.2. \(\theta_a = \infty\). If there are multiple directed path \(P(i,j)\) from node \(i\) to node \(j\), the \(\theta_a\) is obtained by summing the minimum length of the directed path among them, and \(-c_{ij}\) in \(G'=(V,A')\). Let \(P_h(i,j)\) be the length of the both directed path among the multiple directed paths. Then \(l = \min\{P_h(i,j)\}\) and \(\theta_a = l - c_{ij}\). Next we consider \(c_{ij} = c_{ij} - \theta\). Because per unit cost to send are unit of flow along the arc \((i,j)\) is decreased by \(\hat{c}_{ij}\), more flow along the arc \((i,j)\) instead of sending flow along the path from node \(i\) to node \(j\). Here, the meaning of sending more flow along arc \((i,j)\) is equivalent to sending the flow along the path from node \(j\) to node \(i\). Therefore in this case after comparing the per unit cost to send one unit of flow along the path from node \(j\) to node \(i\) with the per unit cost to send one unit of flow along the arc \((i,j)\) from node \(j\) to node \(i\), a flow along the path associated with the cheaper cost of the two cases is created. In this case, \(\theta_i\) is the sum of the length of the directed path \(P(j,i)\) from node \(j\) to
node $i$ and $c_{ij}$ in $G'=(V,A')$. If there are multiple directed paths $P(j,i)$ from node $j$ to node $i$, $\Theta_j$ is calculated by summing the minimum length of the directed path and $c_{ij}$ in $G'=(V,A')$. Here, let $P_h(i,j)$ be the length of the both directed path among the multiple directed paths. Then the length $l$, the shortest path from node $j$ to node $i$ is $l = \min \{P_h(j,i)\}$. Therefore, if the shortest path from node $j$ to node $i$ exists in $G'=(V,A')$ we obtain $\Theta_j = l + c_{ij}$. But if the shortest path does not exist in $G'=(V,A')$, $\Theta_j = \sim$ by lemma 2.2.

In case the number of arcs, with a flow of the $l_y < x_y < u_y$, is greater than or equal to $n-1$, we can easily compute the upper bound or the lower bound. Given a spanning tree optimal solution where the number of arcs with a flow of $l_y < x_y < u_y$ is exactly equal to $n-1$, we can obtain the node potentials $w$ associated with the spanning tree optimal solution because arcs with a flow of $l_y < x_y < u_y$ consist of the optimum basis of the network $G=(V,A)$. By using node potentials $w$ we can compute the length of the shortest path more easily because the node potentials $w$ is the length of the tree path with respect to the optimum basis $O$. We call this case as CS1.

**THEOREM 2.2.** If $O$ is an optimal basis and $w$ is the node potentials with respect to $O$, then the length of the tree path $P(r,s)$ from node $r$ to node $s$ is $w_r - w_s$.

**PROOF.** We compute $P(r,s)$ using the fact that $\tilde{c}_{ij} = c_{ij} - w_i + w_j = 0$, for every arc $(i,j) \in O$.

$$
\tilde{c}_{ij} = \sum_{(i,j) \in P(r,s)} (c_{ij} - w_i + w_j)
= \sum_{(i,j) \in P(r,s)} c_{ij} - \sum_{(i,j) \in P(r,s)} (w_i - w_j)
= \sum_{(i,j) \in P(r,s)} c_{ij} - w_i + w_j
= 0
$$

because all $w$ corresponding to the nodes in the path, other than the terminal nodes $r$ and $s$, cancel each other in the expression $\sum_{(i,j) \in P(r,s)} (w_i - w_j)$ is associated with the length of path $P(r,s)$, the length of the tree path $P(r,s)$ is $w_r - w_s$.

Generally, given the optimal basis associated with spanning tree $O$, because the length of the shortest path from node $i$ to node $j$ is equivalent to the length of the tree path $P(i,j)$ from node $i$ to node $j$, we can compute the sensitivity analysis more easily by calculating the length of the tree path using the node potentials. This case is called CS2.

By this result, CS2 and CS1 are the same in spanning tree solution [2]. Given an optimal solution that the number of arcs with a flow of $l_y < x_y < u_y$ is greater than $n-1$, there arcs with the intermediate flow always form cycle. Let $C$ be a cycle that consist of arcs with intermediate flow, $l_y < x_y < u_y$. Then the results of the sensitivity analysis with respect to an arc $(i,j)$ that belong to cycle $C$ is given in theorem 2.3.

**THEOREM 2.3.** Given a non-tree optimal solution where the number of arcs with a flow of $l_y < x_y < u_y$ is greater than $n-1$, the range $\Theta$ of CS2 with respect to an arc $(i,j)$ that belongs to a cycle $C$ is zero.

**PROOF.** Given a non-tree optimal solution arcs with intermediate flow always form cycles. Let $C$ be a cycle that consists of arcs with the intermediate flow. Then $\sum_{(i,j) \in C} c_{ij} = 0$ if the cost of $c_{ij}$ of an arc $(i,j)$ is changed to $c_{ij} = c_{ij} + \Theta$, then the value of the cycle $C$ is

$$
(c_{ij} + \Theta) = \sum_{(i,j) \in C} c_{ij} = \sum_{(i,j) \in C} c_{ij} + \Theta = \Theta (11)
$$

If $\Theta > 0$ (or $\Theta < 0$) then it is better to send the flow in the opposite direction (or same direction) of cycle to improve the objective function value. In this case, since the flow along cycle $C$ is created, the given optimal solution structure $(P,Q,R)$ may change. Therefore in order to maintain the given optimal solution structure $(P,Q,R)$, the range of $\Theta$ has to be zero. □

3. CONCLUSIONS

In this paper, we have categorized the sensitivity analysis of the minimum cost flow problem into two types. We define CS1 to be the acquirement of a region where the given optimum basis is unchanged. This is the well known method applicable to a tree solution. However CS1 can not be applied to a non-tree solution or a degenerate tree solution. So we proposed the CS2 that finds the region where the upper bound valued arcs in the optimum solution maintain upper bound valued, low bound valued arcs maintain lower bound valued, and intermediate valued arcs maintain intermediate valued. This CS2 provides the
generalized concept of the sensitivity analysis for the optimum solution of the minimum cost flow problem.

4. REFERENCES