Chapter 8, Operations Research (OR)

Kent Andersen

February 7, 2007

1 The transportation problem

In the last chapter, we introduced the transportation problem (P).

$$\begin{aligned}
\text{Minimize } Z = \sum_{i\in S}\sum_{j\in D} c_{i,j}x_{i,j}& \\
\text{s.t.}& \\
\sum_{j\in D} x_{i,j} = s_i, &\qquad \text{for all } i \in S \qquad (u_i)\\
\sum_{i\in S} x_{i,j} = d_j, &\qquad \text{for all } j \in D \qquad (v_j)\\
x_{i,j} \geq 0.\ \text{for all } (i,j) \in S\times D&
\end{aligned}$$

where we have now written dual variables $\{u_i\}_{i\in S}$ and $\{v_j\}_{j\in D}$ next to the constraints. The dual (D) of (P) is given as follows.

$$\begin{aligned}
\text{Maximize } Z = \sum_{i\in S} u_i s_i + \sum_{j\in D} v_j d_j& \\
\text{s.t.}& \\
u_i + v_j \leq c_{i,j}, &\qquad \text{for all } (i,j) \in S\times D \quad (x_{i,j})\\
u_i, v_j \in \mathbb{R}&
\end{aligned}$$

where, similarly, the dual variables $x_{i,j}$ are written next to the dual constraints.

2 Graphs and the transportation problems

The transportation problem is best represented by a *graph* (see Fig. 1). A graph G is a pair $G = (V, E)$ of *nodes* (also called vertices) V and *edges* E. Each node $v \in V$ of a graph is represented as a point in the plane. In Fig. 1, there is a node for every source and a node for every destination. An edge is a *pair* of nodes (v_1, v_2), and an edge is represented in the plane by a line between the two vertices. In Figure 1, we have drawn an edge between the source node 2 and the destination node 3'. For the transportation problem, a variable $x_{i,j}$ can be associated with every edge $(i,j) \in S \times D$, and this edge indicates transportation from i to j.

Hence, for the graph representation of the transportation problem, the set $S \cup D$ represents the nodes, and for every pair $(i,j) \in S \times D$, there is an edge between source node i and destination node j. Observe that there is no edge between two source nodes or two destination nodes.

The understanding of the transportation problem requires the following concepts from graph theory. A *cycle* in a graph is a sequence of connected nodes that form a cycle. For instance, for

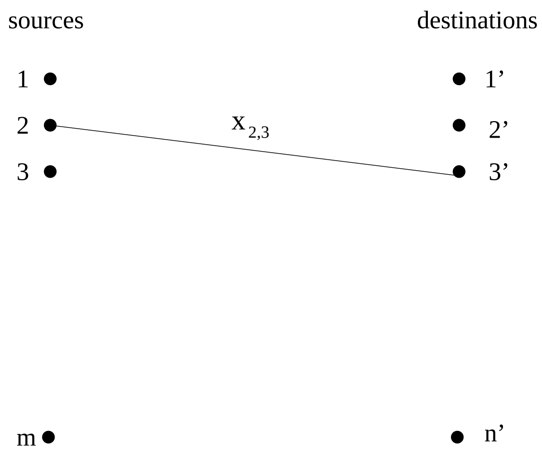

Figure 1: Graph of the transportation problem

the transportation graph, the sequence of nodes given by $1 - 1' - 3 - 5' - 7 - 2' - 1$ forms a cycle. Similarly, a *path* in a graph between nodes u and v is a sequence of nodes that lead from u to v. In the transportation problem, the sequence of nodes given by $1 - 1' - 3 - 5'$ is a path between nodes 1 and 5'. An (edge-induced) *subgraph* of G is a graph obtained from G by deleting some edges of G.

A graph G is said to be a *tree*, if G does not contain any cycles. A graph, which is a tree, is called a *spanning tree*, if every node is connected by some edge.

We first relate the basic solutions of the transportation problem with trees of the graph of the transportation problem

Lemma 1 *Let $B \subseteq S \times D$ denote a basis for the transportation problem. Then the subgraph of the transportation graph obtained by* only *including the edges $(i, j) \in B$ forms a spanning tree.*

Proof: Let x^B be a basic solution for the transportation problem. Then x^B must be the unique solution to the system:

(i) $\sum_{j \in D:(i,j) \in B} x_{i,j} = s_i$ for all $i \in S$.

(ii) $\sum_{i \in S:(i,j) \in B} x_{i,j} = d_j$ for all $j \in D$.

(iii) $x_{i,j} = 0$ for all $(i, j) \in S \times D \setminus B$.

If for some $i \in S$ we have $\{j \in D : (i, j) \in B\} = \emptyset$, then the i^{th} equality reads $0 = s_i \neq 0$, so all nodes $i \in S$ must be connected by some edge $(i, j) \in B$. Similarly, every node $j \in D$ must be connected by some edge $(i, j) \in B$. Hence every node in $S \cup D$ is connected by some edge. Suppose $i_1 - j_1 - i_2 - j_2 - \ldots - i_k - j_k - i_1$ forms a cycle in the graph with nodes $S \cup D$ and edges indexed by B, where $\{i_1, i_2, \ldots, i_k\} \subseteq S$ and $\{j_1, j_2, \ldots, j_k\} \subseteq D$. Observe that this cycle must have an even number of edges. Hence, if we choose $\epsilon > 0$, and let $x'_{i_1,j_1} := x^B_{i_1,j_1} + \epsilon$, $x'_{j_1,i_2} := x^B_{j_1,i_2} - \epsilon$, $x'_{i_2,j_2} := x^B_{i_2,j_2} + \epsilon$, ..., $x'_{i_k,j_k} := x^B_{i_k,j_k} + \epsilon$ and $x'_{j_k,i_1} := x^B_{j_k,i_1} - \epsilon$, then x' is also a solution to the system (i)-(iii). However, this contradicts the assumption that the solution to (i)-(iii) is unique.■

Having characterized the basic solutions of the transportation problem as trees in a graph, we now give some further properties of trees. We first prove that a tree always has a node with exactly one edge connected to it

Lemma 2 *Let $G = (V, E)$ be a graph which is a tree, and suppose $E \neq \emptyset$. Then there is a node $v \in V$ with exactly one edge connected to it.*

Proof: Let $i_1 \in V$ be an arbitrary node. If i_1 only has one edge connected to it, we are done, so we can assume (i_1, i_2) is an edge for some $i_2 \in V \setminus \{i_1\}$. If this is the only edge connected to i_2, we are done, so suppose i_2 is connected to $i_3 \in V \setminus \{i_2\}$. If $i_3 = i_1$, we have a cycle, so we can assume $i_3 \neq i_1$. If i_3 is only connected to i_2, we are done, so we can assume i_3 is connected to some $i_4 \in V \setminus \{i_3\}$. Again, i_4 can not be any of the previously chosen vertices, since then we would have a cycle. This process must end at some point, so either we obtain a node with exactly one edge connected to it. ■

The number of edges connected to a node in a graph is called the *degree* of the node. Lemma 2 shows that a tree always has a node of degree 1.

Observe that, if $G = (V, E)$ forms a tree, then the subgraph $G' = (V, E \setminus \{(i, j)\})$ is also a tree for every $(i, j) \in E$. Why? Well - if the graph G' has a cycle, then clearly G also has a cycle. From this property, we immediately have the following result.

Lemma 3 *Let $G = (V, E)$ be a graph that forms a spanning tree. Then $|E| = |V| - 1$.*

Proof: Let $i_1 \in V$ be a node of degree 1. Consider the spanning tree obtained from G by deleting i_1 and the edge connected to i_1. Since this tree also has a node of degree 1, we can select a node $i_2 \in V \setminus \{i_1\}$ of degree 1 in this graph. We can then delete i_2 and the corresponding edge. This process can be continued until we are left with two nodes, and we will have deleted $|V| - 2$ nodes, and therefore also $|V| - 2$ edges. Since this remaining graph with two nodes is a spanning tree, it will contain exactly one edge, and therefore $|E| = |V| - 1$. ■

Combining what we have shown above, we now have

Every basis B of the transportation problem contains exactly $m + n - 1$ variables

It is possible to show that the converse of Lemma 1 holds. In other words, if we have a spanning tree in the graph with nodes $S \cup D$ and edges $(i, j) \in S \times D$, then we have a basis for the transportation problem - that is, the set of edges B of edges $(i, j) \in S \times D$ that form the spanning tree defines a basis for the transportation problem (P). We illustrate this on an example.

Example 1 *Consider the transportation problem with supply data $(s_1, s_2, s_3, s_4) = (50, 50, 50, 50)$ and demand data $(d_1, d_2, d_3, d_4, d_5) = (30, 20, 70, 30, 60)$ ($m = 4$ and $n = 5$). Also, consider the eight variables $B = \{(1, 1), (1, 2), (2, 2), (2, 3), (3, 3), (3, 4), (3, 5), (4, 5)\}$. The corresponding graph is shown in Fig. 2.*

Clearly the set B forms a spanning tree. Why is the solution to the subsystem obtained from these variables unique. First select a node of degree 1. We arbitrarily select node 1' (we could also have chosen node 4). Since $(1, 1)$ is the only pair in B that connects destination 1, source 1 has to supply all the demand to destination 1. Hence $x^B_{1,1} = d_1 = 30$ (the equality corresponding to destination 1 reads $x_{1,1} = 30$ when deleting all non-basic variables from that constraint). Now delete destination 1 from the graph, delete the edge $(1, 1)$ and update the possible supply from source 1 to $s_1 := s_1 - 30 = 20$. We are now left with a spanning tree on 8 nodes. Next select a node in this remaining graph of degree 1, say source node 4. The only destination node that source 4 can deliver to is destination node 5. Hence destination 5 has to receive all the supply from source 4, so we must have $x^B_{4,5} = s_4 = 50$. After this, we delete source node 4 from the graph, we delete the edge $(4, 5)$, and we finally update the demand of destination node 5 to $d_5 = d_5 - 50 = 10$.

Since a spanning tree always remains after the reduction of the graph, we can continue this process. This then leads to $x^B_{3,5} = 10$, $x^B_{3,4} = 30$, $x^B_{1,2} = 20$, $x^B_{2,2} = 0$, $x^B_{2,3} = 60$ and $x^B_{3,3} = 10$. In this case, at no point did we have to set a variable to a negative value, so this basic solution is feasible for the transportation problem.

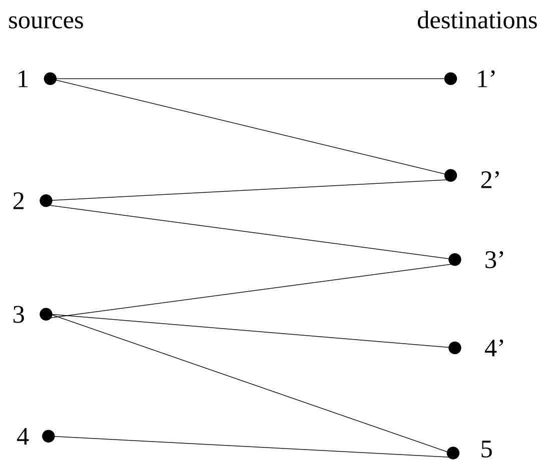

Figure 2: Graph of the transportation problem

3 Computing a dual basic solution

When we presented the simplex method, this was under the assumption that the matrix A in the system "$Ax = b$" had full row rank. In the transportation problem, we have $n + m$ rows, and we just showed that a basis B of the transportation problem consisted of $n + m - 1$ variables, and *not* $n + m$. Why is this so? The reason is that the equality system in the transportation problem has exactly one redundant equality.

Lemma 4 *Consider the equality system $A'x = b'$ obtained from the equality system of the transportation by deleting one of the equalities, and assume that $\sum_{i\in S} s_i = \sum_{j\in D} d_j$. If x' satisfies the $(n + m - 1)$ equalities $A'x = b'$, then x' satisfies* all *equalities involved in the definition of the transportation problem.*

Proof: Without loss of generality, suppose the equality that is deleted is the equality $d_k = \sum_{i\in S} x_{i,k}$, and suppose x' satisfies all remaining equalities. Then we have

$$\begin{aligned} d_k &= \sum_{i\in S} s_i - \sum_{j\in D\setminus\{k\}} d_j \\ &= \sum_{i\in S}(\sum_{j\in D} x'_{i,j}) - \sum_{j\in D\setminus\{k\}}(\sum_{i\in S} x'_{i,j}) \\ &= \sum_{i\in S}(\sum_{j\in D} x'_{i,j}) - \sum_{j\in D\setminus\{k\}}(\sum_{i\in S} x'_{i,j}) + \sum_{i\in S} x'_{i,k} - \sum_{i\in S} x'_{i,k} \\ &= \sum_{i\in S}(\sum_{j\in D} x'_{i,j}) - \sum_{j\in D}(\sum_{i\in S} x'_{i,j}) + \sum_{i\in S} x'_{i,k} \\ &= \sum_{i\in S} x'_{i,k}, \end{aligned}$$

and therefore x' also satisfies the missing equality. ■

So what does it mean that one of the equalities in (P) can be eliminated? Essentially, it means that we could have deleted one of the equalities of (P) before creating the dual, and therefore, we can ignore (or fix to zero) one of the dual variables. For simplicity, suppose we fix $v_n = 0$. Given a basis $B \subseteq S \times D$ of (P), the corresponding dual basic solution can be found by solving the following non-singular system with $m + n$ variables and $m + n$ equalities.

(i) $u_i + v_j = c_{i,j}$ for all $(i,j) \in B$, and

(ii) $v_n = 0$.

Like in the computation of the primal basic solution x^B, the fact that B corresponds to a spanning tree can be exploited to compute the corresponding dual basic solution (u^B, v^B). In other words, due to the special structure of the transportation problem, it is *not* necessary to use the more complicated (and more general) procedure of Gaussian elimination. We illustrate this on the data of example 1.

Example 2 *Consider the transportation problem of example 1, and the same basis B (see Fig. 2). Also, suppose the cost coefficients on the basic variables are given by $c_{1,1} = 16$, $c_{1,2} = 16$, $c_{2,2} = 14$, $c_{2,3} = 13$, $c_{3,3} = 20$, $c_{3,4} = 23$, $c_{3,5} = 100$ and $c_{4,5} = 0$. To compute the dual basic solution corresponding to B, the following system needs to be solved.*

(a) $u_1 + v_1 = 16$,

(b) $u_1 + v_2 = 16$,

(c) $u_2 + v_2 = 14$,

(d) $u_2 + v_3 = 13$,

(e) $u_3 + v_3 = 20$,

(f) $u_3 + v_4 = 23$,

(g) $u_3 + v_5 = 100$,

(h) $u_4 + v_5 = 0$, *and*

(i) $v_5 = 0$

We first exploit the fact that we have fixed v_5^B to zero. Therefore, for all source nodes that are connected to destination node 5 by the edges in B, the corresponding equalities in (a)-(i) now only have one variable. In other words, we obtain $u_4^B = 0$ and $u_3^B = 100$. Proceeding in this manner, we obtain $v_3^B = -80$, $v_4^B = -77$, $u_2^B = 93$, $v_2^B = -79$, $u_1^B = 95$ and $v_1^B = -79$. The fact that B defines a spanning tree ensures that this process finishes and gives the values of the dual variables.

It is important to note that, in order to compute the basic solution x^B for (P), and the dual basic solution (u^B, v^B) for (D), only additions and subtractions of numbers are required. An algorithm with this property is called a *combinatorial algorithm*. It immediately implies the following very important property of the transportation problem.

Lemma 5 *Assume all the data $\{c_{i,j}\}_{(i,j)\in S\times D}$, $\{s_i\}_{i\in S}$ and $\{d_j\}_{j\in D}$ are integer numbers. Then, for every basis B, the basic solution x^B for (P), and the dual basic solution (u^B, v^B) for (D) is integer.*

4 Testing optimality

Given a basic solution x^B for (P), and a basic solution (u^B, v^B) for (D), when do we know that B is an optimal basis? As shown earlier, this is the case when x^B is feasible for (P) and (u^B, v^B) is feasible for (D). Testing feasibility of x^B for (P) simply amounts to checking whether x^B is non-negative. For feasibility of (u^B, v^B) for (D), we need to compute the reduced costs $r_{i,j}^B$ on the variables $x_{i,j}$

$$r_{i,j}^B := c_{i,j} - u_i^B - v_j^B$$

If $r^B_{i,j} \geq 0$ for all variables, then the current basis B is optimal. Otherwise there exists some non-basic variable (i_1, j_1) such that $r^B_{i_1,j_1} < 0$, and it therefore "pays of" to increase the current value (of zero) of x_{i_1,j_1}. The variable x_{i_1,j_1} can then be the entering non-basic variable.

For our example, suppose the cost on the variable $(2, 5)$ is given by $c_{2,5} = 15$. The reduced cost on the variable $x_{2,5}$ is then given by $r^B_{2,5} = 15 - u^B_2 - v^B_5 = -78$. Since this value is negative, B is not optimal, and the variable $x_{2,5}$ can enter the basis.

5 Pivoting

Having decided that the non-basic variable x_{i_1,j_1} will enter the basis, how do we find the leaving basic variable? It turns out that the fact that B defines a spanning tree is very useful for this purpose. Observe that, in a spanning tree, there is a unique path between every pair of nodes. Indeed, if this was not the case, then it would be possible to join two paths to form a cycle, which contradicts the assumption that we had a spanning tree.

Hence, if we add the edge (i_1, j_1) to the spanning tree defined by B, a unique cycle will be created, namely the cycle obtained by joining the unique path from i_1 to j_1 with the edge (i_1, j_1). Furthermore, since every edge connects a source node with a destination node, this cycle will consist of an even number of edges. Let $i_1 - j_1 - i_2 - j_2 - \ldots - i_k - j_k - i_1$ denote the unique cycle obtained by adding the edge (i_1, j_1) to the spanning tree defined by B, where $(i_2, j_1), (i_2, j_2), \ldots, (i_k, j_k), (i_1, j_k) \in B$.

The leaving basic variable can now be found by examining this cycle. Increasing the value of x_{i_1,j_1} forces a decrease in the value of x_{i_2,j_1}, which forces an increase in the value of x_{i_2,j_2} and so on. Furthermore, these are the only changes that are forced on the basic variables by the increase in the value of x_{i_1,j_1}. Hence, to find the leaving basic variable, we only need to figure out which variable on this cycle becomes zero first. We illustrate this on our example.

Example 3 *Suppose we let the variable* $x_{2,5}$ *enter the basis. This creates the even cycle* $2 - 5' - 3 - 3' - 2$.

Hence, when we increase $x_{2,5}$ *to* $x^B_{2,5} + \delta > 0$*, we need to decrease* $x_{3,5}$ *to* $x^B_{3,5} - \delta$*, and we also need to decrease* $x_{2,3}$ *to* $x^B_{2,3} - \delta$*. Also, increasing* $x_{2,5}$ *by* δ *forces an increase of* δ *in the values of the variables* $x_{3,3}$ *and* $x_{2,2}$.

Since $10 = x^B_{3,5} < x^B_{2,3} = 60$*, we can increase* $x_{2,5}$ *by 10 units and still have a feasible solution. The variable* $x_{3,5}$ *will then get the value zero, and therefore leaves the basis. The new basis is therefore* $B' = (B \cup \{(2, 5)\}) \setminus \{(3, 5)\}$*, and the new basic solution becomes*

(i) $x^{B'}_{i,j} = 0$ *for all* $(i, j) \in S \times D \setminus B'$ *(all non-basic variables have value zero).*

(ii) $x^{B'}_{2,5} = 10$ *(value of the new basic variable).*

(iii) $x^{B'}_{2,3} = 60 - 10 = 50$

(iv) $x^{B'}_{2,2} = 0 + 10 = 10$.

(v) $x^{B'}_{3,3} = 10 + 10 = 20$.

Finally the cost of the basic solution corresponding to B' *is* $10|r^B_{2,5}| = 780$ *units less than the cost of the basic solution corresponding to* B.