[AM-05-042] The Classic Transportation Problem

Abstract

The classic transportation problem concerns minimizing the cost of transporting a product from sources/supplies to destinations/demands. It is a network-flow problem that arises in industrial logistics and is often solved by linear programming (LP). The three inputs of the model are total units produced at each source, total units needed at each destination, and the cost to transport one unit from each source to each destination. And the objective is to minimize the total cost of transporting all units produced at sources to meet the demands at destinations. The problem solution includes three basic steps: 1) finding an initial basic feasible solution, 2) checking if the current solution is optimal (with the lowest costs), and improving the current solution through iteration. Solving such a problem relies strongly on the network data models, least-cost path algorithms, other functionalities in GIS. And an integrated framework is often adopted to utilize both GIS and non-GIS linear programming solvers to search for the optimal solution.

Keywords: accessibility, costs, distance, flows, networks, transportation

Author & citation

Song, Y. (2021). The Classic Transportation Problem. The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2021 Edition), John P. Wilson (ed.).
DOI: 10.22224/gistbok/2021.3.5.

This Topic is also available in the following editions:

- Song, Y. (2017). The Classic Transportation Problem. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2017 Edition), John P. Wilson (ed.). DOI: 10.22224/gistbok/2017.4.7.
- DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., and Wentz, E. (2006). Approximating the Earth's shape with geoids. The Geographic Information Science & Technology Body of Knowledge. Washington, DC: Association of American Geographers.

Explanation

1. Definitions
2. Modeling the Classic Transportation Problem
3. Solving the Classic Transportation Problem
4. The Classic Transportation Problem and GIS

1. Definitions

Network: A network can be defined as a directed graph G = (V, A), where V is the set of vertices or nodes, and A is the set of arcs or links corresponding to the feasible routes between nodes. Each arc/link is often assigned with a weight reflecting the movement cost and some additional constraints such as flow capacities.

Linear Programming: A linear programming problem may be defined as a problem of maximizing/minimizing a linear function subject to linear constraints. The constraints can be either equalities or inequalities. And the objective is to find the optimal solution that minimizes/maximizes the linear function.

Industrial Logistics: In general, industrial logistics is the management of the resources between sources and destinations to meet the demands of the market. The resources include both physical items such as food and equipment and abstract ones such as time and resources.

Heuristic: A heuristic trades optimality, completeness, accuracy, or precision for efficiency while solving a problem. Based on available information, a heuristic ranks the alternatives at each step to decide which direction to follow. For instance, a greedy heuristic selects a local optimum value at each stage, attempting to find a global optimum.

2. Modeling the Classic Transportation Problem

The classic transportation problem has long been an indispensable research problem in industrial and transportation logistics. The problem is driven by the need to transport a single commodity from sources (plants and factories) to destinations (warehouses and consumer markets). The total units produced at each source, the total units required by each destination, the cost to transport one unit from each source to each destination are assumed to be known. And the objective is to minimize the total cost of shipping units from sources to meet demands at destinations. The problem is also driven by the need to move people between activity locations such as home and office. The trip distribution step in the traditional four-step modeling is a classic transportation problem that aims to minimize the overall cost (e.g. travel distance or time) for people to move between activity locations (e.g. home and office).



Figure 1. Transporting a commodity from sources to destinations, Si and Dj denotes units supplied and units demanded. Source: author.

The number of units supplied by source i is Si, for i = 1,2, ... , m. The number of units required by destination j is Dj, for j = 1,2, ... , n. The cost of transporting one unit along the link from source i to destination j is cij, and the number of units transported along this link is xij. The total cost of transporting units from source i to all destinations {j} is:



Similarly, the total cost of transporting units to destinations j from all sources {i} is:



Thus, the classic transportation problem can be formulated as a linear equation:

Equation (3) represents the total transportation cost and is the objective function of the classic transportation problem. Equation (4) indicates that the total units shipped from source i should not exceed the available units Si at source i. Equation (5) indicates that the total units shipped to destination j should meet its demand Dj.

The network model above can be balanced or unbalanced. When the total supply at all resources, $\sum_{i=1}^{m} S_i$, is equal to the total demand at all destinations, $\sum_{j=1}^{n} D_j$, the model is balanced. Otherwise, the model is unbalanced, with excess demand or excess supply. For the case of excessive demand, $\sum_{i=1}^{m} S_i < \sum_{j=1}^{n} D_j$, a dummy source is created to reflect the penalty cost, Dp, for each unit of unsatisfied demand. For the case of excess supply, $\sum_{i=1}^{m} S_i > \sum_{j=1}^{n} D_j$, a dummy warehouse is created to reflect the storage cost, Ss, for each unit of surplus production.

The classic transportation problem can also be formulated as matrices. Figure 2 shows a balanced network model with m source and n destinations.

Figure 2. A balanced network model formulated as a matrix. Source: author.

3. Solving the Classic Transportation Problem

The classic transportation problem can be formulated using Equations (3) to (6), which are all linear equations. Hence, the problem can be structured as a linear program and solved using the simplex method. The three steps are 1) finding an initial basic feasible solution, 2) checking if the solution is optimal, and 3) continuing to improve the solution.

3.1 Finding an Initial Basic Feasible Solution

For the balanced transportation problem, the constraint set in Equation (4) and (5) is reformulated as (m + n) linear questions:

Therefore, one of the equations in the constraint set is redundant, and the total number of independent equations becomes (m + n − 1). This is because, if ai, i = 1, 2, ..., m and bj, j = 1, 2, ..., (n − 1) are specified, bn can be calculated as $b_n = \sum^{m}{}_{i=1}\, a_i - \sum^{n-1}{}_{j=1}\, b_j$. In a linear program, the basic variables are variables that are not zeros, and the number of basic variables in a basic feasible solution is the same as the number of independent constraint equations (Luenberger and Ye, 2015). Therefore, solutions to the balanced transportation problem should have exactly (m + n − 1) basic variables; the remaining (mn − (m + n − 1)) variables are non-basic variables and equal to zero.

Common methods to generate an initial basic feasible solution include 1) the northwest corner rule, 2) Vogel's approximation method, and 3) Russell's approximation

method (Hillier and Lieberman, 2010). Given the problem formulated as a matrix (see Figure 2), the northwest corner method simply starts with the northwest corner of the matrix and assigns x11 and sest its value as the smaller one of S1 and D1. Then, given the last assigned basic variable xij, select xi (j+1) as the next basic variable if source i has any supply left; otherwise, select x (i+1)j. The other two methods try to find a basic feasible solution that is closer to the optimal solution. The Vogel's approximation method tries to reduce the risk of failing to allocate the cell with the lowest unit costs in a row or column. And the Russell's approximation method selects the cells with the cost cij that is much smaller than the maximum cost in the row $\bar{u}i$ and the maximum cost in the column $\bar{v}j$, that is, to minimize $\Delta ij = cij - \bar{u}i - \bar{v}j$. It takes a similar form as the simplex method in the next step for the optimality test.

Figure 3 is an example of the northwest corner rule. The method starts with the cell in the northwest corner from Source 1 to Destination 1, x11. Since the demand from the Destination 1 is 5 units and the supply from Source 1 is 8 units, the cell x11 takes the smaller value of 5 units. Since the first row has 3 units left at Source 1, we move one step to the east and processes the cell x12.The second column has a total demand of 9 units and Source 1 can provide 3 units of supply. Therefore, the cell x12 takes the value of 3 units which is smaller than 9 units. Now, Destination 2 needed 6 more units, so we move to the south and processethe cell x22. In each of the remaining iterations, the method moves one step to the east or south and the cell takes the smaller units remaining in its column and row. The process stops when it reaches the southeast corner of the matrix.

Figure 3. Finding an initial basic feasible solution using northwest corner rule. Source: author.

3.2 Check Optimality

Given the initial basic feasible solution, the next step is to check if the current solution is optimal. The optimization criteria can be formulated by defining two sets of shadow prices for the current solution {xij}: 1) a dispatch cost ui to transport one unit from source i to any destination, and 2) a reception cost vj to transport one unit to destination j from any source. These shadow prices should satisfy:

$$cij + ui = vj, \quad \text{for } xij > 0 \qquad (10)$$

Hence, the current solution is optimal if non-basic variables cannot offer lower costs at the destination (markets):

$$cij + ui \geq vj, \quad \text{for } xij = 0 \qquad (11)$$

These optimization criteria can be implemented by determining values of {ui} and {vi} based on Equation (10), calculating cost differences $\Delta cij = cij + ui - vj$ for all the non-basic variables, and checking whether all Δcij is non-negative. Since the number of basic variables in the basic feasible solution is $(m + n - 1)$ and the total number of shadow prices {ui} and {vj} is $(m + n)$, we can always find a basic solution for {ui} and {vj}, and it is common to set $u1 = 0$. Figure 4 shows an example {ui} and {vj} given the initial basic feasible solution in Figure 3. The price difference Δcij are placed on the left- bottom corner of each cell. Given the calculation result, the initial basic feasible solution is not optimal

because the cost differences Δc 13 , Δc 14 and Δc 32 are negative.

3.3 Improving the Solution Through Iteration

If the current solution is not optimal, there is at least one non-basic variable xij = 0 with cost difference Δcij < 0 that can reduce the overall transportation cost of the system.
A direct and possible method to improve the solution is to select the non-basic variable with the largest (in absolute term) negative Δcij and increase its corresponding units xij as much as possible. For instance, the most negative difference is Δc32 in Figure 4. After increasing x32 by ε unit, the sequence used to generate the initial basic feasible solution would be broken and we need to search for the two donor cells and one recipient cell in addition to cell (3, 2) that can complete a chain reaction. In our case, the donor cells (with basic variables) are (2, 2) and (3, 3) and the receipt cell is (2, 3). And we have x32 + ε, x22 − ε, x23 + ε, and x33 − ε as the complete chain reaction. Hence, the ε takes the smaller value of x22 and x33 , which is 1. The total cost change can be calculated as, which is always negative:

$$\begin{aligned} \Delta Z &= (c32 - c22 + c23 - c33) \times \varepsilon \\ &= (c32 - (v2 - u2) + (v3 - u2) - (v3 - u3)) \times \varepsilon \\ &= (c32 + u3 - v2) \times \varepsilon = \Delta cij \times \varepsilon \end{aligned} \quad (12)$$

Unless this current improved solution is optimal, we go to the next interaction.

Figure 4. Checking possible price reduction – initial transportation simplex tableau. Source: author.

Every linear program has a dual. Given the class transportation problem written as equations (3) to (6), its dual problem can be written using dual variables {ui} and {vj}:

The second and third steps in the transportation simplex method described above use this dual form to determine whether the solution in the current iteration is optimal and, if not, identify a better basic feasible solution.

4. The Classic Transportation Problem and GIS

Geographic information systems for transportation deal with information about and related to transportation systems and have been recognized as a major application of geographic information science (GIS) (Miller and Shaw, 2015). Major extensions to the formulations and solutions of the classic transportation problem are: 1) introducing intermediate transfer stops (a.k.a. hubs) (O'Kelly and Miller, 1994); 2) considering changes in supply and demand amounts and dynamic unit transportation costs (a.k.a. dynamic transportation problem)

(Powell, et al. 1995); 3) including constraints to reflect network and loading capacities (Nahmias and Cheng 1993); and 4) setting multiple objectives beside minimizing shipping costs (Ulungu and Teghem 1994). These extensions correspond to real-world planning and operation scenarios and improve the practical merits of the basic model; however, they also bring great challenges, especially when dealing with large networks.

Regarding the classic transportation problem, GIS provides models and algorithms for collecting, managing, analyzing, and distributing spatial information. These include, but are not limited to, estimating supplies and demands, modeling and maintaining the transportation networks, calculating least-cost routes from sources to destinations for use in determining unit costs, and mapping the data and results. However, GIS has very limited capability to solve the optimization problem with linear programming by itself (Murray et al. 2019). An integrated framework is often adopted to utilize both GIS and non-GIS linear programming solvers to search for the optimal solution (Lei, et al. 2016).

Many existing software packages for optimization have linear programming solvers. Some packages require users to buy licenses such as LGO (Lipschitz-continuous Global Optimizer) and MINOPT (A Modeling Language and Algorithmic Framework for Linear, Mixed-Integer, Nonlinear, Dynamic, and Mixed-Integer Nonlinear Optimization). Some packages provide free versions with limited functions and full-versions upon purchase such as GAMS (General Algebraic Modeling System) and LINDO (Linear, Interactive, and Discrete Optimizer). Some are open-source packages free of charge such as AMPL (A Modeling Language for Mathematical Programming). For complex problems and large networks, the LP solvers typically apply heuristics to reduce computational complexity. And various algorithms have been employed to improve the formulation, solution, and evaluation process. For example, the network simplex algorithm based on graph theory and spanning-tree can more efficiently solve the minimum-cost flow problems than the simplex method described in Section 2. The CPLEX Optimizer by IBM (commercial), the NetworkX Python package, and additional software packages all include network simplex solvers. Some other examples include fuzzy and neuro-fuzzy methods, genetic algorithms, tabu search, and branch-and-bound approaches (Srivastava, 2007).

References

Hillier, F. and Lieberman, G. (2010). Introduction to Operations Research, 9th Edition. McGraw-Hill Education.

Lei, T. L., Church, R. L., & Lei, Z. (2016). A unified approach for location-allocation analysis: integrating GIS, distributed computing and spatial optimization. International Journal of Geographical Information Science, 30(3), 515–534.

Luenberger, D. G., and Ye, Y. (2015). Linear and Nonlinear Programming (Vol. 228). Springer.

Miller, H. J., and Shaw, S. L. (2001). Geographic Information Systems for Transportation: Principles and Applications. Oxford University Press.

Miller, H. J., and Shaw, S. L. (2015). Geographic Information Systems for Transportation in the 21st Century. Geography Compass, 9(4), 180-189.

Murray A. T., Xu, J., Wang, Z., and Church, R. L. (2019). Commercial GIS location analytics: capabilities and performance. International Journal of Geographical Information Science 33(5): 1106-1130.

Nahmias, S., and Cheng, Y. (1993). Production and Operations Analysis (Vol. 2). Homewood, IL: Irwin.

O'Kelly, M. E., and Miller, H. J. (1994). The Hub Network Design Problem: A Review and Synthesis. Journal of Transport Geography, 2(1), 31-40.

Powell, W. B., Jaillet, P., and Odoni, A. (1995). Stochastic and Dynamic Networks and Routing. Handbooks in Operations Research and Management Science, 8, 141-295.

Srivastava, S. K. (2007). Green Supply-Chain Management: A State-Of-The-Art Literature Review. International Journal of Management Reviews, 9(1), 53-80.

Ulungu, E. L., and Teghem, J. (1994). Multi-objective combinatorial optimization problems: A survey. Journal of Multi-Criteria Decision Analysis, 3(2), 83-104.