5. THE TRANSPORTATION PROBLEM AND THE ASSIGNMENT PROBLEM

1. The transportation problem

The transportation problem deals with the transportation of any product from m origins, $O_1, \ldots, O_m$, to n destinations, $D_1, \ldots, D_n$, where:

- The origin O_i has a supply of a_i units, $i = 1, \ldots, m$.
- The destination D_j has a demand for b_j units to be delivered from the origins, $j = 1, \ldots, n$.
- c_{ij} is the cost per unit distributed from the origin O_i to the destination D_j, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

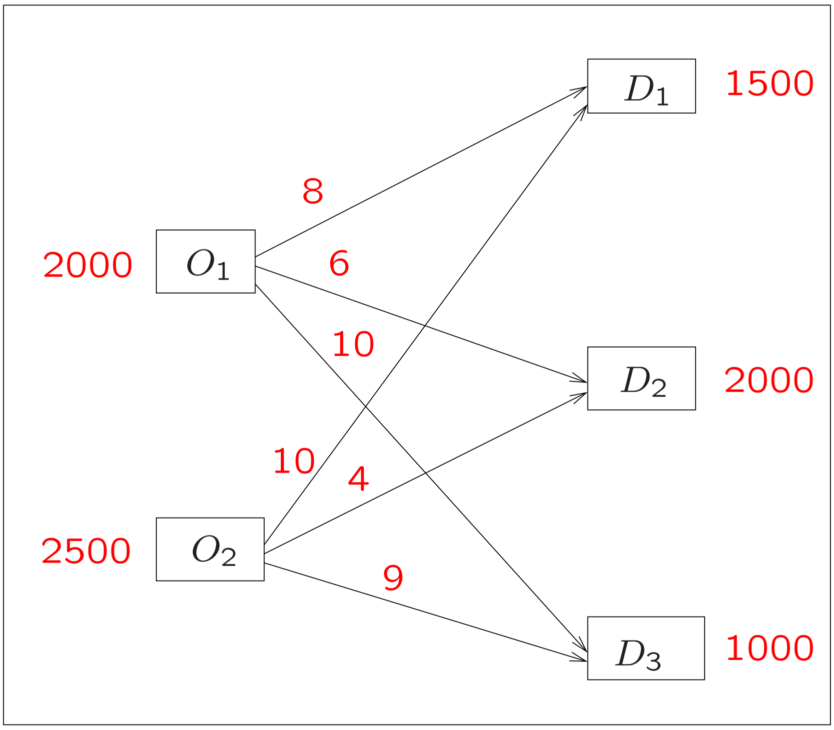
In mathematical terms, the above problem can be expressed as finding a set of x_{ij}'s, $i = 1, \ldots, m$, $j = 1, \ldots, n$, to meet supply and demand requirements.

The aim: to minimize the total distribution cost.

Linear model:

$$\begin{aligned} \text{min} \quad z &= \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \\ &\text{subject to} \\ &\sum_{j=1}^{n} x_{ij} \leq a_i, \quad i = 1, \ldots, m \\ &\sum_{i=1}^{m} x_{ij} \geq b_j, \quad j = 1, \ldots, n \\ x_{ij} \geq 0, \quad i &= 1, \ldots, m, \quad j = 1, \ldots, n \end{aligned}$$

Example.



x_{ij}: the number of loaves of bread to be distributed from the bread factory O_i to the bakery D_j, $i = 1,2$, $j = 1,2,3$.

$$\min \; z = 8x_{11} + 6x_{12} + 10x_{13} + 10x_{21} + 4x_{22} + 9x_{23}$$

subject to

$$\begin{array}{llllllll}
x_{11} & +x_{12} & +x_{13} & & & & = 2000 \\
 & & & x_{21} & +x_{22} & +x_{23} & = 2500 \\
x_{11} & & & +x_{21} & & & = 1500 \\
 & x_{12} & & & +x_{22} & & = 2000 \\
 & & x_{13} & & & +x_{23} & = 1000
\end{array}$$

$$x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23} \geq 0$$

We can write the constraints in equation form, because the total supply is equal to the total demand.

The model in matrix form:

$$\min\ z = (8\ ,\ 6\ ,\ 10\ ,\ 10\ ,\ 4\ ,\ 9)\begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \end{pmatrix}$$

subject to

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \end{pmatrix} = \begin{pmatrix} 2000 \\ 2500 \\ 1500 \\ 2000 \\ 1000 \end{pmatrix}$$

$$x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23} \geq 0$$

2. The matrix format for the transportation problem

The relevant data for any transportation problem can be summarized in a matrix format using a tableau called the transportation costs tableau.

	D_1	D_2	$\cdots$	D_n	Supply
O_1	c_{11}	c_{12}	$\cdots$	c_{1n}	a_1
O_2	c_{21}	c_{22}	$\cdots$	c_{2n}	a_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
O_m	c_{m1}	c_{m2}	$\cdots$	c_{mn}	a_m
Demand	b_1	b_2	$\cdots$	b_n	

Example.

	D_1	D_2	D_3	Supply
O_1	8	6	10	2000
O_2	10	4	9	2500
Demand	1500	2000	1000	

4. Theorems and definitions

Theorem 1 *The necessary and sufficient condition for a transportation problem to have a solution is that the total demand equals the total supply.*

A transportation problem is said to be balanced if $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$. If the transportation problem is unbalanced, we have to convert it into a balanced one before solving it. There are two possible cases:

1. The demand exceeds the supply, $\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j$.

   A dummy origin: O_{m+1}.

   $$a_{m+1} = \sum_{j=1}^{n} b_j - \sum_{i=1}^{m} a_i, \quad c_{m+1,j} = 0, \quad j = 1, \ldots, n$$

2. The supply exceeds the demand. $\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j$.

   A dummy destination: D_{n+1}.

   $$b_{n+1} = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j, \quad c_{i,n+1} = 0, \quad i = 1, \ldots, m$$

Theorem 2 *A balanced transportation problem always has a feasible solution.*

Theorem 3 *A balanced transportation problem always has a basic feasible solution. Such a solution consists of* $m + n - 1$ *positive variables at most.*

5. Finding an initial basic feasible solution

To compute an initial basic feasible solution we will use a tableau of the same dimensions as the transportation costs tableau; the transportation solution tableau.

Each position (i, j) is associated with the decision variable x_{ij}, that is, the number of units of product to be transported from origin O_i to destination D_j.

The transportation solution tableau:

	D_1	D_2	$\cdots$	D_n	Supply
O_1	x_{11}	x_{12}	$\cdots$	x_{1n}	a_1
O_2	x_{21}	x_{22}	$\cdots$	x_{2n}	a_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
O_m	x_{m1}	x_{m2}	$\cdots$	x_{mn}	a_m
Demand	b_1	b_2	$\cdots$	b_n	

The main difference between the Northwest Corner method and Vogel's approximation method lays in the criteria used to select a cell in the solution tableau.

We have to balance the transportation problem before solving it.

5.1 The Northwest Corner method

* **Step 1**. In the rows and columns under consideration, select the cell (i, j) in the northwest corner of the solution tableau.

* **Step 2**. Assign to the variable x_{ij} the maximum feasible amount consistent with the row and the column requirements of that cell, $x_{ij} = \min\{a_i, b_j\}$. Adjust the supply a_i and the demand b_j as follows:

- If a_i happens to be the minimum, then the supply of the origin O_i becomes zero, and the row i is eliminated from further consideration. The demand b_j is replaced by $b_j - a_i$.

- If b_j happens to be the minimum, then the demand of the destination D_j becomes zero, and the column j is eliminated from further consideration. The supply a_i is replaced by $a_i - b_j$.

- If $a_i = b_j$, then the adjusted values for the supply a_i and the demand b_j become both zero. The row i and the column j are eliminated from further consideration.

* **Step 3**. Two cases may arise:

- If only one row or only one column remains under consideration, then any remaining cells (i, j), that is, variables x_{ij} associated with these cells, are selected and the remaining supplies are assigned to them. Stop.

- Otherwise, go to Step 1.

5.2 Vogel's approximation method

- **Step 1**. Compute RD_i and CD_j: The arithmetic difference between the smallest and the next smallest unit cost c_{ij} which remain under consideration in the row i: RD_i. Same definition for the column j: CD_j.

  Find the row or column with the largest difference, and find in it the cell (i, j) with the smallest c_{ij}.

- **Step 2**. Assign to the variable x_{ij} the maximum feasible amount consistent with the row and the column requirements of that cell, $x_{ij} = \min\{a_i, b_j\}$. Adjust the supply a_i and the demand b_j as follows:

  - If a_i happens to be the minimum, then the supply of the origin O_i becomes zero, and the row i is eliminated from further consideration. The demand b_j is replaced by $b_j - a_i$.
  - If b_j happens to be the minimum, then the demand of the destination D_j becomes zero, and the column j is eliminated from further consideration. The supply a_i is replaced by $a_i - b_j$.
  - If $a_i = b_j$, then the adjusted values for the supply a_i and the demand b_j become both zero. The row i and the column j are eliminated from further consideration.

- **Step 3**. Two cases may arise:

  - If only one row or only one column remains under consideration, then any remaining cells (i, j), that is, variables x_{ij} associated with these cells, are selected and the remaining supplies are assigned to them. Stop.
  - Otherwise, go to Step 1.

6. Improvement of a basic feasible solution

The dual transportation problem is used to find an improved basic feasible solution.

Balanced transportation problem:

$$\min \ z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

subject to

$$\sum_{j=1}^{n} x_{ij} = a_i, \qquad i = 1, \ldots, m$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \qquad j = 1, \ldots, n$$

$$x_{ij} \geq 0, \qquad i = 1, \ldots, m, \qquad j = 1, \ldots, n$$

We denote by $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ the dual variables,

The dual model:

$$\max \ G = \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} b_j v_j$$

subject to

$$u_i + v_j \leq c_{ij}, \quad i = 1, \ldots, m, \qquad j = 1, \ldots, n$$

$$u_i, v_j : \text{unrestricted}, \qquad i = 1, \ldots, m, \qquad j = 1, \ldots, n$$

Example. A balanced transportation problem:

$$\min \; z = 8x_{11} + 6x_{12} + 10x_{13} + 10x_{21} + 4x_{22} + 9x_{23}$$

subject to

$$\begin{array}{cccccccc}
x_{11} & +x_{12} & +x_{13} & & & & = 2000 \\
 & & & x_{21} & +x_{22} & +x_{23} & = 2500 \\
x_{11} & & & +x_{21} & & & = 1500 \\
 & x_{12} & & & +x_{22} & & = 2000 \\
 & & x_{13} & & & +x_{23} & = 1000
\end{array}$$

$$x_{11}, x_{12}, x_{13}, x_{21}, x_{21}, x_{23} \geq 0$$

The dual variables: u_1, u_2, v_1, v_2, v_3. The dual model:

$$\max \; G = 2000u_1 + 2500u_2 + 1500v_1 + 2000v_2 + 1000v_3$$

subject to

$$\begin{array}{cccccc}
u_1 & & +v_1 & & & \leq 8 \\
u_1 & & & +v_2 & & \leq 6 \\
u_1 & & & & +v_3 & \leq 10 \\
 & u_2 & +v_1 & & & \leq 10 \\
 & u_2 & & +v_2 & & \leq 4 \\
 & u_2 & & & +v_3 & \leq 9
\end{array}$$

$$u_i, v_j : \text{unrestricted}$$

6.1 Selection of the entering vector

We need to compute the reduced cost coefficients to decide whether the solution in hand can be improved,

$$z_{ij} - c_{ij} = \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{a}_{ij} - c_{ij}.$$

Since $\mathbf{c}_B^T \mathbf{B}^{-1}$ is the vector of dual variables,

$$\mathbf{c}_B^T \mathbf{B}^{-1} = (u_1, \ldots, u_m, v_1, \ldots, v_n),$$

then,

$$z_{ij} - c_{ij} = (u_1, \ldots, u_m, v_1, \ldots, v_n)\mathbf{a}_{ij} - c_{ij}.$$

Since $\mathbf{a}_{ij}$ consists of 1's and 0's, and the only two 1's are in the ith and in the $(m+j)$th positions, we get:

$$z_{ij} - c_{ij} = u_i + v_j - c_{ij}.$$

We compute the dual variables: since $z_{ij} - c_{ij} = 0$ for any basic x_{ij}, we write the system of linear equations, and solve it by setting any dual variable at an arbitrary value.

Taking into account that the objective is to minimize, two cases may arise:

- If $z_{ij} - c_{ij} \leq 0$ for any $i = 1, \ldots, m,\ \ j = 1, \ldots, n$, then the solution is optimal.

- If there is one or more $z_{ij} - c_{ij} > 0$, then the solution may be improved, and the variable with the maximum reduced cost coefficient $z_{ij} - c_{ij}$ is selected to enter the basis.

6.2 Selection of the leaving vector

We need to take into account the following:

1. The $m+n-1$ basic cells corresponding to any basic solution never contain a cycle.
2. The $m+n-1$ basic variables together with the entering variable contain a unique cycle.

To find such a unique cycle: The cell that corresponds to the entering variable is assumed to be basic and is marked with the symbol $\uparrow$. We cross out all the rows and columns which contain only one basic cell: we first cross out the rows, for instance, and afterwards the columns, then we check the rows again and so on. At the end of this process, the basic cells that have not been crossed out as well as the cell that corresponds to the entering variable contain the unique cycle.

To adjust the entries in the cycle: since the entering variable will be assigned a positive value, the basic cells in the cycle that are in the same row or in the same column are marked by the symbol $\downarrow$, the basic cells in the cycle that are in the same row or in the same column as the ones previously marked with the symbol $\downarrow$ are marked with the symbol $\uparrow$ and so on. Stop when all the cells in the cycle are marked.

The first basic variable in the cycle decreased to zero among those marked by the symbol $\downarrow$ becomes the leaving variable. All the basic variables in the cycle are recomputed. The basic variables that do not belong to the cycle remain unchanged. The entering variable is assigned the value the leaving variable had before being adjusted.

7. The transportation tableau

Up to now, we have carried out all the calculations in two different tableaux: the transportation costs tableau and the transportation solution tableau.

To improve a solution, we have computed the values of the dual variables u_i, v_j and the reduced cost coefficients $z_{ij} - c_{ij}$ outside the two tableaux.

The transportation tableau puts together all the calculations needed to solve a transportation problem.

	v_1	v_2	$\cdots$	v_n	
u_1	$z_{11}-c_{11}$ c_{11} x_{11}	$z_{12}-c_{12}$ c_{12} x_{12}	$\cdots$	$z_{1n}-c_{1n}$ c_{1n} x_{1n}	a_1
u_2	$z_{21}-c_{21}$ c_{21} x_{21}	$z_{22}-c_{22}$ c_{22} x_{22}	$\cdots$	$z_{2n}-c_{2n}$ c_{2n} x_{2n}	a_2
$\vdots$			$\ddots$		$\vdots$
u_m	$z_{m1}-c_{m1}$ c_{m1} x_{m1}	$z_{m2}-c_{m2}$ c_{m2} x_{m2}	$\cdots$	$z_{mn}-c_{mn}$ c_{mn} x_{mn}	a_m
	b_1	b_2	$\cdots$	b_n	

8. The transportation algorithm

The objective is to minimize.

- **Step 1**. Balance the transportation problem.

- **Step 2**. Compute an initial basic feasible solution.

- **Step 3**. Compute the values $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ associated with the current basis.

- **Step 4**. Compute the reduced cost coefficient $z_{ij} - c_{ij} = u_i + v_j - c_{ij}$ for every nonbasic variable x_{ij}.

  - If there is one or more $z_{ij} - c_{ij} > 0$, then the current solution may be improved. Select the variable with the most positive $z_{ij} - c_{ij}$ to enter the basis, and go to Step 5.

  - If $z_{ij} - c_{ij} < 0$ for all nonbasic variables, then the current basic feasible solution is optimal and unique. Stop.

  - If $z_{ij} - c_{ij} \leq 0$ for all nonbasic variables, and for one or more of the nonbasic variables $z_{ij} - c_{ij} = 0$ holds, then the optimal solution is multiple. Select anyone of those nonbasic variables to enter the basis, and go to Step 5.

- **Step 5**. Find the cycle which includes some of the current basic variables and the entering variable. Recompute the values of the variables in the cycle to obtain a new basic feasible solution. Go to Step 3.

9. Degeneracy

If a solution to a given balanced transportation problem with m origins and n destinations has less than $m+n-1$ positive variables, that is, if at least one of the basic variables has the value zero, then it is said to be degenerate.

Degeneracy may occur in the following two cases:

- While computing an initial basic feasible solution, either applying Vogel's approximation method or using the northwest corner method. If at Step 2 of either algorithm we find that when assigning to a variable the maximum feasible amount consistent with the row and the column requirements, both the supply and the demand are equal, then they become both zero; a row and a column are eliminated at the same time.

- While applying the transportation algorithm at Step 5, once we have identified the entering variable and found the cycle. If two decreasing variables in the cycle (two cells marked by the symbol $\downarrow$) tie for the minimum, then after recomputing the values in the cycle, they both become zero. However, only one of them can be selected as the leaving variable.

If degeneracy is encountered, we have to distinguish the zero-valued basic variable from the nonbasic ones.

10. The assignment problem

It is a special case of the transportation problem.

It deals with assigning n origins O_i to n destinations D_j with the goal of determining the minimum cost assignment. Each origin must be assigned to one and only one destination, and each destination must have assigned one and only one origin. c_{ij} represents the cost of assigning the origin O_i to the destination D_j.

Decision variables are defined like this:

$$x_{ij} = \begin{cases} 1 & \text{if } O_i \text{ is assigned to } D_j \\ 0 & \text{otherwise} \end{cases}$$

A linear model in standard form for the assignment problem:

$$\begin{aligned} \text{min } z &= \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \\ \text{subject to} & \\ \sum_{j=1}^{n} x_{ij} &= 1, \qquad i = 1, \ldots, n \\ \sum_{i=1}^{n} x_{ij} &= 1, \qquad j = 1, \ldots, n \\ x_{ij} &= 0, 1, \quad i, j = 1, \ldots, n \end{aligned}$$

If the assignment problem has the same number of origins and destinations, it is balanced.

If it is not balanced, we can always add as many dummy origins or dummy destinations as necessary.

The relevant data for any assignment problem can be summarized in a matrix format using a tableau called the assignment costs tableau.

$$\begin{array}{c|cccc|}
 & D_1 & D_2 & \dots & D_n \\
\hline
O_1 & c_{11} & c_{12} & \dots & c_{1n} \\
O_1 & c_{21} & c_{22} & \dots & c_{2n} \\
\vdots & \vdots & \ddots & & \vdots \\
O_n & c_{n1} & c_{n2} & \dots & c_{nn} \\
\hline
\end{array}$$

10.1 The Hungarian method

Theorem 4 *Consider the objective function of an assignment problem:*

$$z = \sum_{j=1}^{n} \sum_{i=1}^{n} c_{ij} x_{ij}.$$

If a solution x_{ij} is optimal for the objective function z, then it is also optimal for the problem in which the objective function has been replaced by z':

$$z' = \sum_{j=1}^{n} \sum_{i=1}^{n} c'_{ij} x_{ij},$$

where $c'_{ij} = c_{ij} - u_i - v_j$ and u_i and v_j are constant, $i, j = 1, \dots, n$.

Theorem 5 *Given* $c_{ij} \geq 0$ *and a solution* x_{ij} *to the assignment problem, if* $z = \sum_{j=1}^{n} \sum_{i=1}^{n} c_{ij} x_{ij} = 0$ *holds, then the values of* x_{ij} *are optimal,* $i, j = 1, \ldots, n$.

If a constant is subtracted from each cost in a row or column, the optimal solution to the problem remains unchanged.

The solution of an assignment problem proceeds by transforming the assignment costs tableau to create zero assignment-costs. Then, if a feasible assignment is found in which all the x_{ij}'s that equal 1 have zero costs and thus $z = 0$, the assignment is an optimal solution to the problem.

10.2 The Hungarian algorithm

The objective is to minimize.

- **Step 1.** Balance the assignment problem.
- **Step 2.** Create zero entries in the rows. For each row in the assignment costs tableau, subtract the row minimum u_i from each element in the row, $u_i = \min_j \{c_{ij}\}$. The new entries in the resulting tableau are $c'_{ij} = c_{ij} - u_i, \quad i, j = 1, \ldots, n.$
- **Step 3.** Create zero entries in the columns. For each column in the resulting tableau, subtract the column minimum v_j from each element in the column, $v_j = \min_i \{c'_{ij}\}$. The new entries are $c''_{ij} = c'_{ij} - v_j, \quad i, j = 1, \ldots, n.$
- **Step 4.** Choose independent zeros. Find the row or column with the smallest number of zero entries. Choose one of its zeros, and cross out all the zeros in the same row or column. Proceed to choose more zeros among the ones that have not been crossed out, starting at the row or column with the smallest number of them, until all zeros are either chosen or crossed out.
  - If n independent zeros have been chosen, an optimal solution is available. Stop.
  - If less than n independent zeros have been chosen, then go to Step 5.

- **Step 5.** Draw the minimum number of lines (horizontal or vertical) that are needed to cover all zeros in the tableau, proceeding like this:

  5.1 Mark with a cross any row where none of its zeros has been chosen.

  5.2 Mark with a cross the columns corresponding to the crossed out zeros in the rows marked in Step 5.1.

  5.3 Mark with a cross the rows corresponding to the chosen zeros of the columns marked in Step 5.2.

  Repeat Steps 5.2 and 5.3, until there are no rows or columns that meet the condition.

  Draw a line through any row not marked and any column marked.

  These lines cover all the zeros in the tableau. Go to Step 6.

- **Step 6.** Create new zeros. Find the smallest nonzero entry (call its value k) that is uncovered by the lines drawn in the previous step. Subtract k from each cost that lies in an uncovered row and add k to each cost that lies in a covered column. Go to Step 4.

10.3 Maximization problems

The Hungarian method can only be applied to assignment problems where the objective is to minimize. In the cases where the objective is given in maximization form, it has to be transformed:

$$\min(-z) = \sum_{i=1}^{n}\sum_{j=1}^{n} -c_{ij}x_{ij}.$$

This transformation of the objective function causes the assignment costs to be negative. The theorem can only be used if costs are nonnegative.

One way to make all entries in the assignment tableau nonnegative is to choose the minimum among the negative entries ($-c_{kl} = \min\{\ -c_{ij}\ /\ -c_{ij} < 0\ \}$), and to subtract it from each cost in the tableau,

$$c'_{ij} = -c_{ij} + c_{kl}.$$

The new entries in the tableau are, hence, nonnegative, $c'_{ij} \geq 0$.