THE DISTRIBUTION OF A PRODUCT FROM SEVERAL SOURCES TO NUMEROUS LOCALITIES

By Frank L. Hitchcock

1. Statement of the problem. When several factories supply a product to a number of cities we desire the least costly manner of distribution. Due to freight rates and other matters the cost of a ton of product to a particular city will vary according to which factory supplies it, and will also vary from city to city.

Denote the cost of one ton of product from the i^{th} factory to the j^{th} city by a_{ij} and the number of tons shipped in each case by x_{ij}. The total cost y will then be

$$y = \sum a_{ij} x_{ij} \tag{1}$$

summed for $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n$ if there are m factories and n cities.

Let the total product shipped from the i^{th} factory to the n cities be f_i; this means we have the equations

$$\begin{aligned} x_{11} + x_{12} + \cdots + x_{1n} &= f_1 \\ x_{21} + x_{22} + \cdots + x_{2n} &= f_2 \\ \cdots\cdots\cdots&\cdots\cdots \\ x_{m1} + x_{m2} + \cdots + x_{mn} &= f_m \end{aligned} \tag{F}$$

which may be called the F-equations for convenience.

Let also the total product consumed by the j^{th} city be c_j tons; this means we have the equations

$$\begin{aligned} x_{11} + x_{21} + \cdots + x_{m1} &= c_1 \\ x_{12} + x_{22} + \cdots + x_{m2} &= c_2 \\ \cdots\cdots\cdots&\cdots\cdots \\ x_{1n} + x_{2n} + \cdots + x_{mn} &= c_n \end{aligned} \tag{C}$$

which may be called the C-equations.

The F-equations and the C-equations taken together form a system of $m + n$ equations which must be satisfied by the mn unknowns x_{ij}. But since the total shipped by factories equals total consumed by cities the sum of the F-equations equals the sum of the C-equations whence the equations are not independent. It can easily be shown that any *one* equation may be omitted and the rest solved for a properly chosen set of $m + n - 1$ of the x's in terms of the other x's as parameters. If the results are substituted in (1) we shall have an expression for y in terms of $mn - (m + n - 1)$ parameters.

From the nature of the problem none of the unknowns x_{ij} can be negative, and no x_{ij} can exceed the smaller of the two numbers f_i and c_j.

2. Geometrical interpretation. A geometrical picture will be helpful. If the mn unknowns x_{ij} are rectangular coordinates in mn dimensions, then since we have $m + n - 1$ independent equations relating these variables, the region corresponding to possible values of the x's will be of $mn - (m + n - 1)$ dimensions in agreement with the number of parameters found above. Since the possible values of the x's are bounded, on the one hand by zero, on the other by f_i or c_j whichever is smaller, the region of possibilities is a closed region in the space of mn dimensions. This region in which meaningful values of the x's occur will be called for brevity the X-region.

The cost y is, by (1), a linear function of position in the X-region. The least value of y will occur at some point on the *boundary* of the X-region, since a linear function cannot attain its least value in free space.

At any point on the boundary of the X-region at least one of the x's must be zero. For at least one of the x's must have one of its extreme values, otherwise the point would not be on the boundary. And if one of the x's has its maximum value one or more other x's must be zero, by the form of the F-equations and C-equations.

Since the X-region is of dimensions $mn - m - n + 1$, a portion of its boundary over which one of the x's is zero will be of dimensions $mn - m - n$. Taking a point on this part of the boundary of the X-region we may move in a direction of diminishing y until we come to a point where some *other* x reaches a limiting value, and there at least one other x must be zero. The region where two x's are zero will be of dimensions $mn - m - n - 1$. Proceeding in this way we shall eventually reach a point where $mn - m - n + 1$ of the x's are zero, i.e. a fixed point on the boundary of the X-region which may be called a *vertex*.

The value of y at a vertex is not necessarily its least value, but the least value must occur at *some* vertex, which may be called conveniently a *best vertex*. If the particular set of $mn - m - n + 1$ x's which are zero at a vertex are taken as parameters and the other x's eliminated from (1), the expression for y will take the form

$$y = M + \sum kx \tag{2}$$

where M is a constant and $\sum kx$ denotes a linear combination of the parametric x's. *If none of the coefficients k in this result is negative the vertex is a "best vertex" and M is the required minimum cost.* For $y = M$ at the vertex, and if any parametric x is different from zero the cost y is increased.

The method will be to find a vertex and form the expression for y in terms of the x's which are zero at this vertex. It can be seen by inspection whether we have a "best vertex". If not, we can always pass easily to a better vertex, i.e. one giving a smaller value of y, and continue in this way until we have a best vertex.

3. Finding a vertex. A vertex can always be quickly found in the following manner. Let the F-equations and the C-equations, taken together, be condensed in form according to the scheme

$$\begin{array}{ccccc} x_{11} & x_{12} & \cdots & x_{1n} & f_1 \\ x_{21} & x_{22} & \cdots & x_{2n} & f_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} & f_m \\ c_1 & c_2 & \cdots & c_n & \end{array} \tag{FC}$$

wherein the sum of the x's in a row equals the f on the right and the sum of the x's in a column equals the c at the bottom. We may suppose the factories and the cities to be so numbered that no f is less than any f above it, and no c is less than any c on its left. We now set all the x's of the first row equal to zero except x_{1n} which then equals f_1; *or else* we take all x's of the first column zero except x_{m1} which then equals c_1. *One* of these two procedures is always possible. Suppose the first case, i.e. $x_{1n} = f_1$. We now discard the first row, replacing c_n at the bottom by $c_n - f_1$ assumed to be positive. If this quantity is negative the first procedure is impossible and instead we take $x_{m1} = c_1$, drop the first column, and replace f_m by $f_m - c_1$ which is surely not negative if $c_n - f_1$ is negative as is easily seen.

We have thus reduced the original array to a smaller array, corresponding to one less factory or else one less city, as the case may be. We then continue in just the same way until we have made $mn - m - n + 1$ of the x's zero and so have a vertex.

To illustrate, take an example of four cities and three factories with the array

$$\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & x_{14} & 25 \\ x_{21} & x_{22} & x_{23} & x_{24} & 25 \\ x_{31} & x_{32} & x_{33} & x_{34} & 50 \\ 15 & 20 & 30 & 35. & \end{array} \tag{3}$$

Either of the procedures mentioned above is possible. Adopting the first we set $x_{11} = x_{12} = x_{13} = 0$, $x_{14} = 25$, drop the first row and replace 35 at the bottom by $35 - 25$ or 10. This gives an array like that for two factories and four cities, namely

$$\begin{array}{ccccc} x_{21} & x_{22} & x_{23} & x_{24} & 25 \\ x_{31} & x_{32} & x_{33} & x_{34} & 50 \\ 15 & 20 & 30 & 10. & \end{array}$$

Since the largest number at the bottom is 30, the third column will now play the part of a right hand column but it is hardly necessary to rewrite the array. We merely set $x_{21} = x_{22} = x_{24} = 0$, $x_{23} = 25$. This entails $x_{31} = 15$, $x_{32} = 20$, $x_{33} = 5$, and $x_{34} = 10$. The whole process may be carried out on the original array (3),

$$\begin{array}{ccccc} (x_{11})0 & (x_{12})0 & (x_{13})0 & (x_{14})25 & 25 \\ \hline [x_{21}]0 & [x_{22}]0 & [x_{23}]25 & [x_{24}]0 & 25 \\ x_{31} & x_{32} & x_{33} & x_{34} & 50 \\ 15 & 20 & [30]5 & (35)10 & \end{array}$$

where, in the first row, the values to the right of each x are those to be used, the parentheses referring to the first step of the process. In the second row also, the values to the right of each x are to be used, and the brackets refer to the second stage of the work. Finally the values remaining at the bottom are those of the x's of the last row. The whole clearly checks, as is seen by adding rows and columns. The method is perfectly general and may be used in any case to find a vertex.

We now choose the x's which were set equal to zero to be parameters, express the other x's in terms of these, and substitute in the cost expression (1). This is a simple calculation and we find

$$x_{14} = 25 - x_{11} - x_{12} - x_{13} \tag{4}$$

$$x_{23} = 25 - x_{21} - x_{22} - x_{24} \tag{5}$$

$$x_{31} = 15 - x_{11} - x_{21} \tag{6}$$

$$x_{32} = 20 - x_{12} - x_{22} \tag{7}$$

$$x_{33} = 5 - x_{13} + x_{21} + x_{22} + x_{24} \tag{8}$$

$$x_{34} = 10 - x_{24} + x_{11} + x_{12} + x_{13} \tag{9}$$

equations which show by their form that a vertex is obtained when the x's on the right are zero, since the others are left with possible values. The configuration of the X-region and the number of its vertices depend on the constants f and c, not on the costs a_{ij}.

Suppose now the respective costs to be

$$\begin{array}{llll} a_{11} = 10, & a_{12} = 5, & a_{13} = 6, & a_{14} = 7, \\ a_{21} = 8, & a_{22} = 2, & a_{23} = 7, & a_{24} = 6, \\ a_{31} = 9, & a_{32} = 3, & a_{33} = 4, & a_{34} = 8. \end{array} \tag{10}$$

Substituting from (4)–(10) into the cost expression (1) we find

$$y = 645 + 2x_{11} + 3x_{12} + 3x_{13} - 4x_{21} - 4x_{22} - 5x_{24} \tag{11}$$

showing a cost 645 at the vertex where the six parametric x's are zero, not a best vertex since the x's with negative coefficients might be given positive values.

4. Finding a better vertex. We may now let x_{21} increase from zero to its maximum value 15, keeping the other five parameters zero. Inspection of (6) shows then $x_{31} = 0$, i.e. we pass along an edge of the X-region to a new vertex where x_{31} is zero instead of x_{21}. Eliminating x_{21} by aid of (6) we obtain the new system

$$x_{14} = 25 - x_{11} - x_{12} - x_{13} \tag{12}$$

$$x_{23} = 10 - x_{22} - x_{24} + x_{11} + x_{31} \tag{13}$$

$$x_{21} = 15 - x_{11} - x_{31} \tag{14}$$

$$x_{32} = 20 - x_{12} - x_{22} \tag{15}$$

$$x_{33} = 20 - x_{11} - x_{13} - x_{31} + x_{22} + x_{24} \tag{16}$$

$$x_{34} = 10 - x_{24} + x_{11} + x_{12} + x_{13} \tag{17}$$

and $$y = 585 + 6x_{11} + 3x_{12} + 3x_{13} + 4x_{31} - 4x_{22} - 5x_{24} \tag{18}$$

showing that at the new vertex the cost is 585 or 60 less than at the first vertex.

Continuing in the same way we next eliminate x_{22} and introduce x_{23} as parameter, using (13). At the same time, for brevity, we can eliminate x_{24} and introduce x_{34}, using (17). We obtain

$$x_{14} = 25 - x_{11} - x_{12} - x_{13} \tag{19}$$

$$x_{22} = x_{31} + x_{34} - x_{12} - x_{13} - x_{23} \tag{20}$$

$$x_{21} = 15 - x_{11} - x_{31} \tag{21}$$

$$x_{32} = 20 - x_{31} - x_{34} + x_{13} + x_{23} \tag{22}$$

$$x_{33} = 30 - x_{23} - x_{13} \tag{23}$$

$$x_{24} = 10 - x_{34} + x_{11} + x_{12} + x_{13} \tag{24}$$

and

$$y = 535 + x_{11} + 2x_{12} + 2x_{13} + 4x_{23} + x_{34}\,. \tag{25}$$

The absence of negative signs in the last equation shows that we have a "best vertex" by setting the parametric x's equal to zero, and the required minimum cost is 535. The problem is therefore solved.

The method is quite general and can always be used in such problems. This example, however, presents some interesting features. First we note the absence of the parameter x_{31} from the cost expression (25). It follows that there is some latitude in the manner of distribution. For, keeping the other five parameters zero, we can vary x_{31} from zero to its maximum 15, with consequent changes in x_{22}, x_{21}, and x_{32}, seen from (20), (21), and (22), without altering the cost from 535. The cost is constant all along an edge of the X-region.

Moreover setting $x_{31} = 15$, still keeping the other five parameters zero, will entail $x_{21} = 0$, by (21). We have arrived at an extremity of the edge along which y is constant, and have another "best vertex" with, of course, the same cost. That is to say, a "best vertex" need not be unique.

Another interesting feature of this example is that, setting the six parametric x's of (19)–(25) equal to zero we *also* have $x_{22} = 0$, that is we have *seven* of the x's zero at this best vertex. Since *any six* of these could be used as parameters, we have various ways for expressing y at this vertex. Suppose we decide to use x_{22} as parameter instead of x_{34}, which we can eliminate by aid of (20). We get

$$y = 535 + x_{11} + 3x_{12} + 3x_{13} + 5x_{23} + x_{22} - x_{31}\,. \tag{26}$$

We might at first think that, by giving to x_{31} a positive value, we could obtain a cost less than 535, contradicting previous results. But if the other five parameters in (26) are kept zero we cannot increase x_{31} because by (20), this would entail a negative value for x_{34}. In other words the *absence* of negative signs in the expression for y at a vertex is sure sign of a "best vertex" but the *presence* of a negative is not sure proof of the contrary. The test, as just illustrated, is whether we are able to *increase* a parameter having a negative coefficient.