

Centroid – A Widely Misunderstood Concept In Facility Location Problems

William V. Gehrlein¹ and Mugdim Pasic²

¹Alfred Lerner College of Business and Economics
University of Delaware
Alfred Lerner Hall
Newark, Delaware 19716

²Industrial Engineering and Management Department
Mechanical Engineering Faculty, University of Sarajevo,
Vilsonovo setaliste 9
Sarajevo, Bosnia & Herzegovina 71000

Corresponding author's e-mail: {Mugdim Pasic, mugdim.pasic@mef.unsa.ba}

The aim of this paper is to show, by use of a complete and exact mathematical model, that the centroid method is a widely misunderstood method in facility location problems and that it is, in fact, normally an inappropriate method to use for such problems. While numerous sources do describe the procedure as minimizing the total shipping cost when transportation costs are linearly proportional to the distances of travel, this study shows that these statements are not valid. The misunderstanding regarding what the centroid method actually does results from an improper interpretation of the notion of the center of gravity. In fact, the centroid method minimizes shipping costs only if transportation costs are proportional to the squares of distances traveled.

Significance: The research findings have a significant impact on the theory and application of the centroid method in finding optimal facility locations. As a result, the centroid method would seem to be completely inappropriate for most, if not all, facility location problems.

Keywords: Facility location, center of gravity, centroid, shipping cost, transportation

(Received: 1 June 2008; Accepted in revised form: 25 February 2009)

1. INTRODUCTION

The problem of identifying the best location in which to place a new facility must be based on many different criteria. Issues such as political stability, the existing infrastructure in regions and the availability of a trained workforce are critical on a 'macro level' in making such decisions. Once a set of feasible regions have been identified for possible location of a new facility, the determination of the ultimate location then takes place on a 'micro level'. One major 'micro level' criterion is to locate a new production facility to minimize total shipping costs to a set of distribution centers that it will supply. The vast majority of reference books in operations management borrow an idea from the physical sciences to address this topic by suggesting the use of the *Centroid Method*, or the *Center of Gravity Method*, to minimize the total cost of shipping. Numerous sources describe the procedure as minimizing the total cost of shipping when transportation costs are linearly proportional to the distances of travel.

Following universal textbook definitions, we describe this procedure by identifying the locations of each of n distribution centers on a map according to their respective Cartesian coordinates (x_i, y_i) , and we let m_i denote the demand requirement at the corresponding distribution center i . The centroid of the distribution system is then located at map coordinates (x_c, y_c) with:

$$x_c = \frac{\sum_{i=1}^n m_i x_i}{\sum_{i=1}^n m_i}, \quad y_c = \frac{\sum_{i=1}^n m_i y_i}{\sum_{i=1}^n m_i} \quad ..(1)$$

2. LITERATURE SURVEY

Quotes from a sample of reference books suggest that a production facility located in this manner will minimize shipping costs to distribution centers. Shim and Siegel (1999), state that: “The center of gravity method is a method used to determine the location of a distribution center that will minimize transportation costs. The method treats shipping cost as a linear function of the distance and quantity shipped. We assume, however, that the quantity to be shipped to each destination is fixed”, (pages 160-161).

Chase, Jacobs and Aquilano (2006) suggest that: “The centroid method is a technique for locating single facilities that considers the existing facilities, the distances between them, and the volumes of goods to be shipped. The technique is often used to locate intermediate or distribution warehouses. In its simplest form, this method assumes that inbound and outbound transportation costs are equal, and it does not include special shipping costs for less than full loads”, (pages 413-415).

Stevenson (2005) says that: “The center of gravity method is a method to determine the location of a distribution center that will minimize distribution costs. It treats distribution costs as a linear function of the distance and the quantity shipped. The quantity to be shipped to each destination is assumed to be fixed (i.e., will not change over time)”, (pages 366-369). Krajewski and Ritzman (2005) is an exception in noting that the centroid location usually is not optimal in minimizing shipping costs. They conclude that: “Testing different locations with the load-distance model (*to minimize total cost*) is relatively simple if some systematic search process is followed. A good starting point is the center of gravity of the target area. ... This location usually is not optimal one for the Euclidian or rectilinear distance measures, but it still is an excellent starting point. Using center of gravity as a starting point, managers can now search in its vicinity for the optimal solution”, (pages 455-456). However the authors stop short of suggesting what the centroid method actually does as an optimization method for minimizing total shipping cost. Heizer and Render (2005) write that: “The center of gravity method is a mathematical technique used for finding the location of a distribution center that will minimize distribution costs. The method takes into account the location of markets, the volume of goods shipped to those markets, and shipping costs in finding the best location for a distribution center”, (pages 251-253). Some theoretical work has been done to suggest what the use of the centroid method actually does purely in terms of an optimization perspective (ReVelle and Eiselt, 2005), but the implications of this work are easily misunderstood by academics who teach in the area of operations management.

Russell and Taylor (2006) state that: “In general, transportation costs are a function of distance, weight and time. The center of gravity, or *weight center*, technique is a quantitative method for locating a facility such as a warehouse at the center of movement in a geographic area based on weight and distance”, (pages 301-303).

The aim of this article is to further contribute to the operations management applied theory of the centroid method and its application in optimal facility location, as well as to show that the centroid method is typically an inappropriate method in facility locations problem. In order to truly understand what the centroid method of facility location actually accomplishes, we must first understand what the center of gravity means in the physical sciences from which the notion was borrowed. This will lead us to the conclusion that the centroid method actually is a very poor method for facility location in most applications, despite the claims in the references given above.

3. AN EXAMPLE OF THE MISUSE OF THE CENTROID

We use an example that is based on a situation that is described in Chase, Jacobs and Aquilano (2006, pgs. 457-458) where a refining company needs to locate an intermediate holding facility between its refining plant at location L1 and its major distributors at locations L2, L3, L4 and L5. Table 1 gives the coordinates of these locations (in miles relative to a given origin position), along with the amount of gasoline that is shipped to or from the refining plant and the distributors. A coordinate map of all of these locations is shown in Figure 1.

Table 1. Coordinates of the locations and the volume shipped.

Coordinates	Location				
	L1	L2	L3	L4	L5
x_i	325	400	450	350	25
y_i	75	150	350	400	450
Millions of Gallons (m_i)	1,500	250	450	350	450

The coordinates of the centroid (x_c, y_c) are calculated by using (1), which results in:

$$x_c = \frac{\sum_{i=1}^4 x_i m_i}{\sum_{i=1}^4 m_i} = \frac{(325 \times 1,500) + (400 \times 250) + (450 \times 450) + (350 \times 350) + (25 \times 450)}{1,500 + 250 + 450 + 350 + 450} = \frac{923,750}{3,000} = 307.9 \quad \dots (2)$$

$$y_c = \frac{\sum_{i=1}^4 y_i m_i}{\sum_{i=1}^4 m_i} = \frac{(75 \times 1,500) + (150 \times 250) + (350 \times 450) + (400 \times 350) + (450 \times 450)}{1,500 + 250 + 450 + 350 + 450} = \frac{650,000}{3,000} = 216.7 \quad \dots (3)$$

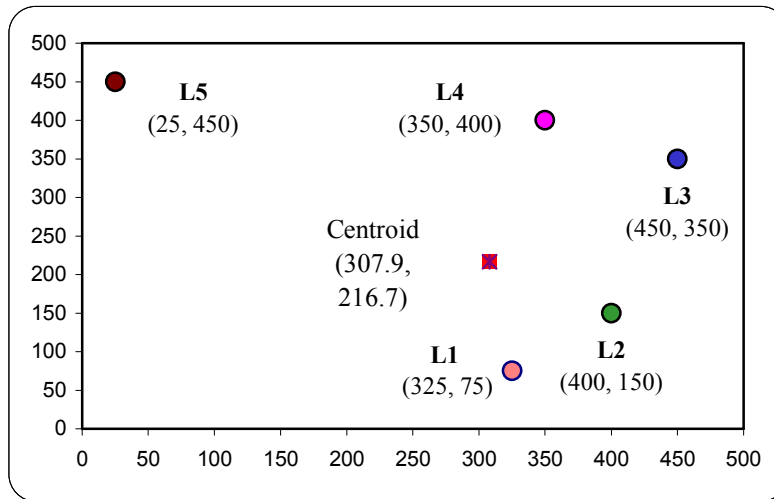


Figure 1. Coordinate map for the facility locations and the centroid.

The total shipping cost (TSC) that results from positioning a distribution facility at a location L^* that is located at coordinates (x^*, y^*) is generally expressed by the following equation:

$$TSC = uc \sum_{i=1}^n d_i m_i \quad \dots (4)$$

where:

uc = cost to ship a unit volume per unit distance

d_i = shipping distance between L^* and the location with coordinates (x_i, y_i) .

In our example, uc would be the cost to ship one million gallons over a distance of one mile. The shipping distance d_i is obtained as the linear distance between L^* and the location with coordinates (x_i, y_i) , with:

$$d_i = \sum_{i=1}^n \sqrt{(x_i - x^*)^2 + (y_i - y^*)^2} \quad \dots (5)$$

Substituting (5) into (4) yields to:

$$TSC = uc \sum_{i=1}^n \sqrt{(x_i - x^*)^2 + (y_i - y^*)^2} m_i \quad . (6)$$

The value of TSC that is obtained from the centroid coordinates is denoted by TSC_c , and by using the centroid coordinates from (2) and (3) for L^* in (6), we find $TSC_c = 561,003 \times uc$. This would be the minimum value of TSC that could be obtained, as claimed in most textbooks, if indeed the centroid location truly minimizes TSC when shipping costs are a linear function of the distance and quantity shipped.

Let us take another perspective of this problem by making a direct attempt to minimize TSC , as it is defined in (6). Since TSC is the function of two unknowns, x^* and y^* , partial derivatives of TSC should be taken with respect to x^* and y^* and each should be set equal to zero, as shown in (7) and (8). The simultaneous solution of (7) and (8) will then yield the optimal location, (x_{op}, y_{op}) , that minimizes TSC .

$$\frac{\partial TSC}{\partial x^*} = 0 \quad \dots (7)$$

$$\frac{\partial TSC}{\partial y^*} = 0 \quad \dots (8)$$

By taking the partial derivatives of TSC in (6) with respect to x^* and y^* and using (7) and (8), the optimal values of x_{op} and y_{op} are obtained by the following two equations:

$$uc \times \sum_{i=1}^n \frac{(x_i - x_{op})}{\sqrt{(x_i - x_{op})^2 + (y_i - y_{op})^2}} m_i = 0 \quad \dots (9)$$

$$uc \times \sum_{i=1}^n \frac{(y_i - y_{op})}{\sqrt{(x_i - x_{op})^2 + (y_i - y_{op})^2}} m_i = 0 \quad \dots (10)$$

Equations (9) and (10) represent a system of two equations with two unknowns, x_{op} and y_{op} , and a standard spreadsheet package was used to obtain x_{op} and y_{op} for the example problem from Table 1 that we have been working on. The calculated coordinates of the optimal location of intermediate holding facility (x_{op}, y_{op}) for the example from Table 1 are given by the following equations:

$$\begin{aligned} x_{op} &= 325 \\ y_{op} &= 75 \end{aligned} \quad \dots (11)$$

The coordinates for (x_{op}, y_{op}) from (11) are quite different than the coordinates (x_c, y_c) for the centroid in (2) and (3). The total shipping cost, TSC_{op} that is obtained from (4) with the coordinates (x_{op}, y_{op}) and data in the Table 1 is found to be $TSC_{op} = 492,624 \times uc$. Recall that the total shipping cost that was obtained with the coordinates of the centroid resulted in $TSC_c = 561,003 \times uc$. The obtained value of total shipping cost is found to be approximately 14 percent greater with the centroid solution than the optimal solution that we obtained directly, regardless of the value of uc that is assumed.

If the coordinates of the centroid location truly minimized total shipping cost, as stated in the quotes from a sample of reference books, then coordinates of the centroid (x_c, y_c) would have been same as (x_{op}, y_{op}) and thus the total shipping cost would have been the minimum possible cost that could be obtained. This example clearly shows that the centroid method for facility location obviously does not really do what the textbooks claim that it does.

As further evidence, a standard software package was used to generate a three-dimensional diagram of the total shipping cost that is obtained as a result of locating the new intermediate holding facility at general coordinates (x, y) in

our example problem from Table 1. This diagram is shown in Figure 2, and one can clearly see that the minimal total shipping cost is at the point when the intermediate holding facility is located at the point (325,75), which is the same point that we obtained above as (x_{op}, y_{op}) . As a result, we can conclude that our value of TSC that resulted from (x_{op}, y_{op}) is a global minimum solution, and not a local minimum, for the example problem that we are considering.

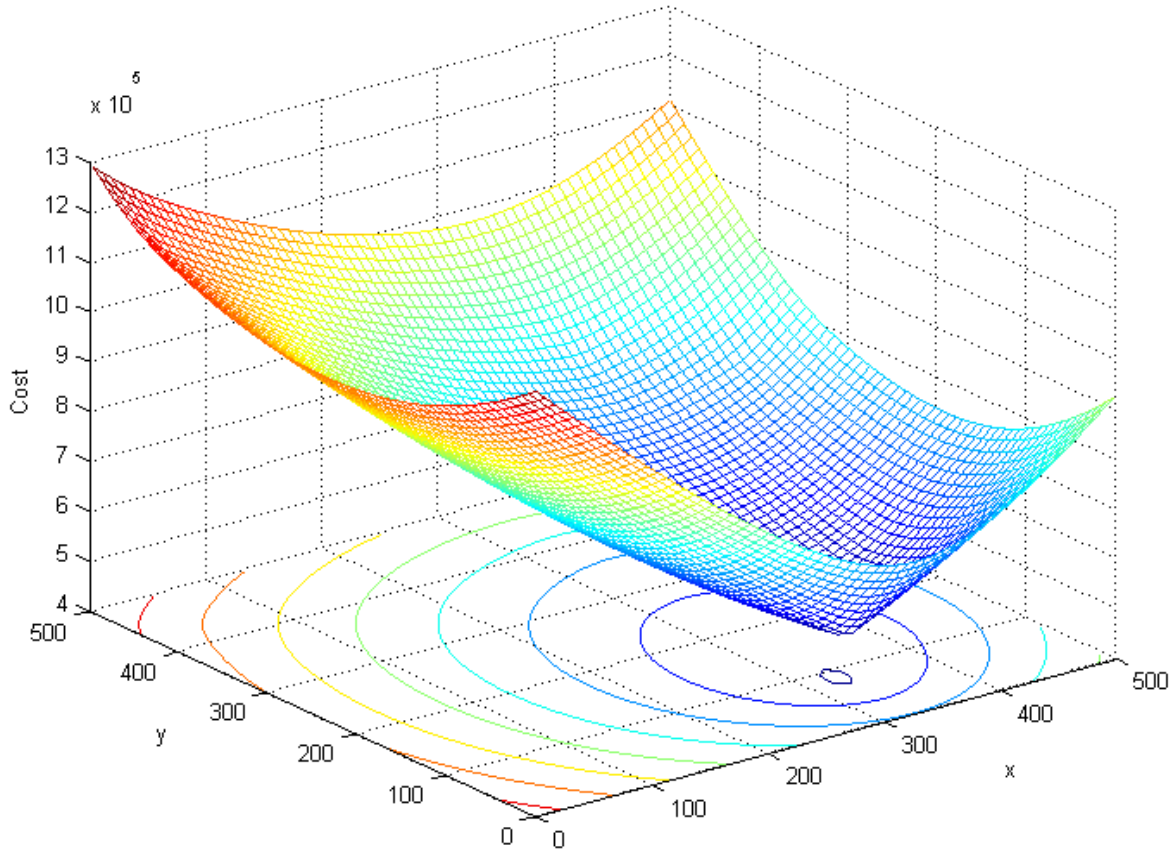


Figure 2. Total shipping cost for location (x,y) .

4. THE PHYSICAL INTERPRETATION OF THE CENTER OF GRAVITY AND ITS MISUSE

We now investigate what goes wrong when we try to apply the centroid method to facility location. Researchers in the area of facility location borrowed the notion of using the centroid from the physical sciences, where the centroid is equivalent to the center of gravity of a set of masses that are distributed in various positions in three-dimensional space. Let us consider Figure 3, where O is the origin of the space, O_i is any other point in the space, and point cg is the center of gravity for n objects with respective masses, m_i . Vector \vec{r}_i connects O and m_i , is defined as:

$$\vec{r}_i = x_i\vec{i} + y_i\vec{j} + z_i\vec{k} \text{ and } r_i^2 = x_i^2 + y_i^2 + z_i^2 \quad \dots (12)$$

where $\vec{i}, \vec{j}, \vec{k}$ are unit vectors along x, y, z axis respectively. Here, r_i^2 is the squared value of the Euclidean length of \vec{r}_i .

The moment of inertia of m_i for O is defined by $m_i r_i^2$, and moment of inertia of all masses m_i for the point O is denoted as I_o and it is given by the sum of their respective moments which is written in equations (13) and (14):

$$I_o = \sum_{i=1}^n m_i r_i^2 \quad \dots (13)$$

or:

$$I_o = \sum_{i=1}^n m_i (x_i^2 + y_i^2 + z_i^2) \quad \dots (14)$$

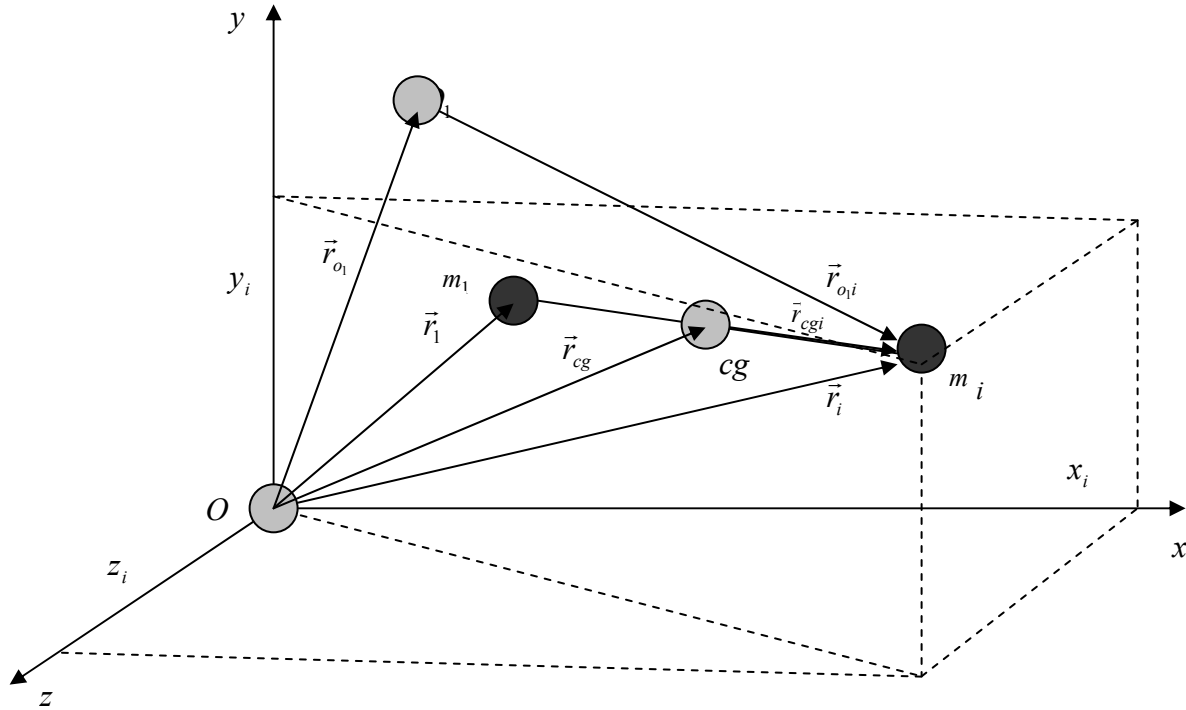


Figure 3. Interpretation of the center of gravity

Equation (14) can be rewritten in the following form:

$$I_o = \sum_{i=1}^n m_i x_i^2 + \sum_{i=1}^n m_i y_i^2 + \sum_{i=1}^n m_i z_i^2 \quad \dots (15)$$

Following the logic of the definition of the location of the center of gravity from the introduction, the coordinates of the center of gravity, cg , of all masses can be extended to three dimensions with

$$x_{cg} = \frac{\sum_{i=1}^n m_i x_i}{\sum_{i=1}^n m_i}; y_{cg} = \frac{\sum_{i=1}^n m_i y_i}{\sum_{i=1}^n m_i}; z_{cg} = \frac{\sum_{i=1}^n m_i z_i}{\sum_{i=1}^n m_i} \quad \dots (16)$$

Let us assume that O is the center of gravity, such that x_{cg}, y_{cg}, z_{cg} are equal to zero. Equation (15) becomes:

$$\sum_{i=1}^n m_i x_i = 0; \sum_{i=1}^n m_i y_i = 0; \sum_{i=1}^n m_i z_i = 0 \quad \dots (17)$$

It can be seen from Figure 3 that in the special case when point O is the center of gravity point, cg , then $\vec{r}_i = \vec{r}_{cgi}$, or:

$$x_i = x_{cgi}; \quad y_i = y_{cgi}; \quad z_i = z_{cgi} \quad \dots (18)$$

Substituting (18) into (17) one can get the following equation:

$$\sum_{i=1}^n m_i x_{cgi} = 0; \quad \sum_{i=1}^n m_i y_{cgi} = 0; \quad \sum_{i=1}^n m_i z_{cgi} = 0 \quad \dots (19)$$

When O is not the center of gravity, suppose that we consider the vector \vec{r}_i from O to m_i in two stages, going first from O to O_1 with \vec{r}_{o1} and then from O_1 to m_i with \vec{r}_{o1i} . It can be seen from Figure 3 that:

$$\vec{r}_i = \vec{r}_{o1} + \vec{r}_{o1i}; \quad r_i^2 = (x_{o1} + x_{o1i})^2 + (y_{o1} + y_{o1i})^2 + (z_{o1} + z_{o1i})^2 \quad \dots (20)$$

Combining equations (13) and (20) one can obtain:

$$I_o = \sum_{i=1}^n m_i (x_{o1} + x_{o1i})^2 + \sum_{i=1}^n m_i (y_{o1} + y_{o1i})^2 + \sum_{i=1}^n m_i (z_{o1} + z_{o1i})^2 \quad \dots (21)$$

Equation (21) can be rewritten in the following form:

$$I_o = \sum_{i=1}^n m_i (x_{o1}^2 + y_{o1}^2 + z_{o1}^2) + \sum_{i=1}^n m_i (x_{o1i}^2 + y_{o1i}^2 + z_{o1i}^2) + 2 \sum_{i=1}^n m_i (x_{o1i} x_{o1} + y_{o1i} y_{o1} + z_{o1i} z_{o1}) \quad \dots (22)$$

or:

$$I_o = \sum_{i=1}^n m_i r_{o1}^2 + \sum_{i=1}^n m_i r_{o1i}^2 + 2 \sum_{i=1}^n m_i (x_{o1i} x_{o1} + y_{o1i} y_{o1} + z_{o1i} z_{o1}) \quad \dots (23)$$

Introducing $I_o = \sum_{i=1}^n I_{o1i}$, or $I_{o1} = \sum_{i=1}^n m_i r_{o1i}^2$, equation (23) becomes:

$$I_o = r_{o1}^2 \sum_{i=1}^n m_i + I_{o1} + 2 \sum_{i=1}^n m_i (x_{o1i} x_{o1} + y_{o1i} y_{o1} + z_{o1i} z_{o1}) \quad \dots (24)$$

Let $O_1 = cg$. From equation (24) one can obtain:

$$I_o = r_{cg}^2 \sum_{i=1}^n m_i + I_{cg} + 2 \sum_{i=1}^n m_i (x_{cgi} x_{cg} + y_{cgi} y_{cg} + z_{cgi} z_{cg}) \quad \dots (25)$$

Equation (25) can be rewritten in the following form:

$$I_o = r_{cg}^2 \sum_{i=1}^n m_i + I_{cg} + 2(x_{cg} \sum_{i=1}^n m_i x_{cgi} + y_{cg} \sum_{i=1}^n m_i y_{cgi} + z_{cg} \sum_{i=1}^n m_i z_{cgi}) \quad \dots (26)$$

Substituting (19) into (26) there is:

$$I_o = I_{cg} + r_{cg}^2 \sum_{i=1}^n m_i \quad \dots (27)$$

Equation (27) shows that from all possible points for which we calculate the moment of inertia, the minimal moment of inertia is when O is the center of gravity cg , with $r_{cg} = 0$.

In general \vec{r}_{cgi} is the vector connecting cg and m_i . Then, vector \vec{r}_{cgi} is defined in equation (28), while scalar product $\vec{r}_{cgi} \cdot \vec{r}_{cgi}$ is given in equations (29) and (30).

$$\vec{r}_{cgi} = \vec{r}_i - \vec{r}_{cg} \quad \dots (28)$$

$$\vec{r}_{cgi} \cdot \vec{r}_{cgi} = (\vec{r}_i - \vec{r}_{cg}) \cdot (\vec{r}_i - \vec{r}_{cg}) \quad \dots (29)$$

$$r_{cgi}^2 = (x_i - x_{cg})^2 + (y_i - y_{cg})^2 + (z_i - z_{cg})^2 \quad \dots (30)$$

The moment of inertia for point cg can be obtained as:

$$I_{cg} = \sum_{i=1}^n m_i r_{cgi}^2 \quad \dots (31)$$

Substituting (30) into (31) yields to:

$$I_{cg} = \sum_{i=1}^n m_i (x_i - x_{cg})^2 + \sum_{i=1}^n m_i (y_i - y_{cg})^2 + \sum_{i=1}^n m_i (z_i - z_{cg})^2 \quad \dots (32)$$

We want to find extreme value (minimum) of the function I_{cg} , defined by the equation (32). Since x_{cg} , y_{cg} , and z_{cg} , are unknown variables in the equation (32), partial derivatives with respect to x_{cg} , y_{cg} , and z_{cg} have to be obtained. Thus we have:

$$\frac{\partial I_{cg}}{\partial x_{cg}} = 2 \sum_{i=1}^n m_i (x_i - x_{cg})(-1); \quad \frac{\partial I_{cg}}{\partial y_{cg}} = 2 \sum_{i=1}^n m_i (y_i - y_{cg})(-1); \quad \frac{\partial I_{cg}}{\partial z_{cg}} = 2 \sum_{i=1}^n m_i (z_i - z_{cg})(-1) \quad \dots (33)$$

In order to obtain extreme value (minimum) of the moment of inertia, partial derivatives of I_{cg} with respect to x_{cg} , y_{cg} , and z_{cg} , given in the equations (33), should be equal to zero, as written in the following equations:

$$2 \sum_{i=1}^n m_i x_{cg} - 2 \sum_{i=1}^n m_i x_i = 0; \quad 2 \sum_{i=1}^n m_i y_{cg} - 2 \sum_{i=1}^n m_i y_i = 0; \quad 2 \sum_{i=1}^n m_i z_{cg} - 2 \sum_{i=1}^n m_i z_i = 0 \quad \dots (34)$$

From equations (34) we get formulas for x_{cg} , y_{cg} , and z_{cg} :

$$x_{cg} = \frac{\sum_{i=1}^n m_i x_i}{\sum_{i=1}^n m_i}; \quad y_{cg} = \frac{\sum_{i=1}^n m_i y_i}{\sum_{i=1}^n m_i}; \quad z_{cg} = \frac{\sum_{i=1}^n m_i z_i}{\sum_{i=1}^n m_i} \quad \dots (35)$$

Equations (35) show that the coordinates of the centroid are obtained by calculating the minimum of the quadratic function $I_{cg} = \sum_{i=1}^n m_i r_{cgi}^2$, and not by calculating the minimum of the linear function $I_{cg} = \sum_{i=1}^n m_i r_{cgi}$, as stated in the center of gravity method which treats total transportation cost as a linear function of both the distance of travel and the quantity shipped.

5. CONCLUSION

The centroid method is widely reported to be a procedure that can be used to locate a central facility to minimize the total shipping cost to supply a product to a number of different locations. The total shipping cost for such a problem is simply the weighted sum of distances over which the product must be shipped to meet the demand requirements at the locations that are to be supplied. Analysis is presented here to show that by calculating the coordinates of the centroid, we are actually minimizing the weighted sum of the squares of distances in facility location problems. As a result, the centroid method is a totally inappropriate model to use for any facility location problem in which shipping costs are proportional to the total distance over which a product is shipped. If transportation costs were actually proportional to the square of the distances between facilities under some very unusual circumstances, then the centroid method would give a correct result. A possible application for which the centroid method might be appropriate is in the determination of the location of a direct service providing facility, such that customers must instead travel from the locations to a central service provider. It is quite plausible that customers would have a non-linear disutility for travel distances that are required to get to this central facility, and the total disutility for all customers would be the 'cost' that the service provider would be attempting to minimize. The centroid method would be applicable in such a scenario if the customers' disutility increased as the square of the distance that must be travelled to get to the central facility. It is quite difficult to determine any general class of situations for which the centroid method would produce an optimal solution to the facility location problem.

6. REFERENCES

1. Chase, R., Jacobs, F. R., Aquilano, N. J. (2006). Operations Management for Competitive Advantage. 11th edition, McGraw Hill Irwin, New York, USA.
2. Heizer, J., Render, B. (2005). Operations Management. 7th edition, Pearson Education, Inc., Prentice Hall, New Jersey, USA.
3. Krajewski, L. J., Ritzman, L. P. (2005). Operations Management – Processes and Value Chains, 7th edition, Pearson Education, Inc., Prentice Hall, New Jersey, USA.
4. ReVelle, C. S., Eiselt, H. A. (2005). Location analysis: A synthesis and survey. European Journal of Operational Research, Vol. 165, 1: 1-19.
5. Russell, R., Taylor, B. (2006). Operations Management – Quality and Competitiveness in a Global Environment. 5th edition, John Wiley & Sons, Inc, New Jersey, USA.
6. Shim, J. K., Siegel, J. G. (1999). Operations Management. Barrons Educational Series Inc., New York, USA.
7. Stevenson, W. (2005). Operations Management. 8th edition, McGraw Hill Irwin, New York, USA.

BIOGRAPHICAL SKETCH



Dr. William V. Gehrlein is Professor of Business Administration at the University of Delaware, where he teaches Operations Management. He did his graduate studies at Pennsylvania State University and is the author of numerous papers and several books in the areas of operations management, graph theory, statistics, decision theory and social choice theory.



Dr. Mugdim Pasic is Professor at the Industrial Engineering and Management Department of the Mechanical Engineering Faculty, University of Sarajevo, where he teaches graduate and undergraduate courses. He received M.Sc. degree at the University of California Los Angeles (UCLA), and Ph.D. degree at the University of Sarajevo. He is the author of numerous papers and several books in the area of Industrial Engineering and Management.
