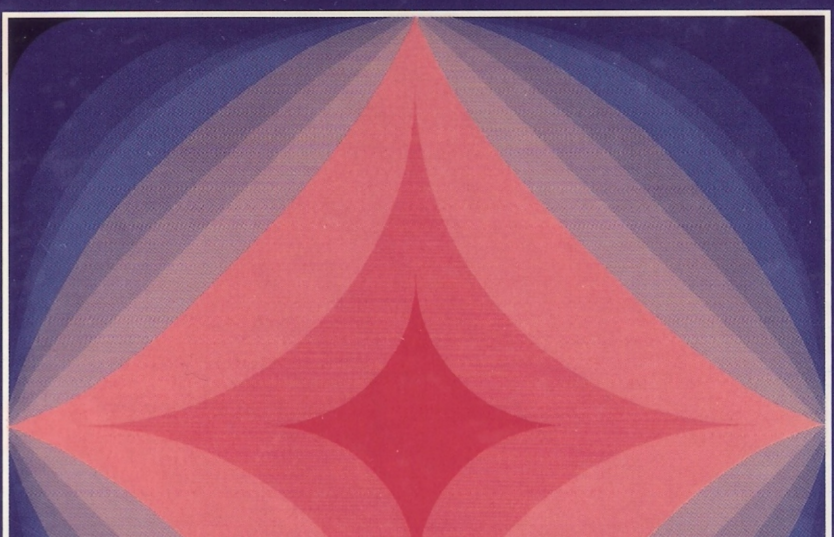
FACILITIES LOCATION

MODELS & METHODS

Robert F. Love
James G. Morris
George O. Wesolowsky



NORTH-HOLLAND

PUBLICATIONS IN OPERATIONS RESEARCH SERIES

Saul I. Gass, *Editor*

Volume I
MANPOWER PLANNING MODELS
Richard C. Grinold and **Kneale T. Marshall**

Volume 2
FIRE DEPARTMENT DEPLOYMENT ANALYSIS
The Rand Fire Project

Volume 3
QUEUEING TABLES AND GRAPHS
Frederick S. Hillier and **Oliver S. Yu**

Volume 4
ADVANCED TECHNIQUES IN THE PRACTICE OF OPERATIONS RESEARCH
Harvey J. Greenberg, Frederic H. Murphy, and **Susan H. Shaw,** *Editors*

Volume 5
IMPACTS OF MICROCOMPUTERS ON OPERATIONS RESEARCH
Saul I. Gass, Harvey J. Greenberg, Karla L. Hoffman, and
R. Warren Langley, *Editors*

Volume 6
A MATHEMATICAL ORGANIZATION THEORY
R. F. Drenick

Volume 7
FACILITIES LOCATION: MODELS AND METHODS
Robert F. Love, James G. Morris, and **George O. Wesolowsky**

FACILITIES LOCATION

MODELS & METHODS

Robert F. Love
Department of Management Science
and Information Systems
Faculty of Business
McMaster University
Hamilton, Ontario

James G. Morris
Department of Quantitative Analysis
School of Business
University of Wisconsin–Madison
Madison, Wisconsin

George O. Wesolowsky
Department of Management Science
and Information Systems
Faculty of Business
McMaster University
Hamilton, Ontario

North-Holland
New York • Amsterdam • London

Elsevier Science Publishing Co., Inc.
52 Vanderbilt Avenue, New York, New York 10017

Distributors outside the United States and Canada:

Elsevier Science Publishers B.V.
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Library of Congress Cataloging-in-Publication Data

Love, Robert F.
Facilities location.

(Publications in operations research series; v. 7)
Bibliography: p.
Includes index.
1. Industry – Location – Mathematical models.
I. Morris, J.G. (James G.) II. Wesolowsky, George O.
III. Title. IV. Series.
HD58.L68 1988 338.6'042'0724 86-748
ISBN 0-444-01031-9

Composition by Impressions, Inc., Madison, Wisconsin.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

Manufactured in the United States of America

To Our Parents

Contents

Endre Vaszonyi Weiszfeld used his second name after immigrating to the United States during the early part of World War II. He is well known to management scientists as Andrew Vaszonyi.* He started his work while studying geometry in secondary school. An earlier paper Weiszfeld (1936) was a geometrical treatment of a four-point location problem in three-dimensional space.

*This information was given verbally to one of the authors by Professor Vaszonyi.

Chapter 2

Introduction to Single-Facility Location

This chapter treats the more common versions of the single-facility location problem. In addition to the usefulness of the models themselves, the methods of their analysis introduce many concepts and techniques that will be employed in treating the more complex models of subsequent chapters.

We will first consider location problems where both the new facility location and the existing facility locations are treated mathematically as points, and where demands and costs are known. Distances are found according to one of the distance metrics discussed in Chapter 1; for additional discussion of distances, refer to Chapter 10. All transportation costs are assumed to be proportional to distance. Finding the optimal location of the new facility is equivalent to solving the following optimization problem:

$$\underset{X}{\text{minimize}}\ W(X) = \sum_{j=1}^{n} w_j \ell_p(X,a_j), \tag{2.1}$$

where

n is the number of existing facilities (or "demand points"),

w_j converts the distance between the new facility and existing facility j into cost and $w_j > 0$,

$X = (x_1,x_2)$ is the location of the new facility on the plane,

$a_j = (a_{j1},a_{j2})$ is the location of existing facility j,

$\ell_p(X,a_j)$ is the distance between the new facility and existing facility j, and

$\ell_p(X,a_j) = (|x_1 - a_{j1}|^p + |x_2 - a_{j2}|^p)^{1/p},\ p \geq 1.$

In the subsequent three sections, $W(X)$ is minimized for straight-line, rectangular, and general (ℓ_p) distances. What follows is a hypothetical example, created out of whole cloth, of a typical location problem.

Example 2.1 Bulk shipments of an industrial chemical arrive in 10-ton modules at a railway depot. The users of the chemical are clustered some distance from the depot and order the chemical in relatively small lots to avoid storage and inventory expenses. The supplier of the chemical is therefore considering the best location on which to construct a warehouse that would receive modules from the depot and then distribute them to the users.

Sacrificing realism for brevity, let us assume that there are only four users. Their locations and demands per year are given in Figure 2.1. Let us further assume that the cost per module per mile is \$20.00 for transportation from the depot and \$8.00 per ton per mile for distribution from the warehouse to the users.

Making the colossal assumption that all relevant cost structures in the system have been specified, $W(X)$ as given in problem (2.1), can now be constructed. There are five existing facilities, including the depot. The weights can be calculated as in Table 2.1. If we read the coordinates (a_{j1}, a_{j2}) of the existing facilities from Figure 2.1, and assume, for example, that $p = 1.7$, we have:

$$\begin{aligned} W(X) = {} & 400(|x_1 - 1|^{1.7} + |x_2 - 9|^{1.7})^{1/1.7} \\ & + 80(|x_1 - 2|^{1.7} + |x_2 - 5|^{1.7})^{1/1.7} \\ & + 240(|x_1 - 6|^{1.7} + |x_2 - 5|^{1.7})^{1/1.7} \\ & + 160(|x_1 - 7|^{1.7} + |x_2 - 10|^{1.7})^{1/1.7} \\ & + 220(|x_1 - 15|^{1.7} + |x_2 - 2|^{1.7})^{1/1.7}. \end{aligned}$$

The problem is to find the location (x_1, x_2) that minimizes $W(X)$.

2.1 THE STRAIGHT-LINE DISTANCE PROBLEM

We now turn to a mathematical description of the straight-line (Euclidean) distance problem and to two of its very basic properties. To solve problem (2.1), a single new facility must be located among n existing facilities on the plane in accordance with the criterion that the sum of weighted distances be minimized. Because the straight-line distance ($p = 2$) between the new facility at (x_1, x_2) and an existing facility at (a_{j1}, a_{j2}) is:

$$\ell_2(X, a_j) = ((x_1 - a_{j1})^2 + (x_2 - a_{j2})^2)^{1/2},$$

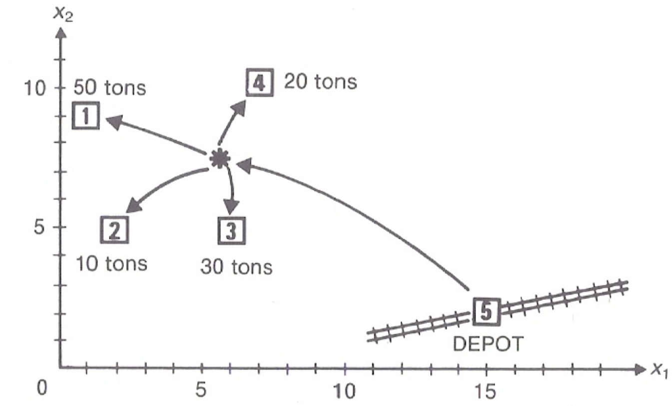


Figure 2.1 Schematic of a Facility Location Problem.

problem (2.1) becomes:

$$\underset{X}{\text{minimize}}\ W(X) = \sum_{j=1}^{n} w_j((x_1 - a_{j1})^2 + (x_2 - a_{j2})^2)^{1/2}. \qquad (2.2)$$

It is useful at this point in the discussion to introduce the concept of a convex function. A function $f(X)$ is said to be *convex* if the line segment between any two points, $[X^1, f(X^1)]$ and $[X^2, f(X^2)]$, on the graph of the function never lies below the graph. Formally, this means $f(X)$ is convex if

$$f[\lambda X^1 + (1 - \lambda)X^2] \leq \lambda f(X^1) + (1 - \lambda) f(X^2)$$

for all X^1 and X^2 and any $\lambda \epsilon [0,1]$. The notation X^1 and X^2 is used to denote two distinct points in the domain of f. A function $f(X)$ is said to be *strictly convex* if the above inequality holds as a strict inequality for all distinct X^1 and X^2 and any $\lambda \epsilon (0,1)$. Strict convexity of f means that the line segment lies strictly above the graph except at the two endpoints of the segment.

Table 2.1 Calculation of Weights.

j	w_j
1	50 × 8 = 400
2	10 × 8 = 80
3	30 × 8 = 240
4	20 × 8 = 160
5	11 × 20 = 220

Property 2.1 [*Convexity of $w_j\ell_2(X,a_j)$*]. *$w_j\ell_2(X,a_j)$ is a convex function of X.*

The proof of convexity is given in the Appendix, mathematical note 2.1. Note that this property would not hold if the weight were negative. It can be shown (see Exercise 2.1) that the sum of convex functions is convex, and hence $W(X)$ itself is convex. This means local optima are global optima for problem (2.2), and $W(X)$ has no inflection points. With this information, we are assured that the extremal equations for $W(X)$ can produce only global optima for problem (2.2). These equations are:

$$\frac{\partial W(X)}{\partial x_k} = \sum_{j=1}^{n} \frac{w_j(x_k - a_{jk})}{\ell_2(X,a_j)} = 0 \quad \text{for } k = 1,2. \tag{2.3}$$

One difficulty immediately presents itself. The derivatives in equation (2.3) are undefined if $\ell_2(X,a_j) = 0$. Therefore, if an optimal location for the new facility coincides with that of an existing facility, equation (2.3) cannot be used to check optimality. Fortunately, we can easily check *each* existing facility location for optimality.

Property 2.2 [*Minimum of $W(X)$ at an existing facility location*]. *$W(X)$ is minimized at the r^{th} existing facility location (a_{r1},a_{r2}) if, and only if:*

$$CR_r = \left[\left(\sum_{\substack{j=1\\ \neq r}}^{n} \frac{w_j(a_{r1} - a_{j1})}{\ell_2(a_r,a_j)}\right)^2 + \left(\sum_{\substack{j=1\\ \neq r}}^{n} \frac{w_j(a_{r2} - a_{j2})}{\ell_2(a_r,a_j)}\right)^2\right]^{1/2} \leqslant w_r. \tag{2.4}$$

A ready explanation for Property 2.2 is provided by the analog model in Exercise 2.2. A derivation is in the Appendix, mathematical note 2.2. There now exist many different iterative methods for finding a solution to problem (2.1). One of the oldest, as well as perhaps the simplest, follows.

In addressing the problem of minimizing $W(X)$, let us temporarily ignore the possibility that the new facility location will coincide with an existing facility location. A procedure for iterating to the optimum location can be obtained by rewriting equation (2.3) so that we have one equation for x_1 and one for x_2:

$$x_k = \frac{\displaystyle\sum_{j=1}^{n} \frac{w_j a_{jk}}{\ell_2(X,a_j)}}{\displaystyle\sum_{j=1}^{n} \frac{w_j}{\ell_2(X,a_j)}} \quad \text{for } k = 1,2. \tag{2.5}$$

Note that x_k is not really isolated on the left-hand side in equation (2.5) because each $\ell_2(X,a_j)$ is a function of x_k. However, equation (2.5) can be used iteratively to approach (x_1^*,x_2^*), the optimum new facility

location. We refer to this procedure as the Weiszfeld procedure (see Chapter 1). Let us imagine that we have just completed iteration ℓ and have obtained the location $(x_1^{(\ell)},x_2^{(\ell)})$. We can then use equation (2.5) to find the estimate of iteration $(\ell+1)$:

$$x_k^{(\ell+1)} = \frac{\displaystyle\sum_{j=1}^{n} \frac{w_j a_{jk}}{\ell_2(X^{(\ell)},a_j)}}{\displaystyle\sum_{j=1}^{n} \frac{w_j}{\ell_2(X^{(\ell)},a_j)}} \quad \text{for } k = 1,2. \tag{2.6}$$

Before beginning the iterations, an initial location $(x_1^{(0)},x_2^{(0)})$ is required. An expedient choice is the solution to the squared Euclidean distance problem, which is the same as problem (2.2) except that each distance $\ell_2(X,a_j)$ is squared. As shown in Exercise 1.2, the center-of-gravity location solves the squared Euclidean distance problem; hence, our starting point for procedure (2.6) is:

$$x_k^{(0)} = \frac{\displaystyle\sum_{j=1}^{n} w_j a_{jk}}{\displaystyle\sum_{j=1}^{n} w_j} \quad \text{for } k = 1,2. \tag{2.7}$$

The iterations will converge to an optimal location, provided that neither an iterate nor an optimal new facility location is at an existing facility location. References for convergence properties are given at the end of this chapter. Though we would expect convergence difficulties with the iterations when (x_1^*,x_2^*) coincides with an existing facility location, these difficulties do not generally materialize. When CR_r in condition (2.4) is much smaller than w_r, convergence is fairly rapid; as $\ell_2(X,a_r)$ in equation (2.5) approaches zero, computational difficulties could eventually arise. Experience shows that when condition (2.4) is near equality (whether or not it is met), convergence may be slow. It is therefore helpful to check all existing facility locations using condition (2.4) *first*. If condition (2.4) is not met, but CR_r is nearly equal to w_r at some existing facility r, iterations could be started near that point. Convergence may also be slow for certain weight structures. We are compensated by the fact that in such cases the cost "bowl" is usually shallow in the vicinity of the optimum solution.

The practical question of when to terminate the procedure can be answered by the use of a stopping criterion that employs a lower bound on the optimum value of $W(X)$. This lower bound is continually updated during the iterations. To derive this bound we need the following property.

Property 2.3 [*Dominance of the convex hull*]. *X^*, an optimal solution to problem (2.1), must lie within Ω, the convex hull of the existing facility locations.*

The *convex hull* is defined as the smallest convex polygon that contains all the existing facility locations. A proof of Property 2.3 is the subject of Exercise 2.5b. It can be shown that the starting point defined by equation (2.7) is in the convex hull. Further, regardless of the starting point used for procedure (2.6), the next point will be in the convex hull (see Exercise 2.5a). We will now give a geometrical rationale for a stopping criterion for the Weiszfeld procedure.

To begin the discussion, recall that $W(X)$ is a convex function. This means that a plane tangent to the convex bowl-shaped graph of $W(X)$ at a given point $X^{(\ell)}$ underestimates $W(X)$ for any X. In particular, if the partial derivatives given in equation (2.3) exist, this means

$$W(X) \geqslant W(X^{(\ell)}) + \sum_{k=1}^{2} [\partial W(X^{(\ell)})/\partial x_k](x_k - x_k^{(\ell)})$$

for any X. The gradient $\nabla W(X^{(\ell)})$ is a vector with components given by the respective partial derivatives. Choosing $X = X^*$ we can write:

$$W(X^*) \geqslant W(X^{(\ell)}) + \nabla W(X^{(\ell)})\cdot(X^* - X^{(\ell)})$$

($\cdot$ denotes scalar product)

$$\geqslant W(X^{(\ell)}) - |\nabla W(X^{(\ell)})\cdot(X^* - X^{(\ell)})|$$

$$\geqslant W(X^{(\ell)}) - \|\nabla W(X^{(\ell)})\| \, \|X^* - X^{(\ell)}\|$$

(since $|u\cdot v| \leqslant \|u\| \, \|v\|$; Schwartz inequality)

where $\|\cdot\|$ denotes the magnitude of a vector. But X^* and $X^{(\ell)}$ are both in the convex hull Ω for $\ell \geqslant 1$. Hence $\|X^* - X^{(\ell)}\|$ cannot be greater than the straight-line distance $\sigma(X^{(\ell)})$, say, between $X^{(\ell)}$ and the point in Ω furthest away from $X^{(\ell)}$. Therefore, an upper bound on the improvement in $W(X)$ must be $\|\nabla W(X^{(\ell)})\|\sigma(X^{(\ell)})$. It is now possible to state the following property.

Property 2.4 [*Lower bound on* $W(X^*)$].

$$W(X^*) \geqslant LB^{(\ell)} = W(X^{(\ell)}) - \|\nabla W(X^{(\ell)})\|\sigma(X^{(\ell)}) \qquad (2.8)$$

where $\sigma(X^{(\ell)}) = \max_{y \in \Omega} \{\ell_2(X^{(\ell)}, y)\}$.

It is, therefore, possible to know an upper bound on further improvement in the objective function value at every iteration in the Weiszfeld procedure. A stopping criterion based on proportional suboptimality can be set using:

$$S^{(\ell)} = \frac{\|\nabla W(X^{(\ell)})\|\sigma(X^{(\ell)})}{LB^{(\ell)}}. \qquad (2.9)$$

Table 2.2 Evaluation of CR_r.

Existing Facility Location	r	CR_r	w_r
(1,1)	1	7.635	1
(1,4)	2	4.931	2
(2,2)	3	4.453	2
(4,5)	4	4.754	4

From inequality (2.8), whenever $LB^{(\ell)} > 0$ we have:

$$S^{(\ell)} \geqslant (W(X^{(\ell)}) - (W(X^*))/W(X^*).$$

Hence,if we wish to find an iterate $X^{(\ell)}$ with relative suboptimality bounded as $(W(X^{(\ell)}) - W(X^*))/W(X^*) < \epsilon$, iterations need only be continued until $S^{(\ell)} < \epsilon$. For example, if iterations are terminated when $S^{(\ell)} < 0.001$, then $W(X^{(\ell)})$ is within 0.1% of the minimal value $W(X^*)$. It should be mentioned that tighter lower bounds can be obtained (see the Appendix, mathematical note 2.5). This one is presented due to its geometrical simplicity.

Example 2.2 Four existing facilities are located at the points (1,1), (1,4), (2,2), and (4,5). The corresponding weights are 1, 2, 2, and 4, respectively. Values of CR_r are given in Table 2.2. Because CR_r is always greater than w_r, the optimum location of the new facility does not coincide with any of the existing facility locations. Table 2.3 shows iterations from the center-of-gravity starting point, which in this example was quite close to the optimum location. As Exercise 2.4 shows, this is not always so.

When iterations were started at (0,1000), $X^{(2)}$ was (2.557,3.669), thus demonstrating the insensitivity of the procedure to the starting point.

Table 2.3 Iterations for Example 2.2.

Iteration Number ℓ	New Facility Location	Cost	$LB^{(\ell)}$	$S^{(\ell)}$
0	(2.556,3.667)	17.646	16.407	7.55×10^{-2}
1	(2.523,3.745)	17.624	17.210	2.41×10^{-2}
2	(2.527,3.772)	17.621	17.382	1.376×10^{-2}
3	(2.536,3.785)	17.620	17.441	1.024×10^{-2}
4	(2.544,3.794)	17.619	17.481	7.934×10^{-3}
5	(2.551,3.799)	17.619	17.511	6.192×10^{-3}
6	(2.557,3.804)	17.619	17.534	4.847×10^{-3}
7	(2.560,3.807)	17.619	17.552	3.802×10^{-3}
8	(2.564,3.810)	17.619	17.566	2.987×10^{-3}
9	(2.567,3.812)	17.619	17.577	2.350×10^{-3}
10	(2.569,3.814)	17.619	17.586	1.851×10^{-3}
15	(2.576,3.818)	17.618	17.608	5.660×10^{-4}
20	(2.577,3.820)	17.618	17.615	1.742×10^{-4}
40	(2.577,3.820)	17.618	17.618	1.580×10^{-6}

Actually, as can easily be verified from a comparison of procedure (2.6) and equation (2.7), if the starting point is very far outside the convex hull, $X^{(1)}$ will be approximately the center-of-gravity solution.

2.2 THE RECTANGULAR DISTANCE PROBLEM

Rectangular distances were described in Chapter 1. In addition to being applicable in a wide variety of location problems, their use greatly simplifies many problems that are quite difficult with straight-line distances. For $p = 1$, problem (2.1) becomes:

$$\underset{X}{\text{minimize}}\ W(X) = \sum_{j=1}^{n} w_j(|x_1 - a_{j1}| + |x_2 - a_{j2}|). \qquad (2.10)$$

There are two properties that are useful in finding (x_1^*, x_2^*). The first will be called separability. Problem (2.10) can be rewritten:

$$\underset{X}{\text{minimize}}\ W(X) = W_1(x_1) + W_2(x_2) \qquad (2.11)$$

where:

$$W_1(x_1) = \sum_{j=1}^{n} w_j|x_1 - a_{j1}|,$$

$$W_2(x_2) = \sum_{j=1}^{n} w_j|x_2 - a_{j2}|.$$

Minimizing $W(X)$ is equivalent to separately finding an x_1 that minimizes $W_1(x_1)$ and an x_2 that minimizes $W_2(x_2)$. The problem now becomes:

$$\underset{x_k}{\text{minimize}}\ W_k(x_k) = \sum_{j=1}^{n} w_j|x_k - a_{jk}| \quad \text{for } k = 1,2. \qquad (2.12)$$

This problem is rather easy to solve. An analog approach is given near the end of this section (in Example 2.5) that makes the use of formulas unnecessary. However, the method of analysis that immediately follows gives mathematical conditions and relationships that will be needed in later chapters; in addition, it validates the analog approach.

We must first verify that $w_j|x_k - a_{jk}|$ is a convex function of x_k. This can be done by simply plotting the term against x_k for any $w_j > 0$ and a_{jk}. A more formal proof is the subject of Exercise 2.7. Because the sum of convex functions is convex, it follows that $W_k(x_k)$ is convex.

To facilitate analysis, a change in notation is made in problem (2.12). Let the values of a_{jk} for $j = 1,\dots,n$ be reordered to produce $a_{(1)k} < a_{(2)k} < a_{(3)k} < \dots < a_{(n_k)k}$, and let $w_1^k,\dots,w_{n_k}^k$ be the corresponding positive weights.

Note that there will now be n_k coordinates, where $n_k \leq n$. The reason for this may be seen in the following example. Suppose $a_{11} = a_{31} = 4$ and $w_1 = 3$, $w_3 = 2$. We can combine the sum of terms $3|x_1 - 4| + 2|x_1 - 4|$ into the single term $5|x_1 - 4|$. We can thus create a sequence of $a_{(j)k}$'s that are strictly increasing in value. As a result, however, we may have fewer $a_{(j)k}$'s than there are a_{jk}'s.

The superscript k in w_j^k denotes ordering and consolidation in the corresponding weights. The reason is that there can now be an ordering of weights in the x_1 dimension different from that in the x_2 dimension, and we use the superscript k to distinguish these orderings.

Problem (2.12) becomes:

$$\underset{x_k}{\text{minimize}}\ W_k(x_k) = \sum_{j=1}^{n_k} w_j^k|x_k - a_{(j)k}| \quad \text{for } k = 1,2. \qquad (2.13)$$

The function $W_k(x_k)$ is now in a form convenient for calculating its derivative (see Exercise 2.8). We can write:

$$W_k'(x_k) = -\sum_{j=1}^{n_k} w_j^k \qquad \text{for } x_k < a_{(1)k} \qquad (2.14a)$$

$$W_k'(x_k) = \sum_{j=1}^{t} w_j^k - \sum_{j=t+1}^{n_k} w_j^k \qquad \text{for } a_{(t)k} < x_k < a_{(t+1)k} \qquad (2.14b)$$

$$W_k'(x_k) = \sum_{j=1}^{n_k} w_j^k \qquad \text{for } x_k > a_{(n_k)k}. \qquad (2.14c)$$

It is easy to see that the slope of $W_k(x_k)$ is made up of linear segments and changes only at points $a_{(j)k}$. To sum up, $W_k(x_k)$ is a continuous, convex, piecewise-linear function with points of discontinuity in the first derivatives occurring at the a_{jk}'s.

Before formally stating optimality conditions for minimizing $W_k(x_k)$, let us consider a numerical example to illustrate the notation and the characteristics of $W_1(x_1)$ and $W_2(x_2)$.

Example 2.3 A new facility must be located among four existing ones at (1,1), (2,4), (2,3), and (4,2). The corresponding weights are 2, 3, 1, and 2, respectively. We are about to solve the associated rectangular distance location problem graphically. Table 2.4 summarizes the data and notation. For example, with $k = 1$ in problems (2.12) and (2.13), we obtain:

$$W_1(x_1) = 2|x_1 - 1| + 3|x_1 - 2| + 1|x_1 - 2| + 2|x_1 - 4|$$

and

$$W_1(x_1) = 2|x_1 - 1| + 4|x_1 - 2| + 2|x_1 - 4|.$$

Table 2.4 Data for Example 2.3.

	Original Coordinates and Weights			Ordering in the x_1 Dimension		Ordering in the x_2 Dimension	
j	a_{j1}	a_{j2}	w_j	$a_{(j)1}$	w_j^1	$a_{(j)2}$	w_j^2
1	1	1	2	1	2	1	2
2	2	4	3	2	4	2	2
3	2	3	1	4	2	3	1
4	4	2	2			4	3

$W_1(x_1)$ and $W_2(x_2)$ are plotted in Figures 2.2(a) and 2.2(b), respectively. It is evident that $W_1(x_1)$ has a minimum value of 6 at $x_1^* = 2$, whereas the minimum value of $W_2(x_2)$ is 9 for x_2^* in the interval [2,3]. Hence, $W(X)$ has a minimum of 15 at (2,[2,3]). We can use equation (2.14) to check the slopes. For example, when $2 < x_1 < 4$, then $t = 2$ and we compute:

$$W'(x_1) = (2 + 4) - 2 = 4,$$

as can be verified in Figure 2.2(a).

Because the slope $W_k'(x_k)$ in equation (2.14) obviously increases with increasing t, the condition for a minimum to $W_k(x_k)$ to occur must be that the slope either changes from negative to positive, or changes from negative to zero at some point. In the latter case, the minimum occurs over a range of values for x_k, as in Figure 2.2(b). However, at least one point $a_{(t)k}$ *must* be a minimizer of $W_k(x_k)$. The following makes these claims precise.

Property 2.5 [*Conditions for a minimum to $W_k(x_k)$*]. *Suppose:*

$$\sum_{j=1}^{t-1} w_j^k - \sum_{j=t}^{n_k} w_j^k < 0 \qquad (2.15a)$$

and

$$\sum_{j=1}^{t} w_j^k - \sum_{j=t+1}^{n_k} w_j^k \geq 0 \qquad (2.15b)$$

are satisfied at some t^*. If condition (2.15b) is met as a strict inequality, then $x_k^* = a_{(t^*)k}$. If condition (2.15b) is met as an equality, then $x_k^* \in [a_{(t^*)k}, a_{(t^*+1),k}]$.

It is possible to express the conditions of Property 2.5 in a form more convenient for finding t^*. We can write condition (2.15a) as:

$$\sum_{j=1}^{t-1} w_j^k + \sum_{j=1}^{t-1} w_j^k - \sum_{j=1}^{n_k} w_j^k < 0,$$

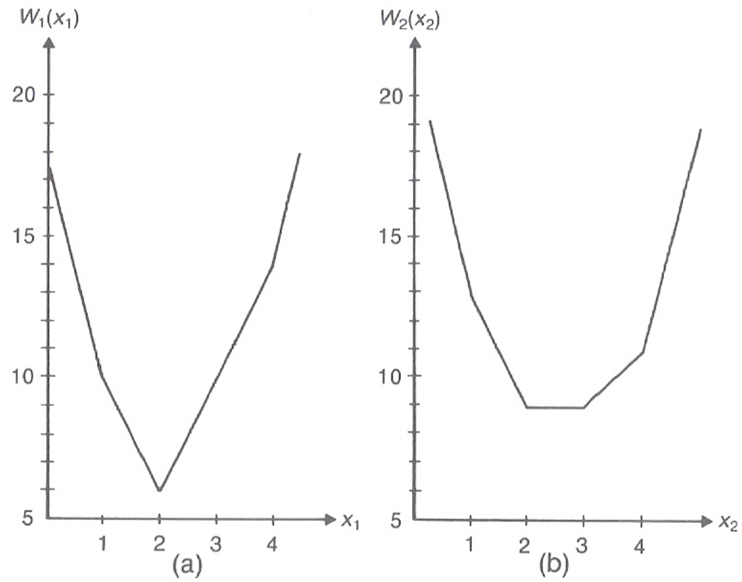


Figure 2.2 Plots of $W_1(x_1)$ and $W_2(x_2)$.

and condition (2.15b) as:

$$\sum_{j=1}^{t} w_j^k + \sum_{j=1}^{t} w_j^k - \sum_{j=1}^{t} w_j^k - \sum_{j=t+1}^{n_k} w_j^k \geq 0.$$

If we let:

$$C = \sum_{j=1}^{n_k} w_j^k \quad \text{for } k = 1,2,$$

then the conditions of Property 2.5 become:

$$-C + 2\sum_{j=1}^{t-1} w_j^k < 0 \qquad (2.16a)$$

$$-C + 2\sum_{j=1}^{t} w_j^k \geq 0. \qquad (2.16b)$$

Inequalities (2.16) now suggest a computational procedure. To $-C$ we add twice the weight of each point $a_{(t)k}$, starting from $t = 1$, until the sum first equals or exceeds zero. Note that we are merely using another expression for the slope of $W_k(x_k)$. We will then have found the optimal range or the optimal point, respectively, for x_k.

Example 2.4 We return to Example 2.3 and the data in Table 2.4 to demonstrate the use of inequalities (2.16). We first calculate:

$$C = \sum_{j=1}^{n_k} w_j^1 = 2 + 4 + 2 = 8.$$

Now, to find x_1^*, we calculate:

$$-C + 2\sum_{j=1}^{1} w_j^1 = -8 + 4 = -4$$

$$-C + 2\sum_{j=1}^{2} w_j^1 = -8 + 12 = 4 > 0$$

and find that condition (2.16b) is satisfied as a strict inequality. So $x_1^* = a_{(2)1} = 2$.

To find x_2^*, we calculate:

$$-C + 2\sum_{j=1}^{1} w_j^2 = -8 + 4 = -4$$

$$-C + 2\sum_{j=1}^{2} w_j^2 = -8 + 8 = 0$$

and find that condition (2.16b) is satisfied as an equality. So $x_2^* \epsilon [a_{(2)2}, a_{(3)2}] = [2,3]$.

But there is an easier way of solving our problem. Example 2.5 illustrates the method.

Example 2.5 We consider the same problem. Recall that there are four existing facilities at points (1,1), (2,4), (2,3), and (4,2) with weights of 2, 3, 1, and 2, respectively.

Let us imagine that the weights are "dropped" on the x_1-axis, and at the same time (in defiance of gravity) the same weights are dropped on the x_2-axis (Figure 2.3). Let us then divide the x_1-axis into two parts such that the weights are bisected. Clearly this point is at $x_1 = 2$ if we assume that weights acting on a mathematical point may be split as we wish. Similarly the weights on the x_2-axis are bisected when $2 \leqslant x_2 \leqslant 3$. As can be checked in the previous examples, these bisections produce x_1^* and x_2^*. As the reader should verify, what we have really done is to apply inequalities (2.15) or (2.16) in slightly disguised form. We can now describe the minimum to $W_k(x_k)$ as occurring at the point(s) of median weight.

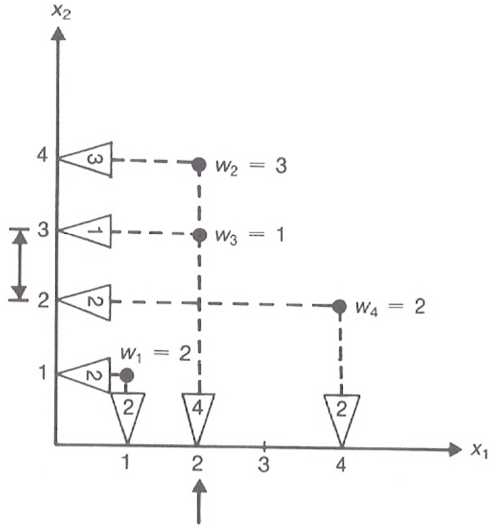


Figure 2.3 Median Weight Illustration.

2.3 THE ℓ_p DISTANCE PROBLEM

Suppose distances are modeled by the ℓ_p function. As discussed in Chapters 1 and 10, ℓ_p distances can often provide a better measure of actual travel distances than either the straight-line or rectangular distances, which, are special cases given by $p = 2$ and $p = 1$, respectively. For convenience, we repeat problem (2.1) here:

$$\underset{X}{\text{minimize}}\ W(X) = \sum_{j=1}^{n} w_j(|x_1 - a_{j1}|^p + |x_2 - a_{j2}|^p)^{1/p}. \qquad (2.17)$$

We first state two properties that characterize $\ell_p(X,a_j)$:

(i) $\ell_p(X,a_j)$ decreases as p increases, ie, for $X \neq a_j$, $(|x_1 - a_{j1}|^p + |x_2 - a_{j2}|^p)^{1/p} > (|x_1 - a_{j1}|^{p'} + |x_2 - a_{j2}|^{p'})^{1/p'}$, for $p < p'$,

and

(ii) as $p \to \infty$, $\ell_p(X,a_j)$ becomes the larger of $|x_1 - a_{j1}|$ and $|x_2 - a_{j2}|$.

The following property bears upon the optimization specified by problem (2.17). A proof is given in the Appendix, mathematical note 2.3.

Property 2.6 *[Convexity of $w_j\ell_p(X,a_j)$]. $w_j\ell_p(X,a_j)$ is a convex function of X.*

As each of the terms of $W(X)$ in problem (2.17) is convex, we can again use the fact that the sum of convex functions is convex to conclude that

$W(X)$ is a convex function of X. Hence, a local minimizer of $W(X)$ is also a global minimizer. As with $p = 2$, we can check to see aforehand whether the new facility would be optimally located at the site of some existing facility. The following criterion, which generalizes Property 2.2, is derived in Juel and Love (1981b).

Property 2.7 [*Minimum of $W(X)$ at an existing facility location; ℓ_p distances*]. *$W(X)$ is minimum at (a_{r1}, a_{r2}) if and only if:*

$$CRP_r = (|R_{r1}|^{p/(p-1)} + |R_{r2}|^{p/(p-1)})^{(p-1)/p} \leq w_r, \text{ for } p > 1, \quad (2.18a)$$

$$\max(|R_{r1}|, |R_{r2}|) \leq w_r, \text{ for } p = 1, \quad (2.18b)$$

where:

$$R_{rk} = \sum_{\substack{j=1 \\ \neq r}}^{n} \frac{w_j \operatorname{sign}(a_{rk} - a_{jk})|a_{rk} - a_{jk}|^{p-1}}{(\ell_p(a_r, a_j))^{p-1}} \text{ for } k = 1,2.$$

We observe that

$$R_{rk} = \frac{\partial}{\partial x_k} \sum_{\substack{j=1 \\ \neq r}}^{n} w_j \ell_p(X, a_j)|_{x_k = a_{rk}} \text{ in (2.18a).}$$

Terms on the left-hand side of inequality (2.18a) are not defined at $p = 1$. Using Property (ii) of ℓ_p distances and letting p decrease toward 1 so that $p' = p/(p-1) \to \infty$, we can easily deduce inequality (2.18b).

Example 2.6 Table 2.5 gives the parameters of the problem with $n = 5$ together with the point optimality calculations for $p = 1$, 1.3, 1.5, 2, and 5. We see that $(x_1^*, x_2^*) = (7,3)$ for $p = 1$ and $p = 1.3$; actually this point is optimal for $p \leq 1.385$.

It is convenient at this point to introduce a hyperbolic (see Exercise 2.11) approximation to $W(X)$ in problem (2.17). This approximation will be especially useful in the more complex multi-facility location problem discussed in a later chapter. We replace each absolute term $|y|$ in problem (2.17) by $(y^2 + \epsilon)^{1/2}$ where ϵ is a small positive number. The approximation is always larger than the original term, but approaches the original term as $\epsilon \to 0$. Problem (2.17) is then approximated by:

Table 2.5 Evaluation of CRP_r.

r	Existing Facility Location	w_r	CRP_r				
			$p=1$	$p=1.3$	$p=1.5$	$p=2$	$p=5$
1	(1,1)	3	10.0	9.380	9.249	9.056	9.218
2	(2,6)	2	5.0	6.081	6.689	7.843	9.821
3	(4,1)	1	9.0	7.804	7.356	6.766	5.999
4	(7,3)	4	3.0	3.791	4.288	5.401	7.757
5	(8,8)	3	10.0	9.674	9.502	9.220	9.249

$$\underset{X}{\text{minimize}}\ WH(X) = \sum_{j=1}^{n} w_j(((x_1 - a_{j1})^2 + \epsilon)^{p/2} + ((x_2 - a_{j2})^2 + \epsilon)^{p/2})^{1/p}. \quad (2.19)$$

We first note that $WH(X)$ is strictly convex (as proved on a do-it-yourself basis in Exercise 2.12), and that all orders of derivatives are continuous at all points (as will become apparent shortly). Therefore, $WH(X)$ is a well-behaved candidate for general nonlinear descent and programming algorithms that converge to the minimizer if the function is smooth and strictly unimodal.

The question arises: If we find the location that minimizes $WH(X)$, what progress have we made in minimizing $W(X)$? We will show that we can come as close as we wish to minimizing $W(X)$ by minimizing $WH(X)$ and simply choosing small values for ϵ.

Property 2.8 [*Maximum difference between $WH(X)$ and $W(X)$*].

$$\max_{X} \{WH(X) - W(X)\} \leq \Delta(\epsilon) = 2^{1/p}\epsilon^{1/2}\left(\sum_{j=1}^{n} w_j\right) \quad (2.20)$$

PROOF: It is shown in the Appendix, mathematical note 2.4, that:

$$(((x_1 - a_{j1})^2 + \epsilon)^{p/2} + ((x_2 - a_{j2})^2 + \epsilon)^{p/2})^{1/p} - (|x_1 - a_{j1}|^p + |x_2 - a_{j2}|^p)^{1/p} \leq 2^{1/p}\epsilon^{1/2}, \quad (2.21)$$

from which the property follows.

As the difference between $WH(X)$ and $W(X)$ never exceeds $\Delta(\epsilon)$, solving problem (2.19) will give us a solution to problem (2.17) that is at most $\Delta(\epsilon)$ from the optimum value. To see this, let X^* be a minimizer for $W(X)$ and X^{**} be the minimizer for $WH(X)$. Because $WH(X^*) - W(X^*) \leq \Delta(\epsilon)$ and $WH(X^{**}) \leq WH(X^*)$, we have $WH(X^{**}) - W(X^*) \leq \Delta(\epsilon)$. Therefore, $W(X^{**}) - W(X^*) \leq \Delta(\epsilon)$.

The Weiszfeld procedure discussed in Section 2.2 can be generalized to the ℓ_p distance case. Differentiating $WH(X)$ with respect to x_1 and x_2, setting the partial derivatives to zero, and then "isolating" x_1 and x_2 on the left-hand side (the details are requested in Exercise 2.13), we obtain:

$$x_k^{(\ell+1)} = \frac{\displaystyle\sum_{j=1}^{n} \frac{w_j a_{jk}}{d'(X^{(\ell)},a_j)d''(x_k^{(\ell)},a_{jk})}}{\displaystyle\sum_{j=1}^{n} \frac{w_j}{d'(X^{(\ell)},a_j)d''(x_k^{(\ell)},a_{jk})}}, \tag{2.22}$$

where

$$d'(X,a_j) = (((x_1 - a_{j1})^2 + \epsilon)^{p/2} + ((x_2 - a_{j2})^2 + \epsilon)^{p/2})^{1-1/p}, \text{ and}$$

$$d''(x_k,a_{jk}) = ((x_k - a_{jk})^2 + \epsilon)^{1-p/2}.$$

When $p = 2$ and $\epsilon = 0$, iterative procedure (2.22) is identical to procedure (2.6). Convergence is guaranteed for $1 \leqslant p \leqslant 2$, provided $\epsilon > 0$. Actually, the procedure will work well with $\epsilon = 0$. However, setting ϵ to a very small quantity (relative to the weights and computer resolution of zero) does protect one against the embarrassing possibility of dividing by zero on a computer. This occurs in procedure (2.6) when $X^{(\ell)}$ coincides with an existing facility location, or in procedure (2.22) when $x_k^{(\ell)}$ coincides with even one coordinate a_{jk} of an existing facility location. It is possible to inadvertently choose a starting point that falls in that category, or it is possible that $X^{(\ell)}$ will move too close to an existing facility location during iteration. In a slow convergence problem, a large initial value of ϵ may get us to the vicinity of an optimum solution relatively quickly. Then a smaller value can be used. However, the use of $\epsilon > 0$ is a practical necessity for Weiszfeld iterations only in the multi-facility case that is discussed in Chapter 4.

Table 2.6a Iterations with $p = 1.5$ and $\epsilon = 0$.

Iteration Number ℓ	$\epsilon = 0$				
	$x_1^{(\ell)}$	$x_2^{(\ell)}$	$WH(X) = W(X)$	$LB^{(\ell)}$	$SH^{(\ell)} = S^{(\ell)}$
0	4.846	4.000	54.849	44.350	2.37×10^{-1}
1	5.114	3.649	54.075	46.141	1.72×10^{-1}
2	5.334	3.448	53.662	47.670	1.26×10^{-1}
3	5.515	3.336	53.449	48.756	9.63×10^{-2}
4	5.665	3.273	53.319	49.486	7.74×10^{-2}
5	5.790	3.236	53.232	49.975	6.52×10^{-2}
10	6.171	3.164	53.050	51.171	3.67×10^{-2}
20	6.447	3.117	52.985	52.184	1.53×10^{-2}
30	6.542	3.098	52.975	52.585	7.42×10^{-3}
40	6.583	3.090	52.973	52.770	3.84×10^{-3}
50	6.604	3.086	52.972	52.864	2.06×10^{-3}
60	6.615	3.083	52.972	52.913	1.12×10^{-3}

Table 2.6b Iterations with $p = 1.5$ and $\epsilon = 0.01$.

Iteration Number ℓ	$\epsilon = 0.01$						
	$x_1^{(\ell)}$	$x_2^{(\ell)}$	$WH(X)$	$LBH^{(\ell)}$	$SH^{(\ell)}$	$W(X)$	$LB^{(\ell)}$
0	4.846	4.000	54.895	44.437	2.35×10^{-1}	54.849	44.350
1	5.114	3.650	54.110	46.273	1.69×10^{-1}	54.059	46.136
2	5.333	3.451	53.722	47.849	1.23×10^{-1}	53.665	47.652
3	5.513	3.341	53.516	48.975	9.27×10^{-2}	53.453	48.720
4	5.660	3.280	53.393	49.732	7.36×10^{-2}	53.324	49.431
5	5.782	3.246	53.313	50.238	6.12×10^{-2}	53.239	49.898
10	6.144	3.186	53.154	51.553	3.11×10^{-2}	53.062	50.942
20	6.370	3.154	53.112	52.667	8.45×10^{-3}	53.000	51.485
30	6.422	3.145	53.109	52.992	2.22×10^{-3}	52.992	51.556
40	6.435	3.143	53.109	53.079	5.71×10^{-4}	52.990	51.567
50	6.438	3.143	53.109	53.101	1.47×10^{-4}	52.989	51.569
60	6.439	3.142	53.109	53.107	3.76×10^{-5}	52.989	51.570

Example 2.7 Table 2.6 illustrates the iteration method [initiated using equation (2.7)] on the single-facility location problem with data given in Table 2.5. The bound $LB^{(\ell)}$ and the $S^{(\ell)}$ that were calculated are those defined in equations (2.8) and (2.9); $LBH^{(\ell)}$ and $SH^{(\ell)}$ were calculated by the same expressions, but using $WH(X)$ instead of $W(X)$. We observe that using $\epsilon = 0.01$ speeded up convergence as measured by $SH^{(\ell)}$ or by movement in $(x_1^{(\ell)},x_2^{(\ell)})$, although using $\epsilon = 0$ gave a better value of $W(X)$ at each iteration.

EXERCISES

2.1 Consider the sum $\sum_{i=1}^{n} f_i(x_1,x_2)$, where each $f_i(x_1,x_2)$ is convex. Using the definition of convexity given in the Appendix, mathematical note 2.1, prove that the sum is also convex.

2.2 The Varignon frame is a mechanical analog for the minimization of $W(X)$ in problem (2.2). It consists of a board with holes drilled in it to correspond to fixed facility locations. A string is passed through each hole j, and the ends are tied together in a knot on top of the board. Under the board a weight is attached to each string, and the weight on string j is proportional to w_j in problem (2.2). In the absence of friction and tangled strings, the knot will come to rest at the optimum new facility location. This analog has been used in actual location studies. The analog can be analyzed in terms of forces acting to move the knot. Assume that the knot is in equilibrium. In Figure 2E.1 the weight w_4, for example, acts with a force w_4, which can be broken up into the orthogonal components w_{x4} and w_{y4}, where:

$$w_{x4} = w_4 \cos\theta_4 \text{ and } w_{y4} = w_4 \sin\theta_4.$$

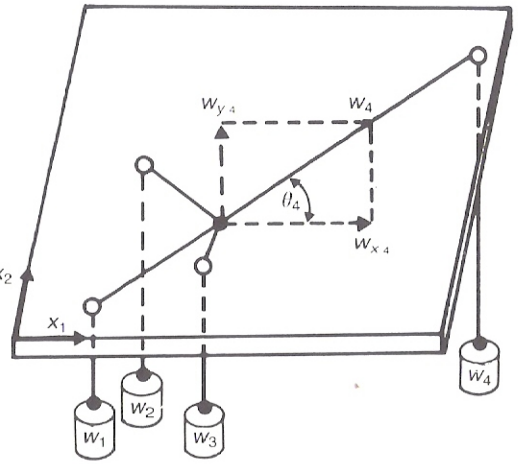


Figure 2E.1 Varignon Frame.

Note that $w_4^2 = w_{x4}^2 + w_{y4}^2$.

a. Show that:

$$\cos \theta_4 = \frac{a_{41} - x_1}{\ell_2(X,a_4)} \text{ and } \sin \theta_4 = \frac{a_{42} - x_2}{\ell_2(X,a_4)},$$

where (x_1,x_2) is the equilibrium position.

b. Show that a balance of x-direction force components and a balance of y-direction force components is, in general, equivalent to conditions (2.3).

c. Explain condition (2.4) in terms of forces. If condition (2.4) holds for point (a_{j1},a_{j2}), what would be observed on the Varignon frame?

2.3 Show that $W(X) = \sum_{j=1}^{n} w_j(\ell_2(X,a_j))^2$ is strictly convex. (Hint: Since the second partial derivatives of W are continuous everywhere, showing strict convexity is equivalent to showing that the Hessian matrix is positive definite everywhere; see Exercise 2.12.)

2.4 Write a computer program to perform the iterations described by procedure (2.6). The program should calculate the lower bound $LB^{(k)}$ and stop when $S^{(k)}$ is below a predetermined value. Use the program to recalculate Table 2.3 for the cases where:

a. $w_2 = 10.0$, using equation (2.7) to define the starting point,
b. $w_2 = 4.9$, using equation (2.7) to define the starting point,
c. $w_2 = 4.9$, using the starting point $X^{(0)} = (0.99,4.01)$.

2.5 a. Show that $X^{(1)}$ in procedure (2.6) will always fall in Ω, the convex hull of of the existing facility locations. (Hint: This is equivalent to showing that $X^{(1)}$ can be written as a weighted sum of the a_j's with positive weights that sum to one.)

b. Show that X^* in problem (2.2) is in Ω. (Hint: If $X^* \neq a_j$, for any j then use conditions (2.3)).

2.6 Apply Property 2.7 to the problem in Example 2.1. Then calculate $(X^{(1)})$ based on $(X^{(0)}) = (0,0)$ using procedure (2.22) with $\epsilon = 0$.

2.7 Using the definition of convexity given in the Appendix, mathematical note 2.1, prove that $W(x) = w|x - a|$ is a convex function for any real positive w and any real a. (Hint: For any real numbers p and q, $|p + q| \leq |p| + |q|$.)

2.8 Derive equation (2.14).

2.9 A facility is to be located on a shop floor where travel is possible only along aisles that are perpendicular to each other. Deliveries will have to be made to two "demand points" located (using rectangular axes corresponding to the directions of the aisles) at (1,1) and (3,2). The weights for the demand points were estimated to be equal (thus taken to be 1 and 1, respectively). Problem (2.10) is to be used as the location model.

a. On a graph showing the locations of the demand points, draw the equal-cost contour where $W(X)$, the total cost, is equal to 5. (Hint: Divide the x_1x_2 plane into regions such that $W(X)$ can be described in each without using absolute values.)
b. Plot the contour $W(X) = 3$.
c. Plot the contour $W(X) = 1$.

2.10 Jack Smooth, an industrial sales and service representative, wants to find a new office location. His business involves several trips a week to seven factories in a large suburban area. From a travel expense diary, he has made up some averages based on several months' experience. He has marked the location of each plant on a map and, using the left and bottom borders of the map as coordinate axes, has given a location to each plant. This information is compiled in Table 2E.1. Jack Smooth's objective is to find an office location that minimizes total travel distance. He considers rectangular

Table 2E.1 Average Trips and Locations of Customers.

Customer	Average Number of Weekly Trips	Location
A	5	(5,20)
B	7	(18,8)
C	2	(22,16)
D	3	(14,17)
E	6	(7,2)
F	1	(5,15)
G	5	(12,4)

distances to be the best representations of the actual travel distances involved.

a. Solve Jack Smooth's office location problem.
b. Show that conditions (2.15) are satisfied at the optimal location.
c. Plot $W_1(x_1)$ against x_1.

2.11 **a.** Graph the function $y = \frac{a}{b}|x - x_0|$ where a and b are positive. Then overlay the graph of the hyperbola whose equation is:

$$\frac{y^2}{a^2} - \frac{(x - x_0)^2}{b^2} = 1.$$

b. Relate a to the vertical distance between the graphs at $x = x_0$.
c. Discard the lower branch of the hyperbola. Solve for y and comment on the result in relation to its use in problem (2.19) in the text.

2.12 Prove that $WH(X)$ in problem (2.19) is strictly convex. Hint: First prove that $g(x_1,x_2) = (((x_1 - a_{j1})^2 + \epsilon)^{p/2} + ((x_2 - a_{j2})^2 + \epsilon)^{p/2})^{1/p}$ is strictly convex by showing that for all $X = (x_1,x_2)$:

$$\frac{\partial^2 g(X)}{\partial x_1^2} > 0$$

and

$$\begin{vmatrix} \frac{\partial^2 g(X)}{\partial x_1^2} & \frac{\partial^2 g(X)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 g(X)}{\partial x_2 \partial x_1} & \frac{\partial^2 g(X)}{\partial x_2^2} \end{vmatrix} > 0.$$

2.13 Derive iterative procedure (2.22) from the extremal conditions (partial derivatives set equal to zero) for $WH(X)$.

2.14 Consider the problem data Table 2E.1 in Exercise 2.10.

a. Check each plant location for candidacy as the optimal office location, assuming ℓ_p distances first with $p = 1$, and then with $p = 1.5$.
b. Use any available computer package for nonlinear optimization to find the optimum location for ℓ_p distances with $p = 1.1$, 1.5, 2, and 5.
c. Write a computer program to perform the Weiszfeld iteration procedure (2.22), and repeat Exercise 2.14b.
d. Write the additional coding for the program written in Exercise 2.14c to begin by testing for optimality at the existing facility locations. Repeat Exercise 2.14b.

2.15 **a.** Prove that:

$$W_L(X_R^*) \geqslant W_L(X_L^*) \geqslant ((W_1(x_{1R}^*))^p + (W_2(x_{2R}^*))^p)^{1/p},$$

where:

$X_R^* = (x_{1R}^*, x_{2R}^*)$ is an optimum location for the rectangular distance problem with $W_k(x_k)$ as defined in problem (2.12); and $X_L^* = (x_{1L}^*, x_{2L}^*)$ is an optimum location for the ℓ_p distance problem with $W_L(X)$ representing $W(X)$ in problem (2.17). (Hint: Consider the Minkowski inequality stated in the Appendix, mathematical note 2.3.)

b. Apply the above result to Example 2.6 with $p = 1.3$.

REFERENCE NOTES

SECTION 2.1 The Weiszfeld procedure appeared in a paper by "Weiszfeld" (1937). It should be noted that it is one of the methods for iteratively solving nonlinear equations that appear in numerical analysis texts (eg, Dahlquist and Björck (1974), Chapter 6). Convergence properties have been discussed by, among others, Katz (1969, 1974), Kuhn (1973), and Ostresh (1978). The stopping criterion presented in this section is one of those given by Love and Yeong (1981). A somewhat tighter lower bound was given by Elzinga and Hearn (1983) and Juel (1984); Love and Dowling (1986) generalized Drezner's rectangular bound to the ℓ_p distance case.

SECTION 2.2 Some of the foundation work with rectangular distances includes that by Bindschedler and Moore (1961) and Francis (1963). The single-facility problem is closely related to the problem of finding a weighted median of a data set, as was seen in Example 2.5. Algorithms that can solve such problems in 0(n) time are available.

SECTION 2.3 The hyperbolic approximation used here is from Wesolowsky and Love (1972). Eyster, White, and Wierwille (1973) used a hyperboloid approximation procedure (HAP) to extend the original Weiszfeld procedure to both Euclidean and rectangular distances. Convergence properties including the extension to ℓ_p distances have been discussed by Morris and Verdini (1979) and Morris (1981). An acceleration of the HAP procedure is discussed by Charalambous (1985).

APPENDIX—CHAPTER 2

Mathematical Notes

2.1 *Prove* Property 2.1
$w_j((x_1 - a_{j1})^2 + (x_2 - a_{j2})^2)^{1/2}$ is convex.

PROOF: As w_j is a positive constant we can, without loss of generality, let $w_j = 1$. As a_{j1} and a_{j2} are constants, we can equivalently prove that $f(y_1,y_2) = (y_1^2 + y_2^2)^{1/2}$ is convex. This merely changes the coordinate system to one with an origin at (a_{j1},a_{j2}). Recall that $f(y_1,y_2)$ is said to be *convex* if, given any two points, (y_1',y_2') and (y_1'',y_2''),

$$f(\lambda(y_1',y_2') + (1-\lambda)(y_1'',y_2'')) \leq \lambda f(y_1',y_2') + (1-\lambda)f(y_1'',y_2'')$$

where $0 \leq \lambda \leq 1$.

Therefore, the convexity requirement is:

$$((\lambda y_1' + (1-\lambda)y_1'')^2 + (\lambda y_2' + (1-\lambda)y_2'')^2)^{1/2}$$
$$\leq \lambda((y_1')^2 + (y_2')^2)^{1/2} + (1-\lambda)((y_1'')^2 + (y_2'')^2)^{1/2}.$$

To show that this must be true, we turn to what is known as the triangle inequality for vectors. For two real-valued vectors $p = (p_1,p_2)$ and $q = (q_1,q_2)$,

$$((p_1 + q_1)^2 + (p_2 + q_2)^2)^{1/2} \leq (p_1^2 + p_2^2)^{1/2} + (q_1^2 + q_2^2)^{1/2}.$$

If we set $p_1 = \lambda y_1'$, $p_2 = \lambda y_2'$, $q_1 = (1-\lambda)y_1''$, and $q_2 = (1-\lambda)y_2''$,

we see that this is equivalent to our requirement.

2.2 *Prove* Property 2.2
The minimum of:

$$W(X) = \sum_{j=1}^{n} w_j((x_1 - a_{j1})^2 + (x_2 - a_{j2})^2)^{1/2}$$

occurs at a_r, if, and only if, $CR_r \leq w_r$.

PROOF: Consider a movement of the new facility from (x_1,x_2) a distance t to $(x_1 + td_1, x_2 + td_2)$, where $(d_1^2 + d_2^2)^{1/2} = 1$. Here d_1 and d_2 are components of a unit direction vector d. Let us find the rate of change of $W(X+td)$ with respect to t as t approaches zero when $X = a_r$:

$$\frac{dW}{dt} = \sum_{j=1}^{n} \frac{w_j((a_{r1} + td_1 - a_{j1})d_1 + (a_{r2} + td_2 - a_{j2})d_2}{((a_{r1} + td_1 - a_{j1})^2 + (a_{r2} + td_2 - a_{j2})^2)^{1/2}}$$

$$= \frac{w_r(td_1^2 + td_2^2)}{((td_1)^2 + (td_2)^2)^{1/2}}$$

$$+ d_1 \sum_{\substack{j=1\\ \neq r}}^{n} \frac{w_j(a_{r1} + td_1 - a_{j1})}{((a_{r1} + td_1 - a_{j1})^2 + (a_{r2} + td_2 - a_{j2})^2)^{1/2}}$$

$$+ d_2 \sum_{\substack{j=1\\ \neq r}}^{n} \frac{w_j(a_{r2} + td_2 - a_{j2})}{((a_{r1} + td_1 - a_{j1})^2 + (a_{r2} + td_2 - a_{j2})^2)^{1/2}}.$$

We can write this derivative as:

$$\frac{dW}{dt} = w_r(d_1^2 + d_2^2)^{1/2} + d_1R_1(t) + d_2R_2(t),$$

where $R_1(t)$ and $R_2(t)$ are defined as implied above.
Therefore,

$$\frac{dW}{dt}\Big|_{t\to 0} = w_r(d_1^2 + d_2^2)^{1/2} + d_1R_1 + d_2R_2$$

$$= w_r + d_1R_1 + d_2R_2,$$

$$\text{where } R_1 = \sum_{\substack{j=1\\ \neq r}}^{n} \frac{w_j(a_{r1} - a_{j1})}{\ell_2(a_r,a_j)}$$

$$\text{and } R_2 = \sum_{\substack{j=1\\ \neq r}}^{n} \frac{w_j(a_{r2} - a_{j2})}{\ell_2(a_r,a_j)}.$$

We can use elementary calculus *and* the condition $d_1^2 + d_2^2 = 1$ to find that the minimum of:

$$\frac{dW}{dt}\Big|_{t\to 0}$$

occurs at:

$$d_1 = -\frac{R_1}{(R_1^2 + R_2^2)^{1/2}}, \quad d_2 = -\frac{R_2}{(R_1^2 + R_2^2)^{1/2}}$$

$$\text{and so } \min \frac{dW}{dt}\Big|_{t\to 0} = w_r - (R_1^2 + R_2^2)^{1/2}.$$

When this derivative is positive, $W(X)$ will increase as (x_1,x_2) is moved in *any* direction from (a_{r1},a_{r2}). Since $W(X)$ is convex, $X^* = (a_{r1},a_{r2})$ if, and only if, $w_r \geq (R_1^2 + R_2^2)^{1/2} = CR_r$, as required.

2.3 *Prove* Property 2.6
$w_j\ell_p(X,a_j) = w_j(|x_1 - a_{j1}|^p + |x_2 - a_{j2}|^p)^{1/p}$ is convex.

PROOF: Following note 2.1 we can equivalently prove that $f(y_1,y_2) = (|y_1|^p + |y_2|^p)^{1/p}$ is convex. We will appeal to a well-known (to mathematicians, at least) inequality called the Minkowski inequality, given by:

$$\left(\sum_{k=1}^{K} |\alpha_k + \beta_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^{K} |\alpha_k|^p\right)^{1/p} + \left(\sum_{k=1}^{K} |\beta_k|^p\right)^{1/p}$$

where $p \geq 1$ and α_k and β_k are real numbers.

We have $f(\lambda y_1' + (1 - \lambda)y_1'', \lambda y_2' + (1 - \lambda)y_2'')$

$$\leq ((|\lambda y_1'| + |(1 - \lambda)y_1''|)^p + (|\lambda y_2'| + |(1 - \lambda)y_2''|)^p)^{1/p} \quad \text{(triangle inequality)}$$

$$\leq (|\lambda y_1'|^p + |\lambda y_2'|^p)^{1/p} + (|(1 - \lambda)y_1''|^p + |(1 - \lambda)y_2''|^p)^{1/p} \quad \text{(Minkowski inequality)}$$

$$= \lambda(|y_1'|^p + |y_2'|^p)^{1/p} + (1 - \lambda)(|y_1''|^p + |y_2''|^p)^{1/p}$$

$$= \lambda f(y_1', y_2') + (1 - \lambda)f(y_1'', y_2''), \text{ as required.}$$

2.4 *Prove* $L_p(X,a_j) - \ell_p(X,a_j) \leq 2^{1/p}\epsilon^{1/2}$, where

$$L_p(X,a_j) = (((x_1 - a_{j1})^2 + \epsilon)^{p/2} + ((x_2 - a_{j2})^2 + \epsilon)^{p/2})^{1/p}.$$

PROOF: Let $y_1 = |x_1 - a_{j1}| \geq 0$, $y_2 = |x_2 - a_{j2}| \geq 0$ and $y_3 = \epsilon^{1/2} > 0$. Then:

$$L_p(X,a_j) = [(y_1^2 + y_3^2)^{p/2} + (y_2^2 + y_3^2)^{p/2}]^{1/p}$$

$$\leq [((y_1 + y_3)^2)^{p/2} + ((y_2 + y_3)^2)^{p/2}]^{1/p}$$

$$= [|y_1 + y_3|^p + |y_2 + y_3|^p]^{1/p}$$

$$\leq [|y_1|^p + |y_2|^p]^{1/p} + [|y_3|^p + |y_3|^p]^{1/p}$$

(Minkowski inequality; see note 2.3)

$$= [|x_1 - a_{j1}|^p + |x_2 - a_{j2}|^p]^{1/p} + [\epsilon^{p/2} + \epsilon^{p/2}]^{1/p}$$

$$= \ell_p(X,a_j) + 2^{1/p}\epsilon^{1/2}, \text{ as required.}$$

2.5 *Prove* [Rectangular bound on $W(X^*)$]. $W(X^*) \geq R(X^{(\ell)})$

$$= \min_{x_1} \sum_{j=1}^{n} w_j' |x_1 - a_{j1}| + \min_{x_2} \sum_{j=1}^{n} w_j'' |x_2 - a_{j2}|,$$

where $W(X^*)$ is defined in problem (2.2) in the text while

$$w_j' = w_j|x_1^{(\ell)} - a_{j1}|/\ell_2(X^{(\ell)},a_j), \text{ and}$$

$$w_j'' = w_j|x_2^{(\ell)} - a_{j2}|/\ell_2(X^{(\ell)},a_j).$$

PROOF: We may write the following inequality:

$$|x_1 - a_{j1}|\,|x_1^{(\ell)} - a_{j1}| + |x_2 - a_{j2}|\,|x_2^{(\ell)} - a_{j2}|$$

$$\leq [|x_1 - a_{j1}|^2 + |x_2 - a_{j2}|^2]^{1/2}[|x_1^{(\ell)} - a_{j1}|^2 + |x_2^{(\ell)} - a_{j2}|^2]^{1/2}$$

(since $|u \cdot v| \leq \|u\|\,\|v\|$; Schwartz inequality)

$$= \ell_2(X,a_j)\,\ell_2(X^{(\ell)},a_j),$$

or equivalently,

$$\ell_2(X,a_j) \geq [|x_1^{(\ell)} - a_{j1}|/\ell_2(X^{(\ell)},a_j)]|x_1 - a_{j1}| + [|x_2^{(\ell)} - a_{j2}|/\ell_2(X^{(\ell)},a_j)]|x_2 - a_{j2}|.$$

Multiplying both sides by w_j, summing over j and then taking the min on both sides with respect to X yields the required result.

(Note: The bound $R(X^{(\ell)})$ is due to Drezner (1984). The solution $X^{(\ell)}$ at each iteration of the Weiszfeld procedure is used to compute the weights w_j' and w_j'' that are used in calculating $R(X^{(\ell)})$. While it may appear that adding another optimization problem and solving it at each iteration has increased the work required to find a lower bound, this approach has several advantages. Using techniques from Section 2.2, each part of the optimization involved in calculating $R(X^{(\ell)})$ can be accomplished rapidly. Also, it is not necessary to find the hull points that are used in Property 2.4.)