Environment and Planning B: Planning and Design, 1989, volume 16, pages 41-50

An algorithm for facility location in a districted region

C ReVelle
Department of Geography and Environmental Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA
D J Elzinga
Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA
Received 15 October 1988

Abstract. The problem of facility siting in a districted region is discussed and a two-stage algorithm proposed. The first stage consists of solving a particular facility siting problem (for example, a p-median problem, or a maximal location covering problem) in each district for all possible allocations of facilities to a district with respect to the number of facilities allocated in each district. The second stage is the allocation of resources (facilities) among the districts. A greedy algorithm is proposed to solve the resource allocation problem with a criterion of minimizing the sum of weighted distances under convex conditions.

Introduction

The problem addressed in this paper is that of optimal allocation of a fixed number of facilities to the districts of a region. This problem would arise in the allocation of schools in a county, in voting-machine siting among voting districts, or large-scale facility-siting studies for regions encompassing many counties or states. A district, for the purposes of this paper, is defined to be a subregion of demand points all of which can be served only by supply points in the same subregion.

The p-median problem (Hakimi, 1965) and the maximal covering location problem (MCLP) (Church and ReVelle, 1974) are two examples of general network location problems for which the number of facilities to be allocated is fixed. As Church and ReVelle (1976) have shown that the MCLP is a special case of the p-median problem, we shall treat the p-median problem and make comments when the additional structure of the MCLP permits simplification of results. The p-median problem is designed to minimize the average distance which users must travel by the optimal siting of a fixed number of facilities. The MCLP is designed to maximize the population covered within a stated distance, again by the allocation of a fixed number of facilities.

The p-median problem, first formulated by Hakimi, has been treated by many researchers since 1965. The exact solution to the p-median problem may be found by numerous techniques, for example, by solving the integer linear program as formulated by ReVelle and Swain (1970) or by the branch-and-bound algorithms due to Jarvinen et al (1972) and to El-Shaieb (1973). In 1977, Narula et al proposed a subgradient optimization procedure for generating solutions to the p-median problem. Rosing et al (1979) demonstrated how row-and-column compression could extend the size of p-median problems solved by linear programming.

Because of the combinatoric nature of the p-median problem it has only been solved exactly for p slightly more than 100. This restriction in the size of the problem which may be handled exactly has led to the development of a number of heuristics for the solution of the p-median problem, including that of Maranzana (1964) and Teitz and Bart (1968).

The assumption of a districted region implies a restriction in the definition of when a supply point is eligible to serve a demand point. As previously stated, a facility site within one of the subregions is only eligible to serve or cover a demand point within that subregion, and, conversely, a demand point in subregion A cannot be served by a facility in subregion B. This limitation on eligible servers holds even if the closest facility to a given demand point is not in the subregion of the demand point. This assumption lends more structure to the problem; the added structure suggests that there may be a faster procedure for exactly solving districted location problems than simply treating them with techniques for solving the undistricted problem.

The problem has two aspects. First, we are to decide the number of facilities to be allocated to each region. Second, we must select the sites at which those facilities are to be located.

Algorithm for facility location in districted regions

Let us assume we have an area A divided into K districts, with s_k eligible sites for facilities in each district k. We are given p facilities to site in A.

To solve the p-median problem in the region, we will find it necessary to solve individually in each district k a number of smaller p-median problems; these solutions can then be integrated to develop the regional solution. The number of facilities to be sited in a particular district is not known in advance. Instead, there may be any number of facilities between 0 and p the number of potential sites in the district. Let the upper limit in region k be p_k. When the p-median problem is solved in district k, a value for the weighted distance (people-miles) achieved by some number of facilities in the region, say x_k, is obtained. Call this value of the weighted distance $\mathrm{f}_k(x_k)$. Solving the associated p-median problem for $x_k = 1$ to $x_k = p_k$ generates a table that shows how the weighted distance declines as the number of facilities in the region increases.

Now, by the definition of weighted distance, $\mathrm{f}_k(x_k)$ must be a monotonically decreasing function of x_k. It is also reasonable that in nearly all cases $\mathrm{f}_k(x_k)$ will be convex in x_k. Economically, the convexity is equivalent to the common behavior of diminishing marginal returns on facility investment. In theory this need not always occur (see the appendix).

Now that we have k monotonically increasing functions, $\mathrm{f}_k(x_k)$, we shall allocate the total number of facilities, R, to minimize total people-miles for the entire population in all regions. Since there is no assignment across districts, we begin with at least one facility in each district. An integer programming formulation of this problem is:

$$\text{minimize} \sum_{k=1}^{K} \mathrm{f}_k(x_k)\,,$$

subject to

$$\sum_{k=1}^{K} x_k = R\,.$$

It is clear that such a problem can be solved by dynamic programming. One can also see this as the following zero-one program:

Let

$$x_{i,k} = \begin{cases} 1\,, & \text{if } i \text{ facilities are allocated to region } k. \\ 0\,, & \text{otherwise,} \end{cases}$$

and let $c_{i,k}$ be the weighted distance for demands in region k when i facilities are allocated to region k;

$$\text{minimize } Z = \sum_{k=1}^{K} \sum_{i=1}^{p_k} c_{i,k} x_{i,k} ,$$

subject to

$$\sum_{i=1}^{p_k} x_{i,k} = 1, \qquad k = 1, 2, \dots, K ,$$

$$\sum_{k=1}^{K} \sum_{i=1}^{p_k} i x_{i,k} = R .$$

The first K constraints state that only one allocation may be made in each district. The last constraint states that only R facilities may be allocated among the K districts.

One method of solution to this problem is by the method of Lagrangian relaxation suggested by Everett (1963). Bringing the last constraint to the objective with a multiplier λ, we have

$$\text{minimize } Z_L = \sum_{k=1}^{K} \sum_{i=1}^{p_k} c_{i,k} x_{i,k} + \lambda \sum_{k=1}^{K} \sum_{i=1}^{p_k} i x_{i,k} = \sum_{k=1}^{K} \sum_{i=1}^{p_k} (c_{i,k} + \lambda i) x_{i,k} ,$$

subject to

$$\sum_{i=1}^{p_k} x_{i,k} = 1, \qquad k = 1, 2, \dots, k .$$

For any specified value of λ, the solution to this Lagrangian relaxation is obvious. Recall that only one $x_{i,k}$ from the sum of the $x_{i,k}$ over i may be one. From the sum

$$\sum_{i=1}^{p_k} x_{i,k}, \qquad \text{for a given } k ,$$

we would choose that element $x_{i,k}$ with the smallest value of $(c_{i,k} + \lambda i)$ to have a value of one. Once this choice has been made, the number of facilities in each region is immediately specified, as is the weighted distance in that region.

If the number of facilities in region k is known, the total number of facilities in all regions is obtained by merely summing these numbers. The total weighted distance is obtained in the same fashion.

It follows that each specification of λ leads to a total number of facilities, allocated by region, and also to a total weighted distance, composed of regional weighted distances. A change in the value of λ will only lead to a change in the total weighted distance and total number of facilities if for some region k the minimum of $(c_{i,k} + \lambda i)$ shifts to some other i. Ranging λ from small to large in sufficiently small steps will lead to a range of solutions.

For small values of λ, there will be many facilities dispersed among the regions. For larger values of λ, there will be fewer total facilities allocated to the regions. Obviously there is only minimal computational effort needed to allocate the facilities among the regions, once the individual regional problems have been solved.

Another way to view the problem is perhaps more understandable and instructive and indeed is likely to be even easier to implement. We define $\Delta c_{i,k}$ as the decrease in weighted distance in going from i facilities to $i+1$ facilities in region k,

$$\Delta c_{i,k} = c_{i,k} - c_{i+1,k} .$$

The definition is appropriate except for the placement of the first facility in region k. We note also that

$$\Delta c_{1,k} > \Delta c_{2,k} > \cdots > \Delta c_{i,k} > \Delta c_{i+1,k} \cdots .$$

Suppose we are dealing with a three-region problem and λ is such that

$$\text{for region 1: } \min_i (c_{i,1} + \lambda i) = c_{3,1} + 3\lambda ,$$

$$\text{for region 2: } \min_i (c_{i,2} + \lambda i) = c_{2,2} + 2\lambda ,$$

$$\text{for region 3: } \min_i (c_{i,3} + \lambda i) = c_{4,3} + 4\lambda .$$

For one of the regions to be allocated an additional facility, λ will have to decrease until the minimum value of $(c_{i,k} + \lambda i)$ shifts for one of the k regions. For that value of λ in the region k where the minimum first shifts,

$$c_{i,k} + \lambda i > c_{i+1,k} + \lambda(i+1) ,$$

$$c_{i,k} - c_{i+1,k} > \lambda(i+1) - \lambda i ,$$

$$\Delta c_{i,k} > \lambda .$$

That is, as soon as λ has achieved a value less than $\Delta c_{i,k}$, there will be at least $i+1$ facilities in region k. Thus, as λ decreases, successively smaller values of $\Delta c_{i,k}$ become larger than λ, and more facilities enter.

It follows that we need not actually utilize λ in solving the problem. It is sufficient to note that, once there is one facility in each region, the *next* facility to enter occurs in that region with the largest value of $\Delta c_{i,k}$. The facility which enters next occurs in the region with the next largest value of $\Delta c_{i,k}$. Facilities thus enter in order of decreasing values of $\Delta c_{i,k}$ until the limit in the number of facilities is reached. Under conditions of convexity, then, this problem has a greedy algorithm solution, namely finding the R largest values of $\Delta c_{i,k}$.

The problem may be thought of in another way. When the problem is viewed in its dynamic programming form, namely,

$$\text{minimize} \sum_{k=1}^{K} \mathrm{f}_k(x_k) ,$$

subject to

$$\sum_{k=1}^{K} x_k = R ,$$

we merely note the convexity of the objective function to arrive at an algorithm. It is clear that the first facility to enter, given that each region already has been allocated one facility, is the one with the greatest payoff or greatest reduction in people-miles, $\Delta c_{i,k}$. The second to enter is the one with the next largest reduction, and so on. Facilities which enter subsequent to the initial allocation of one per region may enter all in the same region or in different regions depending on the relative values of the $\Delta c_{i,k}$.

This procedure suggests an economic interpretation of the greedy algorithm. The $\Delta c_{i,k}$ may be thought of as the measure of the marginal productivity of an additional facility in district k. We would expect from basic economic theory that the most efficient solution occurs when the marginal productivities in each of the districts are equal. Although it may not be possible in the discrete case for all of the $c_{i,k}$ to be equal, a close equivalent is implied by the algorithm.

Discussion and extension of the algorithm to nonconvex *p*-median problems

Let us give an example of the greedy algorithm. Assume a region is partitioned into three districts and there are eight facilities to be sited in the region, with each district to be covered by at least one facility. There are equal populations in each district. We have determined the $c_{i,k}$ functions by solving the relevant p-median problems in the districts and have obtained

	i	1	2	3	4	5
district 1	$c_{i,k}$	14	11	9	8	7
district 2	$c_{i,k}$	13	9	7	6	5.5
district 3	$c_{i,k}$	16	12	9	7	6

The optimal set of $\Delta c_{i,k}$ is (4, 4, 3, 3, 2) which corresponds to the order of entering the five remaining $\Delta c_{i,k}$ after the first three facilities have been placed. This corresponds to an optimal allocation of

3 in district 1, 2 in district 2, 3 in district 3,

or

2 in district 1, 3 in district 2, 3 in district 3,

or

2 in district 1, 2 in district 2, 4 in district 3.

In the case of districted MCLPs, it is certainly true that the greedy algorithm holds. The assumption that at least one facility is sited in each district need not be made for MCLPs.

In actual practice for large districted problems, the $c_{i,k}$ functions need not be generated over the entire interval $i = 1, \ldots, p$. One procedure might use a good heuristic such as that derived by Teitz and Bart (1968) for the p-median problem to obtain an initial allocation among districts. (In lieu of this, the initial allocation of facilities might be proportional to the population in a district.) Next, the $c_{i,k}$ functions would be generated only for values around the initial allocation and tests for optimality would be made. Further generation of the $c_{i,k}$ functions would be based upon movement toward optimality; optimality determined by analysis of the $\Delta c_{i,k}$ values around the allocation. If a small $\Delta c_{i,k}$ is included in the list for summation, whereas a larger $\Delta c_{i,k}$ is rejected, for a particular allocation, then the allocation is suboptimal and a new allocation excluding the small $\Delta c_{i,k}$ and including the larger $\Delta c_{i,k}$ would be created and analyzed.

As an aside, it might be possible in certain special cases to decompose large p-median problems into several smaller ones by pseudodistricting the large problem and using explicit enumeration techniques where districting structure is violated. Unfortunately, the computational burden increases rapidly with total number of violations.

Although in most all practical cases $f_k(x_k)$ will probably be convex, there might arise special cases (as the examples given in the appendix) where $f_k(x_k)$ may be mildly nonconvex for some districts. Under mildly nonconvex cases the algorithm can be modified. Let us say in district k', $\Delta c_{i+1,k'} \geqslant \Delta c_{i,k'}$, but $\Delta c_{i'+2,k'} < \Delta c_{i',k'}$, for a given i. Thus $c_{i,k}$ is nonconvex on the interval $[i', i'+2]$. If the greedy algorithm chooses both $\Delta c_{i',k'}$ and $\Delta c_{i'+1,k'}$, or excludes both, then the algorithm has generated the optimal allocation. If the greedy algorithm chooses $\Delta c_{i'+1,k'}$ and excludes

$\Delta c_{i',k'}$, then there is difficulty. The following procedure will find the optimum:
Step 1 Generate all $\Delta c_{i,k}$ for all $j = 1, .., p$.
Step 2 Follow the greedy algorithm until $\Delta c_{i'+1,k'}$ would be chosen.
Step 3 Choose next largest $\Delta c_{i,k}$ from all $\Delta c_{i,k}$ *except* from all $\Delta c_{i',k'}$; call this Δc_{ik}^1.
Step 4 Choose next largest $\Delta c_{i,k}$ from all $\Delta c_{i,k}$ *except* from all $\Delta c_{i',k'}$; call this Δc_{ik}^2.
Step 5 If $\Delta c_{i,k}^1 + \Delta c_{i,k}^2 > \Delta c_{i,k'} + \Delta c_{i'+1,k'}$, then choose $\Delta c_{i,k}^1$, $\Delta c_{i,k}^2$. Otherwise choose $\Delta c_{i',k'}$, $\Delta c_{i'+1,k'}$ and resume greedy algorithm.
Step 6 Choose next largest $\Delta c_{i,k}$ from all but $\Delta c_{i,k'}$; call this $\Delta c_{i,k}^3$.
Step 7 If $\Delta c_{i,k}^2 + \Delta c_{i,k}^3 > \Delta c_{i',k'} + \Delta c_{i'+1,k'}$, then choose $\Delta c_{i,k}^2$, $\Delta c_{i,k}^3$. Otherwise choose $\Delta c_{i',k'}$, $\Delta c_{i'+1,k'}$ and resume greedy algorithm.
Step 8 Continue choosing $\Delta c_{i,k}^r$ from all but $\Delta c_{i',k'}$, comparing whether $\Delta c_{i,k}^{r-1} + \Delta c_{i,k}^r > \Delta c_{i',k'} + \Delta c_{i'+1,k'}$, in which case choose $\Delta c_{i,k}^{r-1}$, $\Delta c_{i,k}^r$, otherwise choose $\Delta c_{i',k'}$, $\Delta c_{i'+1,k'}$, and resume greedy algorithm. Continue until greedy algorithm is resumed or p values of $\Delta c_{i,k}$ have been chosen.

This procedure may also easily be extended to any number of nonconvexities on any length interval. However, the computational burden increases quickly. The above process is 'almost' greedy for a single nonconvexity.

In conclusion, this proposed methodology can handle much larger problems than heretofore have been treated. If a subregion has fifty nodes, a very large number of subregions may be handled by the algorithm. Although the $f_k(x_k)$ may not be convex in all cases, mild nonconvexities may be handled with little additional difficulty. Experience has shown, however, that the conditions necessary for convexity hold true in most practical examples, especially as p increases.

References

Church R, ReVelle C, 1974, "The maximal covering location problem" *Papers of Regional Science Association* **32** 101–118

Church R, ReVelle C, 1976, "Theoretical and computational links of the *p*-median and maximal covering location problem" *Geographical Analysis* **8** 406–421

El-Shaieb A M, 1973, "A new algorithm for locating sources among destinations" *Management Science* **20** 221–231

Everett H, 1963, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources" *Operations Research* **11** 399–417

Hakimi S L, 1965, "Optimal location of switching centers and the absolute centers and medians of a graph" *Operations Research* **12** 450–459

Jarvinen P, Rajak J, Sinervo H, 1972, "A branch-and-bound algorithm for seeking the *p*-median" *Operations Research* **20** 173–178

Maranzana F, 1964, "On the location of supply points to minimize costs" *Operations Research Quarterly* **15** 261–273

Narula S, Ogbu U, Samuelsson H, 1977, "An algorithm for the *p*-median problem" *Operations Research* **28** 709–711

ReVelle C, Swain R, 1970, "Central facilities location" *Geographical Analysis* **2** 30–42

Rosing K, ReVelle C, Rosing-Vogellar H, 1979, "The *p*-median and its linear programming relaxation: an approach to large problems" *Journal of the Operational Research Society* **30** 815–823

Teitz M B, Bart P, 1968, "Heuristic methods for estimating the generalized vertex median of a weighted graph" *Operations Research* **16** 955–961

APPENDIX

Conditions for convexity in the *p*-median problem

An example of nonconvexity in the *p*-median problem is the following $[a_{ij} w_i]$ matrix where a_{ij} is the distance travelled from supply point i to a facility at j, and w_i is a set of population weights assigned to the supply points:

i \ j	1	2	3	4
1	1	1	2	2
2	1	2	1	2
3	1	2	2	1
4	1	1	2	2
5	1	2	1	2
6	1	2	3	1
7	1	1	2	2
8	1	2	1	2
9	1	2	2	1
10	2	1	2	2
11	2	2	1	2
12	2	3	2	1

By enumeration, we can create *p*-median solutions for one, two, or three facilities (that is, $p = 1, 2, 3$). When $p = 1$, supply point 1 should be chosen with total weighted distance 15. When $p = 2$, supply points 1 and 4 are chosen (or other alternate optima) with total weighted distance 14. However, when $p = 3$, supply points 2, 3, and 4 are chosen with total weighted distance 12. Thus the function $f_k(x_k)$ in this case is not convex, in that $f_k(1) - f_k(2) = 1 < 2 = f_k(2) - f_k(3)$.

The equivalent b_{ij} matrix for the MCLP where $b_{ij} = 1$ if cell j is within k_0 units of supply point i, and 0 otherwise is

i \ j	1	2	3	4
1	1	1	0	0
2	1	0	1	0
3	1	0	0	1
4	1	1	0	0
5	1	0	1	0
6	1	0	0	1
7	1	1	0	0
8	1	0	1	0
9	1	0	0	1
10	0	1	0	0
11	0	0	1	0
12	0	0	0	1

which implies $f_k(x_k)$ is nonconcave in x_k.

Nonconcavity of $f_k(x_k)$ in the MCLP is equivalent to nonconvexity of $f_k(x_k)$ in the *p*-median problem in that the greedy algorithm to be proposed must be modified if nonconvexity occurs in an actual case. It is thought that occurrence of these nonconvex cases will be fairly unusual.

One support for the peculiarity of nonconvex cases results from the observation that, for most practical *p*-median problems, there is often little change in the chosen sites in the optimal solution. If the chosen sites from the p-median problem with x_k facilities is a subset of the chosen sites by the same p-median problem with x_k+1 facilities, $a < x_k < b$, then $\mathrm{f}_k(x_k)$ will be convex over the interval $[a, b+1]$. This results from the fact that the marginal decrease in the coverage function f_k, from opening a new site n given a subset A of sites already open, will not itself decrease as more elements are added to A. In fact, it is true that if the chosen sites from the p-median problem with x_k facilities is a subset of the chosen sites by the same p-median problem with x_k+2 facilities, then the coverage function $\mathrm{f}_k(x_k)$ will be convex over the interval $[x_k, x_k+2]$. This is shown shortly. In other words, if the solution of a p-median problem is 'stable' in some sense with respect to p, then $\mathrm{f}_k(x_k)$ is convex in x_k. As noted before, this 'stability' is common in practical examples.

We can show that a curve of people-miles versus number of facilities is convex over a region of three values of number of facilities if the optimal set of p facilities is included in the optimal set of $p+2$ facilities.

Define J_p is the set of p chosen sites in the solution of the p-median problem with p facilities,
N is the number of all possible sites in the district,
$C\{A\}$ is the average travel time associated with set of chosen sites A, thus $\mathrm{f}_k(x_k) = C\{J_t\}$.

Theorem *If* $J_p \subset J_{p+2}$, *then* $\mathrm{f}_k(x_k)$ *is convex for* $x_k \in [p, p+2]$.

Proof Convexity implies

$$\mathrm{f}_k(x_k)-\mathrm{f}_k(x_k+1) \geqslant \mathrm{f}_k(x_k+1)-\mathrm{f}_k(x_k+2) \, ,$$

or equivalently

$$C\{J_t\}-C\{J_{t+1}\} \geqslant C\{J_{t+1}\}-C\{J_{t+2}\} \, .$$

Thus

$$2(C\{J_t\}-C\{J_{t+1}\}) \geqslant C\{J_t\}-C\{J_{t+2}\} \, . \tag{A1}$$

We shall show that equation (A1) holds for $J_t \subset J_{t+2}$. Choose n_a such that

$$C\{J_t U n_a\} = \min_{n_i \in N} C\{J_t U n_i\} \, . \tag{A2}$$

Thus

$$C\{J_t\}-C\{J_t U n_a\} \geqslant C\{J_t\}-C\{J_t U n_i\}, \qquad \forall\, n_i \in N \, . \tag{A3}$$

Lemma

$$C\{A U n_j\}-C\{A U n_j U n_i\} \leqslant C\{A\}-C\{A U n_i\}, \qquad \forall\, n_i,\, n_j \in N. \tag{A4}$$

This states that the marginal productivity of siting at n_i is smaller if more facilities have already been sited. Intuitively, this is obviously true.

Proof Formally equation (A4) is true by the following,

$$C\{A\} = \sum_{g=1}^{T} \min_h (\mathbf{w})_{gh}, \qquad (\mathbf{w})_{gh} \in \boldsymbol{D}_h, \qquad h = 1, \ldots, S \, ,$$

where $\mathbf{w}_{gh}$ is the weighted distance matrix for the p-median problem in A, with T demand points and S supply points. Variable h subscripts the supply points, g subscripts the demand points. $\boldsymbol{D}_h$ is the vector of weighted distances to supply point h from all demand points. This is the mathematical formulation of the p-median problem.

Similarly,

$$C\{AUn_i\} = \sum_{g=1}^{T} \min_{h'}(\mathbf{w})_{gh'}, \qquad (\mathbf{w})_{gh'} \in \boldsymbol{D}_{h'}, \quad h' = 1, \dots, S+1\,,$$

$$C\{AUn_j\} = \sum_{g=1}^{T} \min_{h''}(\mathbf{w})_{gh''}, \qquad (\mathbf{w})_{gh''} \in \boldsymbol{D}_{h''}, \quad h'' = 1, \dots, S+1\,,$$

$$C\{AUn_i\,Un_j\} = \sum_{g=1}^{T} \min_{h'''}(\mathbf{w})_{gh'''}, \qquad (\mathbf{w})_{gh'''} \in \boldsymbol{D}_{h'''}, \quad h''' = 1, \dots, S+2\,.$$

Now we will define two $\boldsymbol{H}$ vectors which will represent the difference in individual weighted distances for each demand point under different supply-point configurations,

$$\boldsymbol{H} = (h_1, h_2, \dots, h_T)\,,$$

where

$$h_g = \min_{h}(\mathbf{w})_{gh} - \min_{h'}(\mathbf{w})_{gh'}\,;$$

$$\boldsymbol{H}'' = (h_1'', h_2'', \dots, h_T'')\,,$$

where

$$h_g'' = \min_{h''}(\mathbf{w})_{gh''} - \min_{h'''}(\mathbf{w})_{gh'''}\,.$$

Thus

$$C\{AUn_j\} - C\{AUn_j\,Un_i\} = \sum_{g=1}^{T} h_g''\,,$$

$$C\{A\} - C\{AUn_i\} = \sum_{g=1}^{T} h_g\,.$$

Now $h_g \leqslant 0$, $h_g'' \leqslant 0$, $\forall g = 1, \dots, T$. Also, $\forall g$, $h_g = h_g''$, unless

$$\min_{h''}(\mathbf{w})_{gh''} = \min_{h}(\mathbf{w})_{gh}\,,$$

where $h_g'' = 0$, and $h_g \geqslant 0 = h_g''$. Thus

$$\sum_{g=1}^{T} h_g'' \leqslant \sum_{g=1}^{T} h_g\,.$$

Thus

$$C\{AUn_j\} - C\{AUn_j\,Un_i\} \leqslant C\{A\} - C\{AUn_i\}\,.$$

Returning to the theorem, from equation (A3) we obtain

$$2(C\{J_t\} - C\{J_t\,Un_a\}) \geqslant (C\{J_t\} - C\{J_t\,Un_k\}) + (C\{J_t\} - C\{J_t\,Un_h\}) \qquad \forall n_k,\, n_h \in N\,, \tag{A5}$$

and from equation (A4), with $A = J_t$,

$$C\{J_t\} - C\{J_t\,Un_k\} \geqslant C\{J_t\,Un_h\} - C\{J_t\,Un_k\,Un_h\}, \qquad \forall n_k,\, n_h \in N\,. \tag{A6}$$

Combining equations (A5) and (A6) we obtain

$$2(C\{J_t\} - C\{J_t\,Un_a\}) \geqslant C\{J_t\} - C\{J_t\,Un_k\,Un_h\}\,. \tag{A7}$$

Next, by optimality,

$$C\{J_{t+1}\} \leqslant C\{J_t\, Un_a\}\,.$$

Thus

$$C\{J_t\} - C\{J_{t+1}\} \geqslant (C\{J_t\} - C\{J_t\, Un_a\}\,.$$

Substituting into equation (7) we obtain

$$2(C\{J_t\} - C\{J_{t+1}\}) \geqslant C\{J_t\} - C\{J_t\, Un_k\, Un_h\}\,. \tag{A8}$$

Choose n_k, n_h such that

$$J_t\, Un_k\, Un_h = J_{t+2}\,.$$

We may do this since, by assumption $J_t \subset J_{t+2}$. We now have

$$2(C\{J_t\} - C\{J_{t+1}\}) \geqslant C\{J_t\} - C\{J_{t+2}\}\,. \tag{A9}$$

Thus equation (A1) holds.