The Regional Optimal Population Points in Finland

By ASLAK HERVA
Academy of Finland

1. Problem of an optimal population point

A general problem in governmental policy, for example; is to designate the optimal place for a public centre. The optimum is understood here as a minimal total distance to each inhabitant. The sum may be weighted with special attributive measures, and inhabitants can be replaced by other objects which have a fixed location.

Distance is a very variable concept (for instance time, waterway, highway, airline, railway, pipeline, number of intervening opportunities, transport costs, social distance). Because this paper is a basic study we could regard distance as distance by helicopter, which of course is not very realistic in practice.

Demographically the population centre as a function of time describes the process of change in a spatial population distribution as a simple, single number pair: the location coordinates. Additionally the mean distance to the point can be calculated. This measures the compactness of the population living sites.

Isard [6] says: »Clearly such measures could be of considerable help in evaluating the degree of concentration or dispersion of given industries, number of manufacturing workers, etc. However, as Duncan, Cuzzort, and Duncan point out, when one is concerned with changes in a distribution pattern (e.g. over time), it is hardly likely that any single centrographic technique could furnish a complete or adequate description. A series of coefficients based on different sets of areal subdivision might well be preferable.» However I think the measures are useful.

Of course the information obtained from these variables is very simplified but the idea of the problem is very clear, complete, and well established.

The problem has been studied in many scientific sectors, especially in management science and operations research, but demographically I have only seen it applied to 24 Soviet Ukrainian towns [7]. A short reference to the problem is made in [11] and some wider considerations are in [10]. Fermat and Torricelli were interested in the solution more than three centuries ago.

Econometrists speak about the general spatial function which contains such independent variables as quantities of input other than transport, quantities of various transport inputs, and quantities of various outputs. Transport

The algorithm converges quickly compared to the trivial solution method which goes through all pairs (x,y) with constant steps and finally selects the optimum after a complete study of a reasonable area.

According to Kuhn and Kuenne [7] the method is very useful and attains the desired point within less than ten steps. My trials also support this hypothesis.

When using (3) we should utilize the generalized Newton method [13]. The algorithm is now

$$s_{p+1} = s_p + h(s_p) \qquad (p = 0,1,2,\ldots)$$

$$\text{where } h = \Delta^{-1}\begin{pmatrix} \dfrac{\partial\psi}{\partial\lambda}\cdot\dfrac{\partial^2\psi}{\partial\varphi\partial\lambda} - \dfrac{\partial\psi}{\partial\varphi}\cdot\dfrac{\partial^2\psi}{\partial\lambda^2} \\ \dfrac{\partial\psi}{\partial\varphi}\cdot\dfrac{\partial^2\psi}{\partial\lambda\partial\varphi} - \dfrac{\partial\psi}{\partial\lambda}\cdot\dfrac{\partial^2\psi}{\partial\varphi^2} \end{pmatrix} \quad \Delta = \begin{vmatrix} \dfrac{\partial^2\psi}{\partial\varphi^2} & \dfrac{\partial^2\psi}{\partial\varphi\partial\lambda} \\ \dfrac{\partial^2\psi}{\partial\lambda\partial\varphi} & \dfrac{\partial^2\psi}{\partial\lambda^2} \end{vmatrix} \neq 0, \qquad (8)$$

and s_0 is »near» the desired point. Thus we must know the approximate solution. This is generally possible by using the planar method. The matrix corresponding Δ should be negative definite. This is what we are not certain about and this is also the weak point of the method. One must be critical when deciding whether the global (in the mathematical sence) optimum has been reached or not.

Explicitely we have

$$x_{s_{p+1}} = x_{s_p} + \Delta^{-1}(g(D_\lambda,0)g(D_{\lambda\lambda},D_\lambda) - g(D_\varphi,0)g(D_{\lambda\lambda},D_\lambda))$$

$$y_{s_{p+1}} = y_{s_p} + \Delta^{-1}(g(D_\varphi,0)g(D_{\lambda\varphi},D_\varphi) - g(D_\lambda,0)g(D_{\varphi\varphi},D_\varphi))$$

$$g(u_i,v_i) = -\Sigma w_i(u_i(1 - d_i(s;r)^2)^{-1/2} - d_i(s;r)v_i(1 - d_i(s;r)^2)^{-3/2},$$

and D is the derivate of $d_i(s;r)$ with respect to the indices given.
Δ is $g(D_{\varphi\varphi},D_\varphi)g(D_{\lambda\lambda},D_\lambda) - g(D_{\varphi\lambda},D_\lambda)g(D_{\lambda\varphi},D_\varphi)$ and the D-derivates

$$D_\varphi = \sin\varphi_i\cos\varphi - \cos\varphi_i\sin\varphi\cos(\lambda-\lambda_i)$$

$$D_\lambda = -\cos\varphi_i\cos\varphi\sin(\lambda-\lambda_i)$$

$$D_{\varphi\varphi} = -\sin\varphi_i\sin\varphi - \cos\varphi_i\cos\varphi\cos(\lambda-\lambda_i)$$

$$D_{\varphi\lambda} = \cos\varphi_i\sin\varphi\sin(\lambda-\lambda_i)$$

$$D_{\lambda\varphi} = \cos\varphi_i\sin\varphi\sin(\lambda-\lambda_i)$$

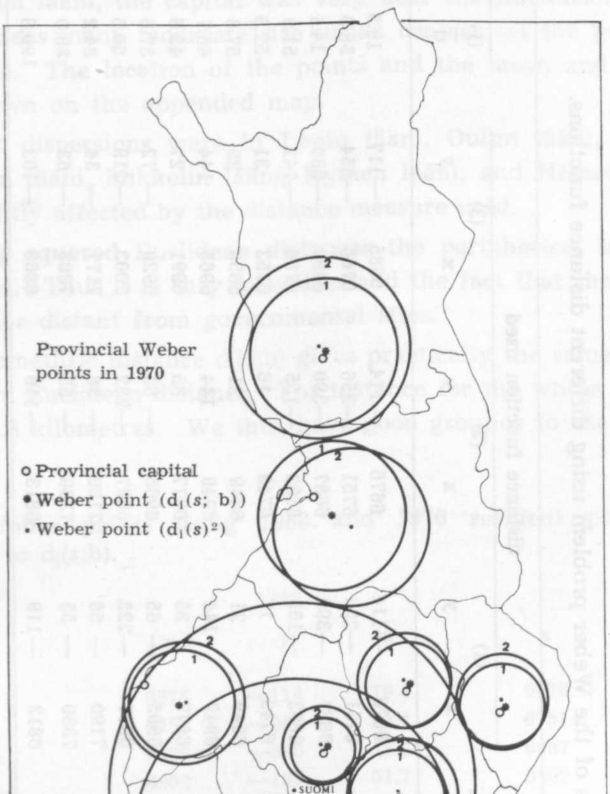
$$D_{\lambda\lambda} = -\cos\varphi_i\cos\varphi\cos(\lambda-\lambda_i).$$

The algorithm converges quadratically within a few steps but the starting point must be selected carefully.

We could also develop formulas for an ellipsoid using geodetical lines but these would be too laborious in relation to the improvement in exactness.

3. An application to Finnish provinces in 1970

An application of previous methods was made to the 1970 Finnish census population in 12 provinces and in the whole country. The units were communes and the whole population was thought to be living in the administrative commune centres. Of course this generates an error which is difficult to estimate. The same analyses were made with the resident populations in 1966 and 1970 but no comparision was made because of the questionable reliability of the data. The coordinate system was Universal Transversal Mercator (Gauss—Hannover or Gauss—Krüger) (see [2]) with 27 degrees east of Greenwich as the mean meridian. The latitude x is expressed in kilometres from the equator to the north and the longitude y east from the mean meridian measured on the earth ellipsoid (i.e., along the geodesian lines).



1 mean distance 2 standard distance

T a b l e 1. Solutions of the Weber problem using different distance functions.

province (lääni)	distance function used						distance f.			distance f.		
	(1)		(2)		(3)		(1)	(2)	(3)	(1)	(2)	(3)
	x	y	x	y	x	y	$\overline{d}$	$\overline{d}$	δ	d_c	d_c	d_c
Uudenmaan	6676	—114	6676	—114	6685	—117	18.5	18.5	31.4	0	0	9
Turun ja Porin	6751	—256	6751	—256	6766	—254	58.9	58.8	63.8	37	36	51
Åland province	6687	—390	6687	—390	6692	—385	14.1	14.1	19.5	2	2	10
Hämeen	6802	—155	6802	—155	6796	—142	51.9	51.9	56.7	39	39	30
Kymen	6743	17	6743	16	6753	31	52.5	52.4	57.6	33	32	46
Mikkelin	6849	21	6849	21	6856	29	57.6	57.5	64.4	10	9	20
Pohjois-Karjalan	6949	144	6949	144	6965	144	45.7	45.6	54.7	2	2	18
Kuopion	6977	35	6977	35	6991	27	43.6	43.5	51.4	1	0	16
Keski-Suomen	6905	— 65	6905	— 65	6920	— 72	37.6	37.5	47.7	0	0	17
Vaasan	6997	—222	6997	—222	7003	—219	60.5	60.5	66.4	50	50	52
Oulun	7190	— 56	7190	— 56	7177	— 34	84.2	84.1	95.7	30	30	54
Lapin	7380	— 55	7380	— 55	7386	— 63	88.0	87.9	108.2	0	0	10
Whole Finland	6813	—119	6813	—119	6868	—103	190.8	190.5	220.5	138	137	192

The data was prepared by Dr. Lauri Hautamäki who kindly gave me an opportunity to utilize it.

A data processing program was prepared and tested by the author for Burroughs B6700 Data Processing System at University of Helsinki Computing Centre using B6700 Extended ALGOL Programming Language.

The results of 1970 census data are given in Table 1. In Table 2 are found results using the 1966 and 1970 resident population.

The distance function in the tables are (1) $d_j(s)$, (2) $d_i(s;b)$, and (3) $d_i(s)^2$. $\overline{d}$ is for the arithmetic mean distance to the optimal point, and d_c the distance between s and the provincial capital. All numbers are in kilometres.

In six of the twelve provinces (lääni), namely Uudenmaan lääni, Åland province, Mikkelin lääni, Pohjois-Karjalan lääni, Kuopion lääni, Keski-Suomen lääni, and Lapin lääni, the capital was very near the Euclidean Weber point. In the other areas many moderate size urban centres set the point clearly in the countryside. The location of the points and the mean and standard distances are shown on the appended map.

The largest dispersions were in Lapin lääni, Oulun lääni, Vaasan lääni, Turun ja Porin lääni, Mikkelin lääni, Kymen lääni, and Hämeen lääni. The order was slightly affected by the distance measure used.

When using squared Euclidean distances the peripherical inhabitants are better observed. Thus it is easy to understand the fact that the Weber points are now further distant from governmental sites.

The trigonomethric distance $d_i(s;b)$ gives practically the same results as the straightforward Euclidean distance. For instance for the whole of Finland the difference is 0.3 kilometres. We thus have good grounds to use the Euclidean

Table 2. Weber points using 1966 and 1970 resident populations and $d = d_i(s;b)$.

province (lääni)	1966			1970		
	x	y	$\overline{d}$	x	y	$\overline{d}$
Uudenmaan	6676	—114	18.4	6676	—114	18.4
Turun ja Porin	6751	—256	59.4	6751	—256	58.8
Åland province	6687	—390	15.2	6687	—390	14.1
Hämeen	6802	—155	51.7	6802	—155	51.9
Kymen	6743	16	52.9	6743	16	52.4
Mikkelin	6849	21	57.8	6849	21	57.5
Pohjois-Karjalan	6949	144	47.5	6949	144	45.6
Kuopion	6977	35	44.8	6977	35	43.5
Keski-Suomen	6905	— 65	39.6	6906	— 65	37.5
Vaasan	6997	—221	60.8	6997	—222	60.5
Oulun	7188	— 53	86.1	7190	— 56	84.1
Lapin	7380	— 55	87.8	7380	— 55	87.9
Whole Finland	6815	—122	194.7	6813	—119	190.6

distance as such in most studies. The spherical study was made because of generality and as a verification check. It can be applied to any mondial area.

4. Additional remarks

When generalizing the problem, e.g. considering more variables than location alone, we come to applications of operations research [11].

Using the Weber point of population as a basic point, comparisions can be made when other weight variables w_i respectively have been used in the criterion function. For example such demographic variables as age and sex may be considered. Semidemographic indicators, such as economic and social variables, are also worth studying.

The functional form of the criterion function can be chosen in many ways. The three represented here may be adequate for most applications. Which one to use, d_i or d_i^2, is a question for an additional study.

References

[1] Bachi, R. Statistical Analysis of Geographic Series. Kaplan School, Hebrew University and Israel Central Bureau of Statistics. Jerusalem (1957).

[2] Bjerhammar, A. Geodesi. Almqvist & Wiksell, Stockholm (1967).

[3] Christaller, W. Central Places in Southern Germany. Englewood Cliffs, N. J. (1966).

[4] Hautamäki, L. »Preliminary methods in seeking optimal system of centres applied to supraprovincial level». Fennia, Societas geographica fenniae, 112 (1972).

[5] Henrici, P. Elements of Numerical Analysis. J. Wiley & Sons, London (1964).

[6] Isard, W. Methods of Regional Analysis: An Introduction to Regional Science. The Technology Press of the MIT and John Wiley & Sons, Inc.; sixth printing, New York (1969).

[7] Kuhn, H. & Kuenne R.: »An Efficient Algorithm for the Numerical Solutions of the Generalized Weber Problem in Spatial Economics», Journal of Regional Sciences, Vol. 4, No 2 (1962).

[8] Kuklinski, A. (ed.). Regional Information and Regional Planning. Mouton, Paris (1974).

[9] Matlin, I. S. Modelirovani'e razmeščeni'a naseleni'a. Problemy sovetskoj ekonomiki. Izdatel'stvo »Nauka», Moskva 1975.

[10] ReVelle, C., Marks D. & Liebman J. C. »An Analysis of Private and Public Sector Location Models», Management Science, Vol. 16, No 11 (July 1970).

[11] Seppälä, Y. Matemaattinen yhdyskuntasuunnittelu. Tammi, Helsinki (1973).

[12] Weber, A. Über den Standort der Industrien. Tübingen (1909). In English, Theory of the Location of Industry. University of Chicago Press, Chicago (1929).

[13] Zangwill, W. Nonlinear Programming: A Unifield Approach. Prentice-Hall, Inc., Englewood Cliffs, N. J. (1969).