GEOMETRIC PROGRAMMING AND ENTROPY MAXIMIZING MODELS

P. NIJKAMP[1] and J.H.P. PAELINCK[2]

1. *Introduction*

This paper is devoted to an analysis of the close relationship between geometric programming and entropy maximizing models. It will be proved that entropy maximizing models are a sub-class of geometric programming models, and that the concept of entropy can be interpreted and generalized by means of geometric programming. In addition, an illustration in the field of spatial analysis will be given.

2. *Geometric Programming.*

Geometric programming is a recently developed method for handling a broad class of nonlinear programming problems (see Duffin et al. (1967)). It is useful for programming problems written as posynomial , i.e. positive sums of multiplicative power functions. The standard format of a (primal) geometric program is :

$$\left.\begin{array}{l}
\min \psi = c_{01}\, x_1^{a^0_{11}}\, x_2^{a^0_{12}} \ldots x_I^{a^0_{1I}} + c_{02}\, x_1^{a^0_{21}}\, x_2^{a^0_{22}} \ldots x_I^{a^0_{2I}} + \ldots \\
\text{s.t.} \\
c_{11}\, x_1^{a^1_{11}}\, x_2^{a^1_{12}} \ldots x_I^{a^1_{1I}} + c_{12}\, x_1^{a^1_{21}}\, x_2^{a^1_{22}} \ldots + \ldots \leq 1 \\
\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
c_{K1}\, x_1^{a^K_{11}}\, x_2^{a^K_{12}} \ldots x_I^{a^K_{1I}} + c_{K2}\, x_1^{a^K_{21}}\, x_2^{a^K_{22}} \ldots + \ldots \leq 1 \\
x_i \geq 0 \; (i = 1, \ldots, I)
\end{array}\right\} \qquad (2.1)$$

The coefficients of this program are denoted by c_{kj} (k = 0, 1, ..., K; j = 1, ..., J), i.e. the coefficient in the j^{th} term of the k^{th}

---

(1) Erasmus University, Rotterdam, and Free University, Amsterdam
(2) Erasmus University, and Netherlands Economic Institute, Rotterdam.

side-condition ($k = 0$ means the objective function itself). These coefficients are assumed to be positive (hence the name 'posynomials'). The exponent a_{ji}^{k} denotes the exponent in the j^{th} term of the k^{th} side-condition related to the i^{th} variable ($i = 1, \ldots, I$).

Geometric programming has been developed first in chemical equilibrium systems, and is now also being used in economics and regional science (see, for instance, Nijkamp (1972) and Nijkamp and Paelinck (1972)).

The previous (primal geometric program is hard to solve, but it has a dual formulation which is more manageable. On denoting the dual variables by p_{kj} the dual program reads as :

$$\left.\begin{aligned}
&\max \tau = \left(\frac{c_{01}}{p_{01}}\right)^{p_{01}} \left(\frac{c_{02}}{p_{02}}\right)^{p_{02}} \ldots \times \left(\frac{c_{11}}{p_{11}}\right)^{p_{11}} \left(\frac{c_{12}}{p_{12}}\right)^{p_{12}} \ldots \times \left(\frac{c_{K1}}{p_{K1}}\right)^{p_{K1}} \left(\frac{c_{K2}}{p_{K2}}\right)^{p_{K2}} \\
&\qquad \times (p_{11} + p_{12} + \ldots)^{p_{11}+p_{12}+\ldots} \ldots \times (p_{K1} + p_{K2} + \ldots)^{p_{K1}+p_{K2}+\ldots} \\
&\text{s.t.} \\
&p_{01} + p_{02} + \ldots + p_{0J} = 1 \\
&a_{11}^{0} p_{01} + a_{21}^{0} p_{02} + \ldots + a_{11}^{1} p_{11} + a_{21}^{1} p_{21} + \ldots + a_{11}^{K} p_{K1} + a_{21}^{K} p_{K2} + \ldots = 0 \\
&\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
&a_{1I}^{0} p_{01} + a_{2I}^{0} p_{02} + \ldots + a_{1I}^{1} p_{11} + a_{2I}^{1} p_{12} + \ldots + a_{1I}^{K} p_{K1} + a_{2I}^{K} p_{K2} + \ldots = 0 \\
&p_{kj} \geq 0
\end{aligned}\right\} \quad (2.2)$$

The dual programme is more tractable owing to the linear equality restrictions. Since the primal and dual programmes are related to each other by means of primal-dual relationships, one needs to solve only the dual programme in order to find the primal solution. The solution of the dual geometric programme and the identification of active or inactive primal constraints makes use of a constrained gradient method, which is a particular kind of a so-called hill-climbing method. This solution algorithm as well as an application to the planning strategy of new activities in an industrial complex is contained in Nijkamp (1972).

It is obvious that there should be a one-to-one relationship between the primal variables x_i and the dual variables p_{kj} at the optimum. These relationships are according to Duffin et al. (1967) equal to :

$$p_{0j} = \frac{c_{0j} x_1^{a_{j1}^0} \cdots x_I^{a_{jI}^0}}{\varphi} \tag{2.3}$$

and

$$p_{kj} = \frac{\lambda_k c_{kj} x_1^{a_{j1}^k} \cdots x_I^{a_{jI}^k}}{\varphi} \; ; \; k = 1, \ldots, K \; ; \tag{2.4}$$

in which λ_k represents the Lagrange multiplier associated with the k^{th} constraint of (2.1).

In case of an optimum solution of a normal programming model, the value of the primal and dual objective function should be equal (the so-called 'duality-theorem'). Similarly, the value of a primal geometric objective function is at the optimum exactly equal to the value of the dual geometric objective function.

3. *The concept of Entropy.*

Modern scientific research is predominantly characterized by a formal, conceptual way of thinking. More and more attempts are being made at modelling phenomena and systems at various scales. The rise of *entropy maximizing models* fits precisely into this pattern.

The concept of entropy is originally derived from thermodynamics (see, for instance, Fast (1970)). Closed physical systems can adopt numerous states. In general, however, the elements of such a system tend towards an arrangement which can be organized in as many ways as possible (a maximum 'disorder'). Such a tendency is called the maximization of the entropy of a system : for a closed physical system a situation of disorder is much more probable than an ordered situation.

In addition to the *physical* concept of entropy, this concept plays an important role in *information* theory (see, for instance, Theil (1967)). In the latter theory entropy represents expected information : it indicates the degree of uncertainty about the realizations of events in information systems, represented by a discrete probability distribution (see, for instance, Jaynes (1957)). The mathematical specification of entropy in information systems bears a close resemblance to that used in physics, for both concepts are based on a common statistical background.

The major part of modern science is concerned with systems, in the sense of entities consisting of specialized, interdependent parts (see Berry (1964)). Therefore, it is understandable that the concept of entropy is frequently being used in order to obtain insight into the uncertainty of systems. In this analogous way, the notion of entropy is introduced into the social sciences (hence the name 'social physics'). Recently the concept of entropy has also found applications in regional systems (for instance, trip distribution, freight flows, migration flows, and activity allocations). The rather involved spatial processes are overloaded with uncertainties. The entropy concept attempts determining the most probable spatial configurations of a system which can adopt numerous uncertain spatial states. This implies that entropy in regional research is *a probability concept, describing the outcome of a stochastic process.*

In accordance with the distinction set out above between entropy in physics and in information theory, a twofold use of entropy in regional research may be distinguished.

In the first place, entropy can be used as a descriptive device, based on the assumption of a spatial equilibrium. Such a most probable state of a system corresponds with a maximum entropy of the system. In this way one arrives at the general specification of a gravity model between spatial entities (see, for instance, Wilson (1970), Batty (1970[a]), Cordey Hayes and Wilson (1971)). The entropy maximizing models permit the determination of "most probable" spatial flows of commodities, migration, etc. By assessing the parameters associated with the resulting gravity model one could use the entropy approach for the projection of future spatial flows, but here some difficulties arise, due to the unknown values of the associated dual variables.

Secondly, entropy can be used as a *measure for the degree of organization in a spatial system.* In this way the concept of entropy is a tool for studying spatial differentiation, for instance, by inspecting whether certain spatial configurations are completely arbitrary and disordered, or whether these configurations show a certain degree of spatial organization or regularity. Attempts in this field have been made among others by Berry (1964), Medvedkow (1966), Gurevitch (1969), and Semple and Gauthier (1972).

In this paper the rather 'mechanical' working of the entropy models will be re-considered. In the first place, it will be shown that *entropy* *maximizing models are a sub-class of a more general class of models*, viz. *geometric programming models*. This implies that the frequently too simple specification of entropy maximizing models can easily be extended by additional constraints on spatial patterns. These more general entropy maximizing models do not result in the simple gravity solutions of the original entropy maximizing models, but a recently developed solution method for geometric programming allows efficient determination of the optimum results.

Furthermore, it will be proved that *entropy maximizing models possess a dual formulation*, which shows the close links between the gravity approach and the entropy approach. This dual formulation of the original entropy maximizing models allows a more appropriate interpretation of entropy in terms of economic preferences and constraints.

4. *Entropy maximizing models in regional research.*

The use of entropy models will be illustrated by means of a spatial application. Entropy models are currently used in many fields of spatial analysis : trip distribution, transport and traffic flows, location models, allocation models for consumer expenditures, interregional freight flows, etc. This entropy concept is related to the *state* of a system; it is characteristic concerning the assignment of elements to a spatial structure. In this paper major attention will be paid to journey-to-work decisions of workers travelling from zone i to zone j. It is obvious, however, that a similar reasoning can be applied to all other problems concerning the (spatial) distribution of flows.

There are numerous possibilities to assign the journey-to-work flows of an origin-destination table, supposing the number of worker residences and jobs is given. Each particular spatial configuration can be obtained in many different ways, dependent on the number of possible states associated with a certain configuration. Anyhow, each distribution of flows within an origin-destination table should satisfy a set of *additivity* conditions. First, one defines the following variables :

T_{ij} : (unknown) number of persons living in zone i and working in zone j;

O_i : (known) total number of workers living in zone i;

D_j : (known) total number of jobs provided in zone j;

c_{ij} : (knwon) unit cost of travelling from zone i to zone j;

C : (known) total travel budget.

Then the additivity conditions of the urban or regional travel system can be written as :

$$\sum_{j=1}^{J} T_{ij} = O_i \qquad (4.1)$$

$$\sum_{i=1}^{I} T_{ij} = D_j \qquad (4.2)$$

$$\sum_{i=1}^{I} \sum_{j=1}^{J} c_{ij} T_{ij} = C \qquad (4.3)$$

Next, one may ask : which spatial distribution of the trips is the most probable arrangement within the system defined by (4.1) - (4.3) ? According to the entropy assumption the most probable arrangement is formed by the spatial configuration which possesses the greatest number of states associated with it. If one denotes the total number of workers within the system by T, i.e.,

$$T = \sum_{i=1}^{I} \sum_{j=1}^{J} T_{ij} , \qquad (4.4)$$

the number of ways the individuals can be assigned to a particular origin-destination table with elements T_{ij} is equal to :

$$\omega(T_{ij}) = \frac{T!}{\prod_{i=1}^{I} \prod_{j=1}^{J} T_{ij}!} \qquad (4.5)$$

The maximum value of $\omega(T_{ij})$ represents the maximum number of assignments of workers to an origin-destination matrix. This maximum number of states of a spatial configuration dominates the number of states of alternative arrangements to such a degree, that the spatial allocation of the trips associated with the maximum is the most likely one. This implies that the entropy approach is a probability approach : the 'optimal' spatial arrangement of a system is formed by the most probable arrangement, which is characterized by the fact that $\omega(T_{ij})$ of this arrangement is at a maximum.

Since $\omega(T_{ij})$ is invariant against a monotonically increasing transformation, the objective function of the spatial system can be written in a logarithmic form as :

$$\ln \omega(T_{ij}) = \ln T! - \sum_{i=1}^{I} \sum_{j=1}^{J} \ln T_{ij}! \tag{4.6}$$

or after Stirling's approximation :

$$\ln \omega(T_{ij}) = \ln T! - \sum_{i=1}^{I} \sum_{j=1}^{J} (T_{ij} \ln T_{ij} - T_{ij}) \tag{4.7}$$

Since ln T! is a constant, it can be left out of consideration in the maximization procedure, so that the ultimate program to be solved is :

$$\left.\begin{aligned} &\max \ln \omega(T_{ij}) = - \sum_{i=1}^{I} \sum_{j=1}^{J} (T_{ij} \ln T_{ij} - T_{ij}) \\ &\text{s.t.} \\ &\sum_{j=1}^{J} T_{ij} = O_i \\ &\sum_{i=1}^{I} T_{ij} = D_j \\ &\sum_{i=1}^{I} \sum_{j=1}^{J} c_{ij} T_{ij} = C \end{aligned}\right\} \tag{4.8}$$

The outcome of program (4.8) is easily obtained by constructing a Lagrangean function L for this constreined maximization problem, i.e.,

$$\begin{aligned} L = \ln \omega(T_{ij}) &+ \sum_{i=1}^{I} \lambda_i (O_i - \sum_{j=1}^{J} T_{ij}) + \sum_{j=1}^{J} \mu_j (D_j - \sum_{i=1}^{I} T_{ij}) \\ &+ \beta (C - \sum_{i=1}^{I} \sum_{j=1}^{J} c_{ij} T_{ij}) , \end{aligned} \tag{4.9}$$

where λ_i, μ_j and β are the Lagrange multipliers associated with (4.1), (4.2) and (4.3), respectively.

The necessary conditions for the maximum are :

$$\frac{\partial L}{\partial T_{ij}} = - \ln T_{ij} - \lambda_i - \mu_j - \beta c_{ij} = 0 \tag{4.10}$$

or :

$$T_{ij} = e^{-\lambda_i - \mu_j - \beta c_{ij}} \tag{4.11}$$

Next, writing :

$$A_i = \frac{e^{-\lambda_i}}{O_i} \tag{4.12}$$

and :

$$B_j = \frac{e^{-\mu_j}}{D_j} \tag{4.13}$$

one obtains :

$$T_{ij} = A_i\ B_j\ O_i\ D_j\ e^{-\beta c_{ij}} \tag{4.14}$$

By making use of (4.1) and (4.2) it is easily seen that :

$$A_i = \{ \ {}_{j=1}^{J}\ B_j\ D_j\ e^{-\ c_{ij}} \}^{-1} \tag{4.15}$$

and

$$B_j = \{ \sum_{i=1}^{I} A_i\ O_i\ e^{-\beta c_{ij}} \}^{-1} \tag{4.16}$$

It can be proved that the second-order conditions for a maximum are also satisfied ($\ln \omega(T_{ij})$ is a concave function), so that there is only one (absolute) maximum.

The final result (4.14) represents the most probable flow between each point of origin i and each point of destination j. This distribution function of the flows within the system concerned appears to possess a *gravity* specification. Such a gravity model is frequently used in regional research in order to analyse or estimate the flows between two spatial masses. The spatial distribution of trips, represented by (4.14), can be used for several purposes.

In the first place, the result can be used to derive an "optimal" (i.e. most likely) trip distribution in a spatial system, given O_i, D_j and C. This implies that the *most probable state* of the system will be determined on the basis of the known totals O_i, D_j and C. Then one has to determine first the parameters A_i, B_j and β. Such a (numerical) estimation can be carried out in an iterative way, for instance by starting with initial values of all B_j's and of , by calculating the resulting values of all A_i's by means of (4.15), by calculating in its turn a new series of B_j by means of (4.14), and so forth, until the procedure converges to an equilibrium point. By substituting the resulting values of A_i and B_j into (4.14), and, next, by substituting T_{ij} into (4.3), one can try to calibrate β, given the value C.

Once A_i, B_j and β have been assessed, the 'optimal' (i.e., most likely) flows T_{ij} can be determined. Such a most probable equilibrium state can be used as a planning objective in a spatial system. Next, if C is unknown (which is frequently the case), A_i and B_j can be estimated up to a multiplicative constant (viz., the exponential cost function). By varying the parameter β, different states of the system are found. In this way one might also approximate the *actual* state of the system for a particular value of β. As a further step formula (4.15) can be used for projecting spatial flows. Then the *actual* observations on T_{ij} from the past are used to estimate the unknown parameters A_i, B_j and β. In this case, the parameters A_i, B_j and β are determined by means of regression methods (or alternative numerical methods; see, for instance, Batty (1970[b]), Batty and Mackie (1972), Chisholm and O'Sullivan (1973), Hyman (1969), Wilson et al. (1969)). One might, for instance, take the natural logarithm from both sides of (4.15). Then, by assuming that T_{ij} can be explained from O_i, D_j and c_{ij}, one can estimate the corresponding parameters. Once the parameters A_i, B_j and β are known, (4.15) can be used as a relationship projecting the most likely future development of the spatial system concerned, given the knowledge of O_i and D_j. One should keep in mind, however, the special assumption implicit in the invariance of λ_i, μ_j and β, whereas it is obvious that a change in the future structure of the system will affect the dual variables.

5. *Entropy and Geometric Programming.*

The entropy model described in the previous paragraph can be represented in an alternative way, which is closely linked up with the probability approach implicit in information theory. If one defines the relative proportion :

$$p_{ij} = \frac{T_{ij}}{T} \tag{5.1}$$

one may interpret p_{ij} as the probability that a trip from i to j will be undertaken. It is obvious that the probabilities add up to one, since according to (4.4) the individual trips T_{ij} satisfy the additional condition :

$$\sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{T_{ij}}{T} = 1 \tag{5.2}$$

Next, if one substitutes (5.1) into $\ln \omega(T_{ij})$, one obtains

$$\begin{aligned}\ln \omega(T_{ij}) &= \ln \omega(p_{ij}) \\ &= -\sum_{i=1}^{I}\sum_{j=1}^{J}\{p_{ij}T \ln(p_{ij}T) - p_{ij}T\} \\ &= -T\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij}\ln p_{ij} - T\sum_{i=1}^{I}\sum_{j=1}^{J}(\ln T - 1)p_{ij}\end{aligned} \tag{5.3}$$

By making use of (5.2) one may write (5.3) as :

$$\ln \omega(p_{ij}) = -T\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} \ln p_{ij} - T(\ln T - 1) \tag{5.4}$$

Because $T(\ln T - 1)$ is a constant, it can be left out of consideration in the maximization procedure. Furthermore, the objective function is invariant against a monotonically increasing transformation, so that the ultimate objective function equivalent to (5.3) is :

$$\Omega = \ln \omega(p_{ij}) = -\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} \ln p_{ij} \tag{5.5}$$

The latter expression is a measure of uncertainty, introduced in information theory by Shannon and Weaver (1949); it is called the entropy of a probability distribution of p_{ij}. Wilson (1970) has proved that (5.5) is a unique, unambiguous measure of uncertainty. The latter expression is also valid in smaller systems, since the derivation of the entropy of a probability distribution can be carried out without making use of Stirling's approximation (Wilson, 1970).

In a similar way the constraints (4.1) - (4.3) can be rewritten by means of (5.1). One finds successively :

$$\sum_{j=1}^{J} p_{ij} T = O_i \tag{5.6}$$

$$\sum_{i=1}^{I} p_{ij} T = D_j \tag{5.7}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J} c_{ij} p_{ij} T = C \tag{5.8}$$

Therefore, the ultimate program to be solved is to maximize (5.5) subject to (5.2) and (5.6) - (5.8). This program can be re-specified as :

$$\left.\begin{aligned}
&\max\ \Omega = -\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij}\ln p_{ij}\\
&\text{s.t.}\\
&\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} = 1\\
&\sum_{j=1}^{J} p_{ij} - \frac{O_i}{T} = 0\\
&\sum_{i=1}^{I} p_{ij} - \frac{D_j}{T} = 0\\
&-\sum_{i=1}^{I}\sum_{j=1}^{J} c_{ij}\,p_{ij} + \frac{C}{T} = 0
\end{aligned}\right\} \qquad (5.9)$$

The previous entropy model appears to be a special member of a broad class of models, viz. geometric programming models. This can easily be seen by taking the logarithm of the dual objective function in (2.2) :

$$\left.\begin{aligned}
&\text{Max }\Omega = p_{01}(\ln c_{01} - \ln p_{01}) + p_{02}(\ln c_{02} - \ln p_{02}) + \ldots\\
&\qquad + p_{11}(\ln c_{11} - \ln p_{11}) + p_{12}(\ln c_{12} - \ln p_{12}) + \ldots\\
&\qquad + p_{K1}(\ln c_{K1} - \ln p_{K1}) + p_{K2}(\ln c_{K2} - \ln p_{K2}) + \ldots\\
&\qquad + (p_{11} + p_{12} + \ldots)\ln(p_{11} + p_{12} + \ldots) \ldots\\
&\qquad + (p_{K1} + p_{K2} + \ldots)\ln(p_{K1} + p_{K2} + \ldots)\\
&\text{s.t.}\\
&p_{01} + p_{02} + \ldots + p_{0J} = 1\\
&a^0_{11}p_{01} + a^0_{21}p_{02} + \ldots + a^1_{11}p_{11} + a^1_{21}p_{12} + \ldots\\
&\qquad\qquad + a^K_{11}p_{K1} + a^K_{21}p_{K2} + \ldots = 0\\
&\vdots\\
&a^0_{1I}p_{01} + a^0_{2I}p_{02} + \ldots + a^1_{1I}p_{11} + a^1_{2I}p_{12} + \ldots\\
&\qquad\qquad + a^K_{1I}p_{K1} + a^K_{2I}p_{K2} + \ldots = 0\\
&p_{kj} \geq 0;\ k = 0, \ldots, K;\ j = 1, \ldots J\ .
\end{aligned}\right\} \qquad (5.10)$$

It is easily seen that (5.10) is a generalization of (5.9) so that an entropy maximizing model is essentially a specific type of a dual geometric model. This implies that in addition to (4.1), (4.2) and (4.3) many other, alternative constraints can be taken into consideration (for instance, congestion, trip distribution, etc.). By adding more constraints it will become extremely difficult to derive a simple gravity solution such as (4.14), but since this extended program corresponds to a geometric program, one can use a constrained gradient method to solve this program, albeit that is this way not an analytical, but a numerical solution will be obtained. Therefore, one may conclude that extensions of the elementary entropy maximizing models can easily be handled by means of geometric programming.

Finally, some attention will be paid to the relationship between entropy, Baysesian statistics and geometric programming. There is a close relationship between entropy maximizing models and Baysesian statistics, because there is a close correspondence between the entropy objective function and the likelihood function of a set of random variables in statistical analysis. Baysesian statistics includes subjective elements, caused by the use of certain prior information. This gives rise to conditional probabilities, viz. the probability that certain events will occur, given a certain state of prior knowledge. Theil (1967) has also dealt with these conditional probabilities in the entropy approach to information theory. The use of conditional probabilities will now be proved to provide more insight into the close relationship between entropy and geometric programming.

It has already been stated that p_{ij} is the probability that a trip originating in zone i will terminate in zone j. The entropy of the probability distribution of p_{ij} is represented in (5.5). Next, one defines :

$$p_{i.} = \sum_{j=1}^{J} p_{ij} \qquad (5.11)$$

as the probability of a trip from i (i.e., the probability of the marginal distribution of the trip). By means of vertain operations derived from probability theory (see Appendix) one may create an expression for the average conditional entropy of the trips terminating in all zones of destination, given the condition that the marginal probabilities of the trips moving from each zones of origin are known. This average conditional entropy is according to (A.2) equal to :

$$\Omega_{AC} = - \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} \ln p_{ij} + \sum_{i=1}^{I} p_{i.} \ln p_{i.} \tag{5.12}$$

$$= - \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} \ln p_{ij} + \sum_{i=1}^{I} \{ (\sum_{j=1}^{J} p_{ij}) \ln (\sum_{j=1}^{J} p_{ij}) \}$$

By comparing (5.12) with (5.10) it is easily seen that the average conditional entropy possesses approximately the standard format of a dual geometric objective function. Inversely, one may conclude that a *dual geometric program is essentially a general formulation of an average conditional entropy expression*. This important conclusion proves also that there is no essential difference between entropy maximizing models and normal programming models.

6. A *dual entropy Model.*

Entropy models are a special sub-class of dual geometric programming models. As each dual programming model is related in a one-to-one way to a primal program, it is worthwhile to consider a primal geometric program, and to inspect whether an entropy model can also be interpreted with the aid of a primal geometric program. This might illuminate the working and the essence of entropy maximizing models. This is completely in line with the classical linear programming models : if such a primal program relates to quantities (for instance, spatial flows), the corresponding dual program describes the price or cost side (by means of shadow prices). In a similar way an attempt will be made here to derive and to interpret shadow (or dual) variables associated with a normal entropy model. Such a *dual* version of an entropy maximizing model will be proved to correspond to a *primal* geometric program.

Suppose, one assumes the most elementary entropy model :

$$\left. \begin{array}{l} \max \ \Omega = - \sum_{j=1}^{J} p_j \ln p_j \\ \text{s.t.} \\ \sum_{j=1}^{J} p_j = 1 \end{array} \right\} \tag{6.1}$$

It is evident that without additional constraints the entropy is at a maximum if all probabilities are equal (see also Richter (1969, p.67)), i.e., if :

$$p_j = \frac{1}{J} \ ; \quad j = 1, \ldots, J \tag{6.2}$$

Then the maximum entropy is equal to :

$$\Omega = \ln J \,, \tag{6.3}$$

whereas the anti-log Ω_a of (6.3) is :

$$\Omega_a = e^{\Omega} = J \tag{6.4}$$

If one writes the entropy model (5.1) as a (degenerated) dual geometric program (see (5.10)), one will obtain :

$$\left.\begin{array}{l} \max \Omega = p_1 \, (\ln 1 - \ln p_1) + \ldots + p_J(\ln 1 - \ln p_J) \\ \text{s.t.} \\ p_1 + \ldots + p_J = 1 \\ 0.p_1 + \ldots + 0.p_J = 0 \end{array}\right\} \tag{6.5}$$

It is obvious that in the simplified entropy model (6.1) all parameters a_{ji}^k from the dual geometric program (5.10) are equal to zero. The *dual* version of the entropy model (6.5) (i.e. the *primal* version of the dual geometric model (6.5)) can easily be found by means of (2.1) :

$$\min \varphi = 1.x_1^0 x_2^0 \ldots x_1^0 + 1.x_1^0 x_2^0 \ldots x_1^0 + \ldots = J \tag{6.6}$$

Evidently, there are no side-conditions in (6.6), as (6.5) does not include the parameters a_{ji}^k $(k = 1, \ldots, K)$, so that (6.6) is also a degenerated program. It should be noted that the 'optimum' value of (6.6), i.e. J, is exactly equal to the optimum value of the dual objective function (5.4). By means of (2.3) one can easily check that the outcome (6.2) is also correct.

Next, the entropy model discussed in section 5 will be considered. In a way similar to that described above, the dual version of this entropy model will be derived and interpreted as a primal geometric program. It has already been stated that the entropy is invariant against a monotonic increasing transformation, so that maximizing (5.4) is equivalent to maximizing (5.5). In order to facilitate the dual version of an entropy model, the following monotonic increasing transformation of (5.4) is chosen as an entropy function :

$$\max \Omega = - \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij}(\ln p_{ij} - \ln T) \tag{6.7}$$

For the sake of clarity, 3 points of origin and 3 points of destination are assumed. Then the entropy model can be written as :

$$\max \Omega = - \sum_{i=1}^{3} \sum_{j=1}^{3} p_{ij} (\ln p_{ij} - \ln T)$$

s.t.

$$\left.\begin{array}{l}
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -O_1/T \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & -O_2/T \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -O_3/T \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -D_1/T \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & -D_2/T \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & -D_3/T \\
-c_{11} & -c_{12} & -c_{13} & -c_{21} & -c_{22} & -c_{23} & -c_{31} & -c_{32} & -c_{33} & C/T
\end{bmatrix}
\begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{31} \\ p_{32} \\ p_{33} \\ 1 \end{bmatrix}
=
\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\
p_{ij} \geq 0 \; ; \; i = 1, 2, 3; \; j = 1, 2, 3
\end{array}\right\} \tag{6.8}$$

Program (6.8) contains 10 variables, viz. 9 unknown variables p_{ij} and one known variable 1. It is easily seen that program (6.8) is equivalent to the dual geometric program (5.10). Therefore, by means of (2.1) the dual version of this entropy model can be derived. Since program (6.8) contains 10 variables, the dual version will contain 10 terms. Furthermore, this dual version will contain 7 dual variables, since program (6.8) contains 7 constraints (apart from the first additivity condition for the probabilities). Then the dual program associated with the (primal) entropy model (6.8) is equal to :

$$\left.\begin{aligned}
\min \varphi = {} & T x_1 x_4 x_7^{-c_{11}} + T x_1 x_5 x_7^{-c_{12}} + T x_1 x_6 x_7^{-c_{13}} + T x_2 x_4 x_7^{-c_{21}} + \\
& T x_2 x_5 x_7^{-c_{22}} + T x_2 x_6 x_7^{-c_{23}} + T x_3 x_4 x_7^{-c_{31}} + T x_3 x_5 x_7^{-c_{32}} + \\
& T x_3 x_6 x_7^{-c_{33}} \\
\text{s.t.} & \\
& x_1^{-O_1/T} x_2^{-O_2/T} x_3^{-O_3/T} x_4^{-D_1/T} x_5^{-D_2/T} x_6^{-D_3/T} x_7^{C/T} = 1
\end{aligned}\right\} \quad (6.9)$$

The dual form of an entropy model appears to be a particular type of a primal geometric program. It is composed of a positive polynomial ('posynomial') with many terms each of which is non-linear. Furthermore, the dual entropy model appears to contain only one equality constraint; this equality is due to the fact that the Lagrange multiplier associated withtthe latter constraint is positive (see the last element of the vector at the left-hand side of (6.8). The equality constraint is a positive linear combination of power function products.

An interpretation to the dual entropy model can be given by considering the separate terms of the objective function and the constraint. Each term of the objective function is a power product composed of 3 (dual) variables. The first dual variable relates to the 'push' effect of a zone of origin, the second one to the 'pull' effect of a zone of destination, and the last one to the cost between these zones. Therefore, each term is a gravity model, indicating the volume of flows between corresponding zones. Concluding, the objective function attempts to minimize the aggregate flows within the system, subject to a certain condition. This condition states that the weighted (geometric) average of all push and pull-effects should be equal to the shadow variable of the costconstaint (x_7), weighted with the average travel cost. The implicit assumption of an entropy model appears to be indeed a gravity model.

APPENDIX

Derivation of a conditional entropy

The probability p_{ij} can be conceived of as the probability that a trip originating in zone i will terminate in zone j. The entropy of the probability distribution of p_{ij} is represented by (5.5). Next, one may define the *conditional probability* that a trip will terminate in zone j, given the condition that this trip originates from zone i. This conditional probability is denoted by $p_{ij}/p_{i.}$, where the marginal probability $p_{i.}$ is defined in (5.11). The entropy Ω_C of the conditional probability $p_{ij}/p_{i.}$ is equal to :

$$\Omega_C = - \sum_{i=1}^{I} \sum_{j=1}^{J} (p_{ij}/p_{i.}) \ln(p_{ij}/p_{i.}) \tag{A.1}$$

Next, one may calculate the average conditional entropy of the trips attracted by all zones of destination, given the fact that the marginal probabilities of the trips originating in each zone i are known. Such an average conditional entropy requires weights to be assiegned to each component. If one uses $p_{i.}$ as weights, the average conditional entropy Ω_{AC} is calculated as :

$$\begin{aligned} \Omega_{AC} &= - \sum_{i=1}^{I} p_{i.} \sum_{j=1}^{I} (p_{ij}/p_{i.}) \ln(p_{ij}/p_{i.}) \\ &= - \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} \ln(p_{ij}/p_{i.}) \\ &= - \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij}(\ln p_{ij} - \ln p_{i.}) \\ &= - \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} \ln p_{ij} + \sum_{i=1}^{I} \{(\sum_{j=1}^{J} p_{ij})\ln(\sum_{j=1}^{J} p_{ij})\} \end{aligned} \tag{A.2}$$

References

BATTY, M., "An Activity Allocation Model for the Nottinghamshire-Derbyshire Subregion", *Regional Studies*, vol.4, 1970(a), pp.307-332.

- "Models and Projections of the Space Economy : A Sub-regional Study in Northwest England", *Town Planning Review*, vol.41, 1970(b), pp. 121-146.

BATTY, M., and MACKIE, S., "The Calibration of Gravity, Entropy and Related Models of Spatial Interaction", *Environment and Planning*, vol.4, 1972, pp.205-233.

BERRY, B.J.L., "Cities as Systems within Systems of Cities", *Papers and Proceedings of the Regional Science Association*, vol.13, 1964, pp.147-163.

CHISHOLM, M. and O'SULLIVAN, P., *Freight Flows and Spatial Aspects of the British Economy*, Cambridge University Press, Cambridge, 1973.

CORDEY HAYES, M., and WILSON, A.G., "Spatial Interaction", *Socio-Economic Planning Sciences*, vol.5, 1971, pp.73-95.

DUFFIN, R.J., PETERSON, E.L. and ZENER, C.M., *Geometric Programming*, Wiley, New York, 1967.

FAST, J.D., *Entropy*, MacMillan, London, 1970.

GUREVITCH, B.L., "Geographical Differentiation and its Measures in a Discrete System", *Soviet Geography*, vol.10, nr 9, pp.387-412.

HYMAN, G.M., "The Calibration of Trip Distribution Models", *Environment and Planning*, vol.1, 1969, pp.105-112.

JAYNES, E.T., "Information Theory and Statistical Mechanics", *Physical Review*, vol.106, 1957, pp.620-630.

MEDVEDKOV, Y.V., "The Concept of Entropy in Settlement Pattern Analysis", *Papers and Proceedings of the Regional Science Association*, vol.18, 1967, pp.165-168.

NIJKAMP, P., *Planning of Industrial Complexes by Means of Geometric Programming*, Rotterdam University Press, Rotterdam, 1972.

NIJKAMP, P., and PAELINCK, J.H.P., "An Interregional Model of Environmental Choice", *Papers and Proceedings of the Regional Science Association* (to appear), 1972.

RICHTER, K.J., *Verkehrsökonometrie*, Westdeutscher Verlag, Cologne, 1969.

SEMPLE, R.K. and GAUTHIER, H.L., "Spatial-Temporal Trends in Income Inequalities in Brazil", *Geographical Analysis*, vol.4, nr 2, 1972, pp.169-180.

SHANNON, C., and WEAVER, W., *The Mathematical Theory of Communication*, University of Illinois Press, Illinois, 1949.

THEIL, H., *Economics and Information Theory*, North-Holland Publishing Co., Amsterdam, 1967.

WILSON, A.G., *Entropy in Urban and Regional Modelling*, Pion Limited, London, 1970.

WILSON, A.G., HAWKINS, A.F., HIU, G.J., and WAGON, D.J., "Calibration and Testing of the SELNEC Transport Model", *Regional Studies*, vol.3, 1969, pp.337-350.