ON THE FORMULATION OF A STOCHASTIC USER EQUILIBRIUM MODEL CONSISTENT WITH THE RANDOM UTILITY THEORY - A CONJUGATE DUAL APPROACH -

Toshihiko Miyagi

Department of Civil Engineering
Gifu University
1-1 Yanagido, Gifu 5Ø1-11
Japan

1. INTRODUCTION

The purpose of this paper is first to provide a theoretically valid foundation for combining two current approaches to the modeling of travel choice behavior, viz., the entropy maximizing approach (1,2) and the random utility maximizing approach(3,4), and second to show that the stochastic network equilibrium model with logit-based loading proposed by Fisk(5) is naturally derived from the random utility theory. Then a solution algorithm of the stochastic network equilibrium problem based on the contraction mapping theorem will be proposed.

Random utility and Entropy models appear to be quite different. This is because the random utility model is probabilistic in nature and deals with discrete events and utility characterized by attributes of alternatives, while the entropy model, although its behavioral interpretation in terms of a relation to utility has not been discussed so far, is deterministic in nature and treats continuous flows. So strong criticism that its theoretical ground is merely analogous to the statisticial theory or no more than ad hoc computational procedures, has been leveled at it.

On the other hand, however, from the view point of network equilibrium models, the entropy model including cumulative cost functions seems to be capable of predicting various flow patterns incurred according to trip makers' information for their route choices. The traditional equilibrium model deals with too extreme conditions from a behavioral view point in the sense that all routes utilized between each O-D pair must have the same route costs, and all routes with higher costs are never used. Such conditions are overly simplistic and not supported by the random utility theory. On the contrary, it has not been considered that the random utility model is able to account for the Wardrop equilibria.

With these ideas in mind, our purpose is first to show that the relationship between entropy and utility maximization approaches is far more foundamental than that has been believed so far. Indeed, both approaches are essentially identical and are two different representation methods of the same choice problem. The conjugate theory gives us more rigorous and general results to interpret this equivalency than any one that currently exists (including Williams (6), Anas (7) and Leonardi (8). A theoretical interpretation for the solutions to entoropy maximization models is provided by smith (9), however, his approach is different from the viewpoint that is taken in this paper in that his theory is based on the cost-efficiency principle developed by himself.

The next section establishes a Correspondence relationship which shows that the entropy maximization model can be derived within the theoretical framework of the random utility theory, and that its extreme value function exactly corresponds to the expected maximal utility (EMU) that each individual (decision-maker) derives from a set of alternatives.

The expected maximum utility has some important properties. One of those is the similarity with the constant elastisity substitution (CES) function, that is, if the dispersion parameter included in the EMU approaches infinity, then the EMU function exhibits the L-letter utility function (i.e. The Leontief utility function). Then, the EMU function describes the strict utility maximization behavior. In other words, if one defines the strict utility as the negative travel costs, then the EMU function accounts for the behavior of pure cost-minimizer. In section 3, using this property together with the conjugate theory, we derive the network equilibrium model equivalent in its form to the stochastic network equilibrium (SNE) model proposed by Fisk (5) within the random utility framework. Thus, this paper may serve to show that the random utility framework can interpret the Wardrop equlibria as being a special case of the EMU function. Furthermore, it will be shown that Fisk's model is equivalent to the SUE model by Danganzo (10).

As a computation procedure for the SNE model the successive average (SA) method proposed by Powell and Sheffi (11) is well known. Their method is based on the Blum's theorem. In section 4 we propose a procedure based on the Contraction Mapping theorem, in which the convex combinations parameter is determined so as to satisfy the contraction mapping for successively induced flow variables. This method is different from the SA method in the underlying theory, but is similar in that the in each iteration the update values are determined by the convex combinations of the previous link flows and the transformation values of the previous ones.

2. CONJUGATE CORRESPONDENCES BETWEEN THE SATISFACTION FUNCTION AND THE ENTROPY FUNCTION

2.1 Random Utility Model and the Satisfaction Function

We will here briefly review the random utility model and address the relationship between the expected maximum utility and choice probability formula, the properties of the expected maximum utility.

Let M denote the relevant set of alternatives for each individual and suppose that the utility of alternative $m \in M$ to individual i comprises two parts as is defined in eq.(1).

$$\tilde{U}_{mi} = v_{mi} + \tilde{y}_{mi} \tag{1}$$

where v_{mi} is the measurable(strict) utility of alternative m for individual i and y_{mi} represents the unmeasurable utility of alternative m for individual i or the error term and is the random variable. In order to simplify the notation, hereafter, the subscript denoting to individual will be dropped. The multinomial logit(MNL) model is derived by assuming that each $\tilde{y}_m$ is independently and identically distributed over the population for each individual according to the Gumbel distribution with the distribution function:

$$\Pr[\tilde{y} \leq y] = \exp[-\exp\{-\theta(y-\alpha)\}] \tag{2}$$

where α represents the mode; θ is the parameter associated with the variance σ^2 as is given by

$$\theta^2 = \pi^2/(6\sigma^2) \tag{3}$$

Then the choice probability p_m for alternative m is given by

$$P_m = \exp(\theta v_m)/\Sigma_j \exp(\theta v_j) \tag{4}$$

and the expected maximum utility is obtained as

$$E[\max_j . U_j] = \frac{1}{\theta} \ln \Sigma_j \exp(\theta v_j) + \alpha + \gamma/\theta \tag{5}$$

where γ is Euler's constsnt. Since the selection of the mode parameter is arbitrary and does not affect the choice probability, we can set it so that s(v) is represented as

$$s(v) = \frac{1}{\theta} \ln \Sigma_j \exp(\theta v_j) \tag{6}$$

Hereafter, we will call s the satisfaction(SF) function following Daganzo(10). The SF function has many interesting properties, but only those useful for later analysis will be listed below without proof except for the limiting property(property 4).

property 1 : Derivative (Wiliams(6))

$\partial s(v) / \partial v_j = P_j$

property 2 : Convexity (Daganzo(10))

s(v) is convex with respect to v

property 3 : Nondecreasing submodularity (Leonardi(8))

$s(L \cup \{j\}) - s(L) \geq s(T \cup \{j\}) - s(T) \geq 0$

for all L,T,j, $L \subseteq T$, $j \notin T$

property 4 : Limiting behavior

(i) $\lim_{\theta \to \infty} s(\mathbf{v}) = \max_j v_j$

(ii) $\lim_{\theta \to 0} s(v) = \Sigma v_j / |M| \quad j \in M$

(iii) $\lim_{v_j \to \infty} s(v) = v_j \quad j \in M$

(proof) see Appendix

2.2 Graphical Explanation of Conjugate Theory

To establish a Correspondence relationship between entropy and SF functions, let us introduce the conjugacy theory which has been developed in the convex analysis. The basic idea of conjugacy theory grows out of the fact that there are two ways of viewing a classical curve or surface like a conic, either as a locus of points or as an envelope of tangents. More generally, this implies that any continuous,convex function can be defined as a closed convex set of points and the closed convex set in the real values set is represented as the intersection of the closed half-spaces containing it. This property gives the general notion of duality. More formally, the definition of the convex conjugata function of f is as follows (12).

Definition. Let f be a convex function defined on a convex set F in a norm space X. The conjugate set F^* is defined as

$F^* = \{x^* \in X^* : \sup_{x \in F} [\langle x, x^* \rangle - f(x)] < \infty \}$

Then the function f^* defined by

$$f^* = \sup_{f^* \in F^*} [\langle x,x^* \rangle - f(x)] \qquad (7)$$

or

$$-f^* = \inf_{f^* \in F^*} [f(x) - \langle x,x^* \rangle] \qquad (7')$$

is said to be the conjugate convex function of f or simply the conjugate of f.

In the case of satisfaction function, the strict utility correspond to the point variable, while the choice probability is essentially the slope variable as is shown in the derivative property. So the theory of conjugacy suggests that the satisfaction function can be redefined via choice probabilities. A graphical depiction will help to clarify this theory more.

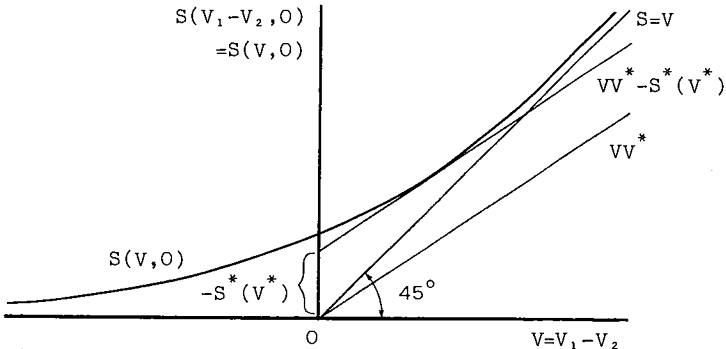
To get an intuitive understanding for a correspondence relation between the entropy and the SF function , we shall consider a simple binary choice situation that trip makers would be confronted with and observe how the epigraph of the satisfaction function s(v) is converted into another equivalent function defined by the choice probabilities.

The satisfaction function is illustrated in Figure 1 (a). In this figure the horizontal axis takes the value v which is the value of measurable utility of the first alternative v_1 subtracted by that of the second alternative v_2. From the property 4(iii),the curve s(v) may close asymptotically to the fourty-five degree line as the value of strict utility of the first alternative is tending to the infinity. Conversely, as v_1 becomes relatively smaller compared with v_2, s(v) approaches asymptotically to the horizontal axis.

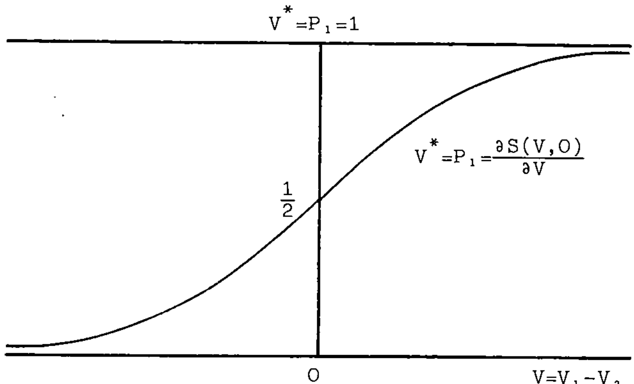
To find the conjugate function s(v), we must find $\sup\{\langle v^*,v \rangle - s(v)\}$. In the first figure, we want to maximize $\langle v^*,v \rangle - s(v)$ or equivalently minimize the $s(v) - \langle v^*,v \rangle$ for a given v^*. For this value v^*, $\langle v^*,v \rangle - r = k$ is a equation of a hyperplane through the intersect k. So as k varies, the solution(r,k) of the equation $\langle v^*,v \rangle - r = k$ describes a parallel closed hyperplane . The number $s^*(v^*)$ is the supremum of the value of k for which the hyperplane intersects the epigraph of s(v). Thus the hyperplane $\langle v^* \rangle - r = s^*(v^*)$ is a support hyperplane of the epigraph.

To minimize the difference between s(v) and $\langle v^*,v \rangle$,we move the hyperplane vertically until it supports the epigraph of s. Thus, we can see that the intersect of this hyperplane on the vertical axis would give us a negative conjugate function.

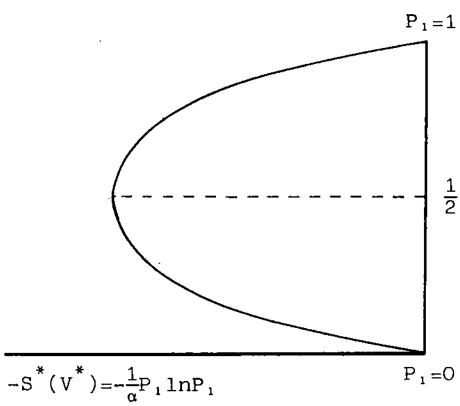
From the derivative property of the satisfaction function, the slope of the function implies the choice probability and furthermore it corresponds to the conjugate variable v^*. Figure 1 (b) depicts a choice probability distribution for the first alternative over the same horizontal axis as the upper figure. If v_1 equals v_2, the choice probability for each alternative is the same, one over two. If v_1 becomes smaller than v_2, the slope of s(v) becomes smaller, so the choice probability for the first alternatives becomes smaller. On the other hand, the greater v_1 becomes as compared with v_2, the



(a) Satisfaction Function



(b) Choice Probability Distribution



(c) Negative Conjugate of SF Function

Fig.1 Conjugacy between Satisfaction Function and Entropy greater increase of the frequency the first alternative is selected. So we get the curve depicted in Figure 1(b).

Next, if we draw the curve of $-s^*$ against the value of the vertical axis of Figure 1 (b), that is, choice probability, we can get the bottom figure. When we compare Figure 1(c) with Figure 1(a) and 1(b), the following relations are obvious. When choice probabilities for both alternatives are the same, according to Figure 1 (a) the intersect of the supporting hyperpane on the vertical axis has the greatest value. When the choice probability that the first alternative is selected becomes smaller, the intersect of the supporting hyperplane also becomes small, and finally equals zero at the point when the slope of s(v) is zero. On the other hand, when the choice probability for the first alternative approaches one, the intersept of the supporting hyperplane equals zero and we get Figure 1 (c). At first glance, we notice that this figure appears to be quite similar to an entropy curve. If we assume the satisfaction function has the log-sum form, then it can be assured that this graph indeed corresponds to the well-known entropy curve.

2.3 Conjugate Correspondences

To make the notion of conjugate correspondence between the entropy and the SF function precise, we begin by derivating the entropy formula from the SF function. Since the SF function is convex, the negative conjugate function is obtained by solving the following mathematical program defined on a strict utility space V which is assumed to be a nonempty set.

[DØ] For any given $v^* \in V^*$,

$$\min : \Phi(v)=s(v)-\langle v,v^*\rangle \quad , \text{ s.t. } v \in V \tag{8}$$

The optimal solution $\hat{v}$ for this program must satisfy the next equation.

$$v_j^*-\exp(\theta\hat{v}_j)/\Sigma_j\exp(\theta\hat{v}_j)=0 \qquad j \in M \tag{9}$$

$$\Sigma v^*_j = 1 \tag{10}$$

It is apparent that the conjugate variable v^*_j for the strict utility v_j corresponds to the choice probability P_j. Since the conjugate function $s^*(v^*)$ is given as the optimal value function defined by conjugate variables, by substituting v_j satisfying the relation (9) for the objective function in (8) , we can get the following entropy formula.

$$\Phi(\hat{v})=-\frac{1}{\theta}\,\Sigma_j v_j^* \ln v_j^* = -s^*(\mathbf{v}^*) \tag{11}$$

If we let the entropy function be H and the conjugate variable v_j^* be p_j, then the <u>First Conjugate Correspondence</u> can be expressed as :

$$\frac{1}{\theta}H(p) = s(\hat{v}) - \Sigma_j p_j \hat{v}_j \tag{12}$$

The first Correspondence shows that the strict utility consistent with given choice probability distribution should be determined by the problem [DO] and that even if the strict utility is determined by [DO] , the SF function includes the unmeasurable part which is not interpreted by the expectation of measurable utility $\langle p,v \rangle$, and is equivalent to an uncertainty measure, the entropy. These imply that [DO] is useful for estimating parameters included in the strict utility and that from the behavioral viewpoint the entropy (more precisely the entropy divided by the dispersion parameter) is interpreted as the expectation of unmeasurable utility (i.e., the expected surplus).

Next, let us consider the conjugate function of $s^*(v^*)$.-- namely, the conjugate function of the negative entropy which is the conjugate of the SF function. Since the negative entropy is a convex function, we can formulate the conjugate problem in a similar way to [DØ]. That is,

[PO] For given $p^* \varepsilon P^*$

$$\max.\ \Gamma(p)=\Sigma_j p_j p_j^* - s^*(p), \qquad \text{s.t.} \quad \Sigma p_j = 1 \tag{13}$$

The optimal solution $\hat{p}$ for this program has to satisfy the next logit formula.

$$\hat{p}_j=\exp(\theta p_j^*)/\ \Sigma_j \exp(\theta p_j^*) \tag{14}$$

That p_j^* corresponds to the strict utility v_j is apparent. Expressing the objective function by p_j which satisfies the equation (14), we get the conjugate function of the negative entropy, and can show that it is equivalent to the SF function. That is,

$$s^{**}(p^*)=s(v) \tag{15}$$

Thus, we have <u>the Second Conjugate Correspondence</u> :

$$s(v)= \Sigma_j \hat{p}_j v_j + \frac{1}{\theta} H(\hat{p}) \tag{16}$$

The second correspondence shows that the choice probability consistent with a given strict utility level should be given by the logit formula and that from the behavioral viewpoint the entropy maximization model (13) is interpreted as the process of maximizing the total expected utility comprising the expectations of measurable utility and unmeasurable utility.

3. FORMULATION OF THE STOCHASTIC USER EQUILIBRIA

3.1 Notations on Networks and Aggregate SF Measures

We begin by describing the basic notation for representing the route choice problem for each traveler. If I and J denote the relevant sets of origins and destinations, respectively, then the decision problem for each traveler from $i \varepsilon I$ to $j \varepsilon J$ is the choice of a route r from the relevant set R_{ij} of feasible routes from i to j so as to maximize his utility or minimize his disutility.

Let f_{ijr} denote the path flow on each route $r \varepsilon R_{ij}$, the vector f the resulting flow pattern on the network, and let q_{ij} denote the number of O-D trips from i to j . then, we define the feasible flow pattern F by

$$F=\{ f\geq 0 \mid \Sigma_r f_{ijr}=q_{ij}, \quad \text{for all } i \in I,\ j \in J\} \tag{17}$$

The network itself is represented by a direct graph that includes sets of consecutively numbered nodes N and links A. If we let the flow on link a be x_a and the cost on that link be c_a , then the relationship between flow and travel cost for link a is characterized by the link performance function , $c_a=c_a(x_a)$, and the travel cost on a particular path is expressed as:

$$c_{ijr}= \Sigma_a \delta_{ij,ar} c_a(x_a) \tag{18}$$

where $\delta_{ij,ar}=1$ if link a is a part of route r connecting O-D pair i-j, and $\delta_{ij,ar}=0$, otherwise. Using the same indicator variable, the link flow can be expressed as a function of the path flow, that is,

$$x_a= \Sigma_i \Sigma_j \Sigma_r \delta_{ij,ar} f_{ijr}, \qquad a \in A \tag{19}$$

By considering relation (19), an alternative expression for the route cost is also possible by using the flow pattern, f,

$$c_{ijr}(f)= \Sigma_a \delta_{ij,ar}\, c_a(f) \tag{20}$$

The above expression is followed hereafter.

In order to describe the route-choice decision for a traveler entering the network at i destined for j, it may be appropriate that we define the strict utility value that governs the route choice for each traveler among the feasible route set R_{ij} by

$$v_{ijr}=-c_{ijr} \tag{21}$$

Then, the satisfaction measure for each trip between O-D pair ij can be obtained in a manner identical to the previous section as:

$$S_{ij}(c) = \frac{1}{\theta} \ln \Sigma_r \exp(-\theta c_{ijr}) \tag{22}$$

It should be noted that S_{ij} is defined in terms of route costs in spite of the strict utility and is significant within the finite domain of the dispersion parameter. Defining the overall measure of satisfaction on the network as the sum of the satisfaction for individual trips, the law of large numbers ensures that the overall satisfaction S is given by

$$S(c) = \Sigma_i \Sigma_j q_{ij} S_{ij}(c) \tag{23}$$

We will call S the aggregate Satisfaction (ASF) measure.

From the limiting behavior of SF function (Property 4 indicated in section 2.1), at $\theta=0$, the ASF function reduces to the linear SF function defined by the following equation.

$$S_{ijo} = - \Sigma_r c_{ijr}/|M| = -\bar{c}_{ij} \tag{24}$$

This implies that since alternative routes are no longer distinguishable regarding to utility of choice, route-choice behavior might be maximally dispersed. On the other hand, at the opposite extreme, namely, as approaches to infinity, the SF function describes the strict utility maximizer. Within the present context, this means that all travelers might choose the cheapest route with perfect information ($\sigma^2 \to 0$ as $\theta \to \infty$). More formally, this situation is reflected in the following limiting form of the SF function.

$$S_{ij\infty} = \max_r[-c_{ijr}] = \min_r[c_{ijr}] \tag{25}$$

It is appropriate to comment that eq.(25) must hold regardless of flow level because that route costs are a function of flow levels and thus describes behavior of a pure cost- minimizer who always seeks to find the cheapest route.

3.2 The Conjugates of the ASF Functions

As is indicated in 3.1. these are three different forms of SF functions. Our central results here show that the stochastic user equilibrium model consistent with the random utility framework is naturally derived by taking the conjugate of each of these different forms of SF functions and by embedding these conjugate functions into a unified conjugat function.

We begin by constructing the conjugate function of an aggregate SF measure S with a positive, finite dispersion parameter. By definition, the negative conjugate function for the ASF measure defined as a function of cost variables is obtained by solving the next minimization problem.

[D1] For any given $c^* \in C^*$,

$$\min. \quad \Phi(c) = S(c) - \Sigma_i \Sigma_j \Sigma_r c^*_{ijr} c_{ijr} \tag{26}$$

where c^*_{ijr} is the conjugate variable corresponding to the path flow. The optimal solution of this program must satisfy the following logit formula.

$$-c^*_{ijr} = q_{ij} \exp(-\theta c_{ijr}) / \Sigma_r \exp(-\theta c_{ijr}) \tag{27}$$

For clarity of notation the sign denoting the optimal is omitted from the above expression. This practice is followed hereafter in this paper. From eq.(27) it is apparent that the negative conjugate variables $(-c^*_{ijr})$ correspond to the path flow f_{ijr} and satisfy the feasible flow condition

(17). As a result, program [D1] is expressible as:

[D1] For any $f \in F$,

$$\min: \quad \Phi(\mathbf{c}) = S(\mathbf{c}) - \Sigma f_a c_a(f) \tag{28}$$

Since the conjugate function is given as the extremal value function of $\Phi(c)$, the conjugate function S^* is given by

$$-S^*(f) = -\frac{1}{\theta} \Sigma_i \Sigma_j \Sigma_r f_{ijr} \ln f_{ijr} + \frac{1}{\theta} \Sigma_i \Sigma_j q_{ij} \ln q_{ij} \tag{29}$$

Defying the network entropy associated with the route choice as:

$$H(f) = -\frac{1}{\theta} \Sigma_i \Sigma_j \Sigma_r f_{ijr} \ln f_{ijr} \tag{30}$$

then the conjugate of the ASF function can be expressed by the next equation by using H(f) and making the secoi.d term constant.

$$\overset{*}{S}(f) = -\frac{1}{\theta} H(\mathbf{f}) + K, \quad K = -\frac{1}{\theta} \Sigma_i \Sigma_j q_{ij} \ln q_{ij} \tag{31}$$

Next,we consider getting the conjugate function of S^*. This can be done by solving the following maximization program.

[P 1] For any given $f^* \varepsilon F^*$

$$\max. \quad \Gamma(f) = \Sigma_i \Sigma_j \Sigma_r f^*_{ijr} f_{ijr} + \frac{1}{\theta} H(f) - K \tag{32}$$

The optimal solution of this program has to satisfy the following logit formuia.

$$f_{ijr} = q_{ij} \exp(\theta f^*_{ijr}) / \Sigma_r \exp(\theta f^*_{ijr}) \tag{33}$$

The above equation ensures that the conjugate variables f^*_{ijr} of path flow correspond to the negative path costs, and program [P 2] is expressible as:

[P 2] For any given $c \varepsilon C$,

$$\max. \quad \Gamma(f) = \frac{1}{\theta} H(f) - \Sigma_a x_a C_a, \qquad \text{s.t. } f \varepsilon F \tag{34}$$

It can be easily shown that the extreme value function of this program is equivalent to the ASF measure S. We notice at first glance that the Problem [P2] is equivalent to the entropy model proposed by Wilson(1). But it should be noted that the model presented here neither requires the total-cost constraint ,nor accounts for θ as being the Langrange multiplier. Since the model is directly derived from the SF function, θ is interpreted as the dispersion parameter which is deeply concerned with the traveler's decision on their route choices.

Now we proceed to the analysis for the case when the dispersion parameter θ asymptotically approaches an infinite value. The satisfaction measure for this case is not continuous as is shown in eq.(25). The conjugate problem is given by [P 3] by making $\lambda_{ij} = \min c_{ijr}$.

[P 3] For any given $c^* \varepsilon C^*$,

$$\min_{c} \quad \Phi(c) = \Sigma_i \Sigma_j \Sigma_r c_{ijr} c^*_{ijr} - \Sigma_i \Sigma_j \Sigma_r f_{ijr} \lambda_{ij} \tag{35}$$

The solution for this program has to satisfy the following relations.

$$\begin{aligned} &\text{if} \quad c_{ijr} = \lambda_{ij}, \quad \text{then} \quad c^*_{ijr} = f_{ijr} \\ &\text{if} \quad c_{ijr} = \lambda_{ij}, \quad \text{then} \quad c^*_{ijr} = \emptyset \end{aligned} \tag{36}$$

Again, the conjugate variables correspond to the path flow, consequently, we can see that program [P 3] is rewritten as the following form.

[P 3'] For any given $f \varepsilon F$,

$$\min. \quad \Phi(c) = \Sigma_a x_a c_a - \Sigma_i \Sigma_j \Sigma_r f_{ijr} \lambda_{ij} \tag{37}$$

However,in this case the extremal value function is not explicitly derived from solution (36) because that λ_{ij} is not explicitly expressed by the conjugate variables (i.e. the path flow or link flow variables). However, relations (36) show the behaviour of pure cost-minimizer, and as is shown by Smith(9) and Miyagi(13),it can be proved that if each traveler seeks to find the minimum cost route at any flow level,then their behavior can be described by the well-known minimization problem of cumulative cost function, as is defined by

$$\min. \Sigma_a \int_o^{x_a} c_a(x)dx \tag{38}$$

Furthermore, the cumulative cost function is a function of the path flow, that is the conjugate variable. Accordingly, we can expect that S^*_∞ ($\mathbf{f}$) amounts to the sum of cumulative cost function with the flow satisfying relation (36),i.e., the user equilibrium flow.

Consequently, the conjugate of S^*_∞ (f) is defined as below because of the cumulative cost function being convex.

[P 4] For $f^* \epsilon F^*$

$$\max. \quad \Gamma(f) = \Sigma_i \Sigma_j \Sigma_r f_{ijr} f^*_{ijr} - \Sigma_a \int_o^{x_a} c_a(x)dx \tag{39}$$

This programe yields the conjugate relation between f^*_{ijr} and a route cost btween O-D pair ij.

$$f^*_{ijr} = \Sigma_a \delta_{ij,ar} c_a = c_{ijr} \tag{40}$$

As a result, the conjugate of $S^*_\infty(\mathbf{f})$ is obtained as:

$$S_\infty(\mathbf{c}) = \Sigma_a \int_{c_{a,min}}^{c_a} c_a^{-1}(t)dt \tag{41}$$

3.3 Formulation of the SUE model by the ASF measure maximization approach

We combine two different forms of ASF measures $S(\mathbf{c})$, $S_\infty(\mathbf{c})$, into a more general class of ASF measure, $W'(\mathbf{c})$. Let us suppose that this can be done by adding these measures.

$$\begin{aligned} W'(\mathbf{c}) &= S(\mathbf{c}) + S_\infty(\mathbf{c}) \\ &= \frac{1}{\theta} \ln \Sigma_r \exp(-\theta c_{ijr}) + \Sigma_a \int_{c_{a,min}}^{c_a} c_a^{-1}(t)dt \end{aligned} \tag{42}$$

$W'(\mathbf{c})$ is a convex function, and includes the first term representing the disutility which is a negative value of the percieved travel cost and the second term representing the travel cost. This function generates the following the surplus maximization problem without no explicit constraints.

$$\max. \quad W(\mathbf{c}) = -\frac{1}{\theta} \ln \Sigma_r \exp(-\theta c_{ijr}) - \Sigma_a \int_{c_{a,min}}^{c_a} c_a^{-1}(t)dt \tag{43}$$

For taking the negative of the disutility for travel, it represents the utility and the value subtracted by the travel cost means the surplus for travel. It is clear from the functional form in eq.(43) that the maximization problem here includes the case of $\theta = 0$, and is equivalent to the SUE program by Daganzo(10). The conjugate of $W(\mathbf{c})$ is straightforward from the development shown in 3.2, and is given by

$$Z(\mathbf{f})=\frac{1}{\theta}\Sigma_i\Sigma_j\Sigma_r f_{ijr}\ln(f_{ijr}/q_{ij})+\Sigma_a\int_o^{x_a} c_a(x)dx \tag{44}$$

$Z(\mathbf{f})$ is also convex. The first term of $Z(\mathbf{f})$ corresponds to $S(\mathbf{c})$, the disutility, and the second corresponds to $S_\infty(c)$, the travel cost. Again, we have the following surplus maximization problem with the feasible flow constraints.

$$\max.\ Z(\mathbf{f})=-\frac{1}{\theta}\Sigma_i\Sigma_j\Sigma_r\ f_{ijr}\ln(f_{ijr}/q_{ijr})-\Sigma_a\int_o^{x_a} c_a(x)dx,\ \text{s.t. } f\in F \tag{45}$$

This problem is essentially the same as the maximization of W (c): it is no more than the problem represented by flow variables instead of cost variables and includes the case of the dispersion parameter being zero. Furthermore this program is indeed the SUE model by Fisk (5). Thus two SUE models, max.W(c), min.Z(f), which are quite different at first glance, have a strict correspondence within the framework of the conjugate theory and are consistent with the random utility theory.

4. ALGORITHM

4.1 General Discription of Algorithm

Consider the following equivalent minimization problem to the original maximization program (45).

$$\min.Z(f)=\frac{1}{\theta}\Sigma_i\Sigma_j\Sigma_k f_{ijk}\ln f_{ijk}+\Sigma_a\int_o^{x_a} c_a(x)dx\ ,\quad \text{s.t. } f\in F \tag{46}$$

It is evident from the form of $Z(\mathbf{f})$ that this problem is a strict convex programe and yields a unique solution defined by

$$f_{ijk}=q_{ij}\exp(-\theta c_{ijk}(\mathbf{f}))/\Sigma_r\exp(-\theta c_{ijr}(\mathbf{f})) \tag{47}$$

Each component f_{ijr} of the resulting flow pattern f is always positive, and is a decreasing function of its corresponding cost $c_{ijr}(\mathbf{f})$ as long as the dispersion parameter θ is finite. Furthermore, as is shown in the previous section ,as $\theta\to\infty$, the solution $\mathbf{f}$ continuosly approaches to the Wardrop equilibrium solution f_e. However, the typical iterative procedures of finding the equilibrium flow such as the convex combinations method based on the Frank-Wolfe decomposition principle can not be easily applied to this program according to the reasons suggested by Sheffi(14).

In spite of this difficulty, equation (47) shows the possibility of existance of solution procedures which guarantee to converge to the unique solution since the equation characterizes f as the fixed point of continuous transformation from F into itself. Even if we rewrite eq.(47) in terms of the link flow as is shown as :

$$x_a=\Sigma_i\Sigma_j\Sigma_k q_{ij}\exp[-\theta(\Sigma_a\delta_{ij,ak}c_a(x_a))]/\Sigma_k\exp[-\theta(\Sigma_a\delta_{ij,ak}c_a(x_a))] \tag{48}$$

the essential feature of having the fixed point is unchanged.

In its general form the classical method of successive approximation applies to an equation of the form $x=T(x)$ where T is the transformation from X on which x is defined into itself. The Contraction mapping theorem ensures the existance of the unique vector $\mathbf{x}$ satisfying $x=T(x)$ and that $\mathbf{x}$ can be obtained by the method of successive approximation starting from an arbitary initial vector $\mathbf{x}_o \in X$(15). The transformation T is said to be contraction mapping if there is an arbitary k , $o \leq k<1$, such that

$$\|T(y)-T(x)\| < k \|y-x\| \quad \text{for all } x,y \in X \tag{49}$$

The algorithm described in this section is based on the contraction mapping.

The outline of the algorithm is as follows. Select an initial feasible solution $\mathbf{x}_o$, which is the link flow vector calculated by the Dial method with prespecified link costs. Let y denote the link flow vector generated by the transformation T which is carried out by the Dial algorithm(16) with costs updated by the previous link flow x_{n-1}. Thus we define the transformation T as being constructed by the Dial algorithm. Then if the next inequality holds for x_{n-1} and the candidate of the next solution y,

$$\| T(y)-T(x) \| \leq k \| x_n-x_{n-1}\| \tag{5Ø}$$

we determine the updated solution as $\mathbf{x}_n=y$. If this is not so, that is, for a candidate of the next solution, $x_n=\mathbf{y}$, the following inequality holds,

$$\| T(\mathbf{y})-T(x_{n-1}) \| > k \| x_n-x_{n-1}\| \tag{51}$$

then by using a step size parameter β satisfying

$$\beta \| T(y)-T(x) \| \leq k \| x_n-x_{n-1}\| \tag{52}$$

the updated solution is determined through the convex combinations

$$x_n=(1-\beta)x_{n-1}+ \beta y \tag{53}$$

By this operation, the following inequality relation holds.

$$\| T(y)-T(x_{n-1}) \| \leq k \| y-x_{n-1}\| \tag{54}$$

It should be noted that the above inequality relation for ensuring contraction mapping does not always hold for the improved solution x_n, except for the choice function being linear. Thus this method does not guarantee that the updated solution always decreases its norm compared with the previous one. However,as will be shown in 4.2, this method is expected to converge.

4.2 Convergence of Algorithm

To show how the algorithm converge, we begin by assuming the case that contraction mapping does not hold among the initial solution x_o, its transformation $\mathbf{y}_o$ and the transformation of $\mathbf{y}_o$, $T(\mathbf{y}_o)$. Suppose that for $T:\mathbf{x}_o \to y_o$ and $T:x_1 \to \mathbf{y}_1$

$$\|T(\mathbf{y}_o)-T(\mathbf{x}_o)\| > k \|\mathbf{x}_1-\mathbf{x}_o\| , Ø\leq k< 1$$

holds, in which $\mathbf{x}_1=\mathbf{y}_o$ is assumed, then, by using the step size parameter β satisfying

$$\beta \| T(y_o)-T(x_o) \| \leq k \| x_1-x_o\|$$

the next point is recalculated as:

$$x_1=(1-\beta)x_o+\beta y_o$$

This operation gives the inequality relation

$$\| T(y_o)-T(x_o)\| \leq k\| y_o-x_o\|$$

For the simplicity of interpretation, suppose that for the succeeding steps the inverse inequality always holds between the previous solution x_{n-1}, a candidate of the desired point y_n generated by the transformation $T(x_{n-1})$ such that

$$\| y_n-T(x_{n-1})\| > k\| x_n-x_{n-1}\|$$

By the same operation as the first step, the step size parameter β is introduced and the following relation is assumed to hold:

$$\beta\| y_n-T(x_{n-1})\| \leq k\| y_{n-1}-x_{n-1}\|$$

where $x_n=(1-\beta)x_{n-1}+\beta y_{n-1}$. Since the step size parameter determined at each iteration is always smaller than one, we set it constant as being a value less than one. Then we have the following sequence of inequalities.

For the first step

$$\| T(y_o)-T(x_o)\| \leq k\| y_o-x_o\|$$

For the second step

$$\begin{aligned}\| T(y_1)-T(x_1)\| &\leq k\| y_1-x_1\|\\ &\leq k\| y_1-y_o+(1-\beta)(y_o-x_o)\|\\ &\leq k\| y_1-y_o\| +k(1-\beta)\| y_o-x_o\|\\ &= \{k^2+k(1-\beta)\}\| y_o-x_o\|\end{aligned}$$

Since similar inequality relations hold for succeeding steps, for the nth step the inequality relation as below holds.

$$\begin{aligned}\| T(y_{n-1})-T(x_{n-1})\| &< k\| y_{n-1}-x_{n-1}\|\\ &< k\| y_{n-1}-y_{n-2}+(1-\beta)(y_{n-2}-x_{n-2})\|\\ &< k\| y_{n-1}-y_{n-2}\|+k(1-\beta)\| y_{n-2}-x_{n-2}\|\\ &= k(1+k-\beta)^{n-1}\| y_o-x_o\| \qquad (55)\end{aligned}$$

Consequently, suppose $n\to\infty$, then it is expected that if the constant term in the right-hand side of eq.(55) approaches zero such that $\| y_n-T(x_{n-1})\|$, $\| x_n-x_{n-1}\|$ and $\| y_{n-1}-x_{n-1}\|$ approach zero. As the results,

$T(x_{n-1})=y_{n-1}$, $T(x_n)=x_n$ hold and we have the fixed point. However, if $k>\beta$ holds at each iteration, the convergency of this algorithm will fail because the constant term becomes greater than one. For this method to converge, it requires less value of k than that of β. It is difficult to find such a constant value, however, if we change the k-value according to determined by eq.(52) at each iteration , the algorithm steps are expected to almost exactly converge.

5. CONCLUSION

The main results developed in this paper include (1) establishing the correspondence relation between the entropy maximization model and the random utility model, (2) deriving the user equilibrium model within the context of conjugate correspondence developed here and (3) proposing a computation procedure based on the contraction mapping method. All of these results are expected to serve for making possible further research. Hence it is appropriate to touch on a few of the remaining problems of each result mentioned above for further research.

The correspondence relation has been developed here under the assumption that the error term of random utility is distributed according to the Gumbel distribution. These correspondence relations may hold under the more general form of random utility distribution such as the general extreme distribution proposed by McFadden, but which should be examined.

The user equilibrium model derived from the conjugate correspondence has been formulated by bearing a single transportation model in mind. Such a structual assumption serves to motivate the basic theory in terms of its simplest decision problem such as route-choice behavior. The fundamental theory does not depend on this assumption and it is directly extendable to a class of multimodal network equilibrium as is derived by the author(17). However, since in the previous formulation on the multimodal network equilibrium, a conveinient technique for deriving the user equilibrium is taken, the extension should be done in considering this point. The remaining problem is how the joint congestion effect created by the interaction of each mode should be incorporated into the framework of conjugacy theory.

Finally, we address the computation procedure presented here. Problems arise in selection of value of k and whether the efficientpaths in the Dial algorithm should be fixed or not. The first problem may be avoidable if one sets it as a variable value according to the -value at each iteration. However, the efficiency of convergence may decrease. Another way of selecting the k-value is to set it appropriately prespecified constant such that when the value of the dispersion parameter is large, one makes the k-value small and for the opposite, one sets the k-value larger. Such a selection method has no theoretical background, and is no more than experimental. This method has similar characteristics to the method of successive average by Powell and Sheffi, but has some cases of rapid convergence. As for the second problem, from the view point of convergency it is preferable to fix the efficient paths. However, for the large value of the dispersion parameter, the equilibrium flow pattern generated by this assumption does not correspond to the usual Wardrop equilibria. To get the equilibrium solution consistent with the Wardrop equilibria, one should assume the efficient paths are unfixed, but the efficiency of convergence may remarkably decrease.

REFERENCES

(1) Wilson, A.G.,(1967), "A Statisical Theory of Spatial Distribution Models", Transpn. Res. Vol. 1, pp. 253-269.

(2) Sasaki,T.,(1966), "Estimation of O-D Traffic through the Transitation Probability Method: Entropy Approach",Road, 1966-8, pp.70-77.

(3) Domencich, T., and McFadden, D., (1975), "Urban Travel Demand : A Behavioral Analysis", North-Holland, Amsterdam.

(4) Charles River Associates, (1972), "A Disaggregate Behavioral Models of Urban Travel Demand", US Department of transportation, Washington DC.

(5) Fisk,C., (1980), "Some Developements in Equilibrium Traffic Assignment", Transpn. Res. Vol.14B, pp. 243-256.

(6) Williams, H. C. W. L., (1977), "On the Formulation of Travel Demand

Models and Economic Evalution Measures of User Benefit", Envnmt. Plan. A.9, pp. 285-344.
(7) Anas, A.,(1983), "Discrete Choice Theory, Imformation Theory and the Multinominal Logit and Gravity Models", Transpn. Res. B, Vol.17B, No.1, pp. 13-23.
(8) Leonardi,G., (1983), "The Use of Ramdom-Utility Theory in Building Location-Allocation Models", Location Analysis of Public Facilities (J. F. Thisse and H.G.Zoller,edts.), North-Holland, Amsterdam, pp.357-383.
(9) Smith, T.E., (1983), "A Cost-Efficiency Theory of Dispersed Network Equilibria", Working Paper No.78, Dept. of Reasional Science, Univ. of Pennsylvania.
(10) Daganzo, C., (1979), "Multiominal Probit : The Theory and Its Application to Demand Forecasting", Academic Press, London .
(11) Powell, W. B., and Y. Sheffi, (1982), "The Convergence of Equilibrium Algorithms with Predetermined Step Sizes", Transpn. Sci., Vol. 16, No.1, pp. 45-55.
(12) Rockafellar, R. T., (1970), "Convex Analysis", Princeton University Press, USA.
(13) Miyagi,T.,(1985), "Traffic Equilibrium Models: Theories and Algorithms", Infrastructure Planning Reviews, No.2, pp.13-28. (in Japanese)
(14) Sheffi, Y., (1985), "Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods", Printice-Hall, Inc., New Jersey.
(15) Luenberger, D. G., (1969), "Optimization by Vector Space Method", John Wiley & Sons, Inc., New York.
(16) Dial, R. B., (1971), "A Probablistic Multipath Traffic Assignment Algorithm which Obviates Path Enumeration", Transpn. Res., Vol.5, No.2, pp.83-111.
(17) Miyagi,T.,(1984),"A Conjugate Dual Approach to Travel Demand Modeling", Proc. of the 12th PTRC Annual Meeting, Vol.257,pp.323-336.

APPENDIX: Proof of Property (4) of s(v)

(i) $\lim_{\theta\to\infty} s(v)=\max_{\cdot j} v_j$

Let us suppose that $v_j=\max_{\cdot k} v_k$ and $|M| \geq 1$. We want to show that $\lim_{\theta\to\infty} s(v)=v_j$. First, from the nondecreasing submodularity,

$$\frac{1}{\theta}\ln[\exp(\theta v_j)] \leq \frac{1}{\theta}\ln[\exp(\theta v_j+\theta v_1) \leq$$
$$\frac{1}{\theta}\ln[\exp(\theta v_j+\theta v_1+\theta v_2)] <\ldots< \frac{1}{\theta}\ln\Sigma_k\exp(\theta v_k) = s(v)$$

Thus we have $v_j< s(\cdot)$. From the assumption together with the nondecreasing submodularity,

$$\frac{1}{\theta}\ln\Sigma_k \exp(\theta v_k) \leq \frac{1}{\theta}\ln[|M| \exp(\theta v_j)] = \ln|M|/\theta + v_j$$

Consquently, the following inequality holds.

$$v_j \leq \lim_{\theta\to\infty} s(v) \leq \lim_{\theta\to\infty} \{ \ln|M|/\theta + v_j\} = v_j$$

Thus we have

$$\lim_{\theta\to\infty} s(v)=v_j=\max_k v_k$$

(ii) $\lim_{\theta \to o} s(v) = \Sigma_k v_k / |M|$

Consider the behavior of the function

$$h(\theta)=\{\theta s(v)-\ln \Sigma_j \exp(\theta v_j)\}/\theta$$

as $\theta\to\infty$. At $\theta=\emptyset$, the value of function is indeterminate, but $\lim_{\theta\to o} h(\theta)$ can be calculated by L'Hopital's rule:

$$\lim_{\theta \to o} h(\theta)=s(v)-\Sigma_j v_j \exp(\theta v_j)/\Sigma_j \exp(\theta v_j)$$
$$=s(v)-\Sigma_j v_j p_j$$

Thus, evaluating this expression at $\theta=\emptyset$, we have

$$\lim_{\theta \to o} h(\theta)=s(v)-\Sigma_j v_j / |M|$$

(iii) $\lim_{V_j\to\infty} s(v)=v_j$

First of all, we rewrite s(v) as the form

$$s(v)=\frac{1}{\theta}\ln[\exp(\theta v_j)\{1+\Sigma_{k\neq j}\exp(\theta(v_k-v_j)]\}$$
$$=v_j+\frac{1}{\theta}\ln \Sigma_{k\neq j}\exp[\theta(v_k-v_j)]$$

Accordingly, the relation desired is straightforward as :

$$\lim_{V_j\to\infty} s(v)=\lim_{V_j\to\infty} v_j$$