A Neuro-Fuzzy Approach to Data Analysis of Pairwise Comparisons

H. Ichihashi
University of Osaka Prefecture, Osaka, Japan

I. B. Türksen
Institute of Technology, Yokohama, Japan.
On leave from University of Toronto, Toronto, Ontario, Canada.

ABSTRACT

Artificial neural networks provide iterative on-line learning schemes for modeling non-linear systems. An iterative learning algorithm in fuzzy models, which is called neuro-fuzzy, has been recently developed within the framework of fuzzy modeling in the sense of M. Sugeno. In this paper, using neuro-fuzzy approach, two quantification methods of pairwise comparisons are presented in order to derive the associated weights of different objects. The proposed methods can be applied even in the case of incomplete pairwise comparisons. A simplified fuzzy reasoning model is obtained in the form of Gaussian radial basis functions.

The psychological sensation responses of human beings to minute vibrations are analyzed by the newly proposed neuro-fuzzy approach. The proposed approach is compared with Guttman's method and Saaty's analytic hierarchy process (AHP). In our two neuro-fuzzy approaches, psychological values are obtained with the interval and the ratio scale properties. They are represented by smooth functions of class $C^\infty$.

KEYWORDS: Hierarchical fuzzy model, neural network, pairwise comparison, Guttman's method, AHP, psychological sensation

1. INTRODUCTION

Within the framework of fuzzy modeling in the sense of Takagi and Sugeno [1] and Sugeno and Kang [2] for approximating non-linear functions, a learning algorithm based on the steepest descent method has been

Address correspondence to I. B. Türksen, Department of Industrial Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada.

Received June 5, 1991; accepted March 12, 1993.
proposed (e.g., Ichihashi and Watanabe [3–6]). It is called a neuro-fuzzy approach since the generalized delta rule in artificial neural networks is adopted as the learning rule. A simplified fuzzy model using Gaussian membership functions (Ichihashi [7–9], Ashida and Ichihashi [10]) is of the Gaussian radial basis functions (GRBF) type (Broomhead and Lowe [11], Poggio and Girosi [12]); and learning by the steepest descend method does not need the error back propagation rule (Rumelhart et al. [13]). The basic concern in this approach is the generation of a simplified fuzzy model in the form of Gaussian radial basis functions from numerical data pairs.

In Section 2, we briefly describe a simplified fuzzy model using Gaussian membership functions and derive a learning algorithm in hierarchical networks. Two quantification methods of pairwise comparisons using the neuro-fuzzy approach are presented to derive the associated weights of different objects. The proposed methods can even be applied in the case of incomplete pairwise comparisons.

The fuzzy quantification method was proposed by Watada and Tanaka [14] where the quantification theory for evaluating 0 or 1 data categories is extended to real numbers in interval [0, 1]. In a similar way the Guttman’s quantification method [15] is extended to the fuzzy quantification method in Section 3. In Section 4 eigenvector of pairwise comparison matrix in Saaty's analytic hierarchy process (AHP) [16] is briefly surveyed, which is also widely used as a ratio scale of weights of different objects. AHP is important as it offers a tool for eliciting membership functions (Saaty [17]).

The evaluation of psychological responses of human beings to vibrations by the category judgment method was reported by Maeda et al. [18, 19]. In Section 5 the psychological sensation responses of human beings to minute vibrations are analyzed by Guttman’s method, Saaty’s AHP, and the newly proposed neuro-fuzzy model. In our two neuro-fuzzy approaches, psychological values that are represented by smooth functions of class $C^\infty$ are obtained with the property of the interval and ratio scales.

2. NEURO-FUZZY MODEL

Let $A_{ik}$ denote the membership function of the $k$-th fuzzy rule in the domain of the $i$-th input variable $x_i$. The $k$-th rule is written as:

$$
\text{If } x_1 \text{ is } A_{1k} \text{ and } x_2 \text{ is } A_{2k} \ldots \text{ and } x_n \text{ is } A_{nk}, \text{ then } y = w_k.
$$

The conclusion part of the fuzzy reasoning rule that infers output $y$ is simplified as a real number $w_k$ as the Sugeno type “Singleton Consequent” model [1, 2]. The compatibility degree of the premise part of the $k$-th fuzzy rule for an observed system state $x = (x_i | i = 1, \ldots, n)$ is computed with
the algebraic product operator of the compatibility degrees, $A_{ik}(x_i)$'s, as
\[
\mu_k = \prod_i A_{ik}(x_i)
\]  
(1)

The graph of $\mu_k$'s ($k = 1, \ldots, 9$) is shown in Figure 1. Each unimodal function corresponds to a fuzzy set in the premise part of the rule. Since the input space is divided fuzzily as in Figure 1, it is called a fuzzy partition. And, the final output $y$ for an observed system state $x = (x_i | i = 1, \ldots, n)$ is written as
\[
y = \sum_k \mu_k \cdot w_k
\]  
(2)

As it should be observed that minimum operators in conventional fuzzy reasoning method are replaced by product operators in Eq. (1). This model is called simplified fuzzy model (Ichihashi and Watanabe [3, 4]).

In general, Gaussian membership function is defined by
\[
A_{ik}(x_i) = \exp\left(-\frac{(x_i - a_{ik})^2}{b_{ik}}\right)
\]
for any $x_i$, where the parameters $a_{ik}$ and $b_{ik}$ ($i = 1, \ldots, n$) are given for each $k$ and are changed in the training procedure. $\mu_k$'s are products of Gaussian functions each of which is a function of only one component of vector $x$. Since exponential functions can be factored, $\mu_k$ have the form of Gaussian functions $F(||x - a||)$; and $y$ in Eq. (2) is its linear combination, i.e., a Gaussian radial basis function (GRBF) [11, 12]. The sum of $\mu_k$'s in the linear combination of Eq. (2) does not equal unity; and $y$ does not have the same scale as the $w_k$. Though this is important in fuzzy modeling by linguistic descriptions, the normalization by the sum of $\mu_k$ (i.e., the

![Figure 1. A fuzzy partition of two dimensional input space. ($i = 1, 2$; and $k = 1, \ldots, 9$).](image)
centroid defuzzification method) in conventional fuzzy reasoning is omitted for simplicity in the fuzzy modeling based on supervised learning.

2.1. Learning Rule and Hierarchical Network

When the number of input variables increases, the number of parameters \((w_{ik}, a_{ik}, b_{ik})\) in the simplified fuzzy model increases rapidly. To reduce the number, a network of fuzzy models has been proposed in [7–9]. An example of a hierarchical fuzzy model in Figure 2 has two layers with four inputs and one output.

The sub-models (i.e., the nodes in the network) illustrated by the rectangles are the simplified fuzzy models.

Let \(k\)-th rule in the \(m\)-th model be

\[
\text{If } x_i^m \text{ is } A_{1k}^m \text{ and } x_i^m \text{ is } A_{2k}^m, \text{ then } y \text{ is } w_k^m.
\]

\(x_i^m\) denotes the \(i\)-th input variable of the \(m\)-th model. \(A_{ik}^m\) denotes the compatibility degree of the \(i\)-th input of the \(k\)-th fuzzy rule in the \(m\)-th model. \(w_k^m\) denotes the real singleton number in the conclusion part of the \(k\)-th rule in the \(m\)-th model. Let \(\mu_k^m\) be the compatibility degree of the \(k\)-th rule of the \(m\)-th model, i.e., a membership function of two variables (Figure 1).

\[
y^m = \sum_k \mu_k^m \cdot w_k^m = \sum_k \prod_i A_{ik}^m(x_i^m) \cdot w_k^m
\]

for \(m = 1, 2, 3\), with \(x_1^3 = y^1\) and \(x_2^3 = y^2\). In the case of iterative training procedure (i.e., the delta rule [13]) for function approximation, the objec-

![Figure 2. A hierarchical fuzzy model with four input variables. (\(m = 1, 2, 3; i = 1, 2;\) and \(k = 1, \ldots, 9\).)
tive function is defined as

$$E_r = \frac{1}{2} (y^3 - y^*)^2$$

where \(y^*\) is the desired output. \(E_r\) is the square error corresponding to the \(r\)-th input pattern. Let \(\delta^3 = y^3 - y^*\). If the hierarchical fuzzy model is as in Figure 2 and the membership function is a Gaussian type, then the error back propagation rule [13] can be written as:

$$\delta^m = -2 \cdot \left\{ \delta^3 \sum_k \mu_k^3 \cdot \frac{(x_m^3 - a_{mk}^3)}{b_{mk}^3} \cdot w_k^3 \right\}$$

for \(m = 1, 2\). This rule is only for the hierarchical model in Figure 2 where \(a_{i,k}^m\) and \(b_{i,k}^m\) are \(a_{i,k}\) and \(b_{i,k}\) in the \(m\)-th model, respectively. It should be noted that the back-propagation errors at the \(m\)-th model \((\delta^m, m = 1, 2)\) are computed from \(m\)-th input of the third model \((x_3^3)\). Hence, to denote the \(m\)-th input variable, the subscript \(m\) is used instead of \(i\) such as \(x_m^3\), \(a_{mk}^3\), and \(b_{mk}^3\).

The training procedures based on the steepest descent method for changing \(w\) is

$$w_{k}^{m\text{NEW}} = w_{k}^{m\text{OLD}} - \tau \mu_k^m \cdot \delta^m$$

for \(m = 1, 2, 3\), where \(\tau\) denotes the learning constant. And, for \(a_{ik}^m\) and \(b_{ik}^m\) in the case of Gaussian membership functions, in a similar manner we have:

$$a_{ik}^{m\text{NEW}} = a_{ik}^{m\text{OLD}} - \tau \mu_k^m \cdot w_k^m \cdot \frac{2(x_i^m - a_{ik}^m)}{b_{ik}^m} \cdot \delta^m$$

and

$$b_{ik}^{m\text{NEW}} = b_{ik}^{m\text{OLD}} - \tau \mu_k^m \cdot w_k^m \cdot \frac{(x_i^m - a_{ik}^m)^2}{(b_{ik}^m)^2} \cdot \delta^m$$

respectively. The \(a_{ik}^m\) and \(b_{ik}^m\) in the Gaussian membership functions are not in the same scale, but by normalizing input and output values in interval [0, 1], we have a fast convergence by the same step size \(\tau\) (Ichihashi [8]).

2.2. Quantification from Pairwise Comparisons

Let \(r\)-th object \(O_r\) and \(s\)-th object \(O_s\) be compared pairwise. Suppose each numerical vector value that represents the objects is assumed to be known. When a set of paired input numerical vector values and difference
or ratio of corresponding output values is given for supervised learning, the proposed new algorithm may be used to learn an input–output mapping hidden behind the given data. With our proposed method, a continuous function of Class $C^\infty$ is represented by fuzzy rules. This function interpolates approximate values obtained from pairwise comparisons.

For simplicity we explain a derivation of iterative learning rule by using the hierarchical fuzzy model (HFM) shown in Figure 2. First we describe the case when we are given a set of data whose values are differences of the two outputs corresponding to the paired inputs. This difference values are used as desired differences of two outputs in supervised learning. Thus the structure of the model is based on the assumption of an interval scale property. Let the objective function be

$$E = \frac{1}{2} (y^s_r - y^s_s - y^*)^2$$

where $r$ and $s$ represent the $r$-th and $s$-th object respectively, and $y^s_r$ is the output of the hierarchical fuzzy model (Figure 2) corresponding to the input vector value of the $r$-th object. $y^*$ is the desired difference of the two outputs of the $r$-th and $s$-th objects. $y^*$'s are the given numerical data for training.

If $m = 3$,

$$\delta_{rs}^3 = y^r_r - y^s_s - y^*$$

Hence, the learning rule of $w^3_k$ is:

$$w^{3\text{NEW}}_k = w^{3\text{OLD}}_k - \tau \cdot \delta_{rs}^3 \cdot (\mu_{kr}^3 - \mu_{ks}^3)$$

Similarly, the learning rule of $a^{3\text{NEW}}_i$ and $b^{3\text{NEW}}_i$ based on the steepest descent method are, for $i = 1, 2$,

$$a^{3\text{NEW}}_i = a^{3\text{OLD}}_i - 2 \cdot \tau \cdot w^3_i \cdot \delta_{rs}^3 \cdot \left\{ \mu_{kr}^3 \left( \frac{x^r_{ir} - a^{3\text{OLD}}_i}{b^3_{ik}} \right) - \mu_{ks}^3 \left( \frac{x^s_{is} - a^{3\text{OLD}}_i}{b^3_{ik}} \right) \right\}$$

and

$$b^{3\text{NEW}}_i = b^{3\text{OLD}}_i - \tau \cdot w^3_i \cdot \delta_{rs}^3 \cdot \left\{ \mu_{kr}^3 \left( \frac{(x^r_{ir} - a^{3\text{OLD}}_i)^2}{(b^3_{ik})^2} \right) - \mu_{ks}^3 \left( \frac{(x^s_{is} - a^{3\text{OLD}}_i)^2}{(b^3_{ik})^2} \right) \right\}$$
When $m = 1, 2$,

$$\frac{\partial E}{\partial w_k^m} = \frac{\partial E}{\partial x_{mr}^3} \cdot \frac{\partial x_{mr}^3}{\partial w_k^m} + \frac{\partial E}{\partial x_{ms}^3} \cdot \frac{\partial x_{ms}^3}{\partial w_k^m}$$

where $x_{mr}^3$ is the output of the $m$-th model (i.e., the $m$-th input variable of the third model). Hence, for the input of the $r$-th object

$$\frac{\partial E}{\partial x_{mr}^3} = \frac{\partial E}{\partial y_r^3} \cdot \frac{\partial y_r^3}{\partial x_{mr}^3}$$

Let $\partial E/\partial x_{mr}^3$ be a back-propagation error $\delta_m^r (m = 1, 2)$ for the $r$-th input. The amount of the change in one repetition becomes

$$\Delta w_k^m = -\tau (\delta_r^m \mu_{kr}^m + \delta_s^m \mu_{ks}^m),$$

$$\Delta a_{ik}^m = -2 \cdot \tau \cdot w_k^m$$

$$\cdot \left\{ \mu_{kr}^m \frac{(x_{ir}^m - a_{ik}^m)}{b_{ik}^m} \delta_r^m + \mu_{ks}^m \frac{(x_{is}^m - a_{ik}^m)}{b_{ik}^m} \delta_s^m \right\}$$

and

$$\Delta b_{ik}^m = -\tau \cdot w_k^m$$

$$\cdot \left\{ \mu_{kr}^m \frac{(x_{ir}^m - a_{ik}^m)^2}{(b_{ik}^m)^2} \delta_r^m + \mu_{ks}^m \frac{(x_{is}^m - a_{ik}^m)^2}{(b_{ik}^m)^2} \delta_s^m \right\}$$

for $m = 1, 2$, where $x_{ir}^m$ is the $i$-th input variable of the $m$-th model corresponding to the $r$-th object’s input.

Next, we describe the case when we are given a set of data whose values are the ratio of two outputs corresponding to paired input vector valued data. Thus we structure a ratio scale model. Let the objective function be

$$E = \frac{1}{2} \left( \log \frac{y_r^3}{y_s^3} - \log y^* \right)^2$$

where $r$ and $s$ represent the $r$-th and $s$-th object, respectively. $y_r^3$ is the output of the hierarchical fuzzy model (Figure 2) corresponding to the $r$-th object’s input vector value. $y^*$ is the desired ratio of the two outputs of the $r$-th and $s$-th objects. $y^*$’s are the given numerical data for training. The logarithmic transformation is introduced to simplify the learning rule. This does not imply a transformation of physical values to psychological values.
If \( m = 3 \),
\[
\frac{\partial E}{\partial w_k^3} = \delta_{rs}^3 \cdot \left( \frac{1}{y_r^3} \cdot \mu_{kr}^3 - \frac{1}{y_s^3} \cdot \mu_{ks}^3 \right)
\]
where \( \delta_{rs}^3 = \log y_r^3 - \log y_s^3 - \log y^* \). Hence, the learning rule of \( w_k^3 \) is:
\[
w_k^{NEW} = w_k^{OLD} - \tau \cdot \delta_{rs}^3 \cdot \left( \frac{1}{y_r^3} \cdot \mu_{rk}^3 - \frac{1}{y_s^3} \cdot \mu_{sk}^3 \right)
\]
For \( a_{ik}^3 \) and \( b_{ik}^3 \), we have:
\[
a_{ik}^{NEW} = a_{ik}^{OLD} - 2 \cdot \tau \cdot w_k^3 \cdot \delta_{rs}^3
\]
\[
\cdot \left( \frac{1}{y_r^3} \cdot \mu_{kr}^3 \left( x_{ir}^3 - a_{ik}^3 \right) - \frac{1}{y_s^3} \cdot \mu_{ks}^3 \left( x_{is}^3 - a_{ik}^3 \right) \right)
\]
and
\[
b_{ik}^{NEW} = b_{ik}^{OLD} - \tau \cdot w_k^3 \cdot \delta_{rs}^3
\]
\[
\cdot \left( \frac{1}{y_r^3} \cdot \mu_{kr}^3 \left( \frac{x_{ir}^3 - a_{ik}^3}{b_{ik}^3} \right)^2 - \frac{1}{y_s^3} \cdot \mu_{ks}^3 \left( \frac{x_{is}^3 - a_{ik}^3}{b_{ik}^3} \right)^2 \right)
\]
When \( m = 1, 2 \),
\[
\frac{\partial E}{\partial w_k^m} = \frac{\partial E}{\partial x_{mr}^m} \cdot \frac{\partial x_{mr}^m}{\partial w_k^m} + \frac{\partial E}{\partial x_{ms}^m} \cdot \frac{\partial x_{ms}^m}{\partial w_k^m}
\]
For the \( r \)-th object,
\[
\frac{\partial E}{\partial x_{mr}^m} = \frac{\partial E}{\partial y_r^3} \cdot \frac{\partial y_r^3}{\partial x_{mr}^m}
\]
Denoting this value as \( \delta_r^m \) (\( m = 1, 2 \): back-propagation error) we have:
\[
\Delta w_{mk} = -\tau \left( \delta_r^m \cdot \frac{1}{y_r^m} \cdot \mu_{kr}^m + \delta_s^m \cdot \frac{1}{y_s^m} \cdot \mu_{ks}^m \right)
\]
\[
\Delta a_{ik}^m = -2 \cdot \tau \cdot w_k^m \cdot \delta_{rs}^m
\]
\[
\cdot \left( \frac{1}{y_r^m} \cdot \mu_{kr}^m \frac{x_{ir}^m - a_{ik}^m}{b_{ik}^m} \cdot \delta_r^m + \frac{1}{y_s^m} \cdot \mu_{ks}^m \frac{x_{is}^m - a_{ik}^m}{b_{ik}^m} \cdot \delta_s^m \right)
\]
and
\[
\Delta b_{ik}^m = -\tau \cdot w_k^m \cdot \delta_{rs}^m
\]
\[
\cdot \left( \frac{1}{y_r^m} \cdot \mu_{kr}^m \frac{(x_{ir}^m - a_{ik}^m)^2}{(b_{ik}^m)^2} \cdot \delta_r^m + \frac{1}{y_s^m} \cdot \mu_{ks}^m \frac{(x_{is}^m - a_{ik}^m)^2}{(b_{ik}^m)^2} \cdot \delta_s^m \right)
\]
Neuro-Fuzzy Approach to Data Analysis

Figure 3. Output of the hierarchical fuzzy model after learning from a set of data of pairwise comparisons generated by Eq. (4). ($x_1^2 = 0.5, x_2^2 = 0.5$).

In Figure 3 the graph of $y$ in Eq. (4) is shown together with the output of the hierarchical fuzzy model after learning with $\tau = 0.01$ (after 1,000 iterations $\Sigma (\delta^3)^2 = 0.0151$). In Figure 3, $x_1^2$ and $x_2^2$ are fixed at 0.5. For the hierarchical fuzzy model shown in Figure 2, we used a set of 400 paired input–output data chosen randomly in interval [0, 1] and $y^*$'s generated by Eq. (4). The variable $x_3^2$ in Figure 2 is redundant because Eq. (4) does not contain the variable $x_3^2$.

$$y = 0.5 \cdot \sin(\pi \cdot x_1^1) \cdot \sin(\pi \cdot x_3^1) \cdot \sin(\pi \cdot x_3^2)$$

(4)

$y^* = y_r - y_s$ is a desired value of the difference of the two output values corresponding to randomly chosen paired input data ($x_{ir}^m, x_{is}^m, m = 1, 2, i = 1, 2$). Nine membership functions of two variables (i.e., Gaussian radial basis functions as in Figure 1) are used in every submodel. Initial values of the parameters are: $a_{ik} \in \{0.0, 0.5, 1.0\}, b_{ik} = 0.2, w_{ik}^m (m = 1, 2)$ are chosen randomly from interval [0, 1] and $w_k^3 = 0$ for all $k$. The result shown in Figure 3 is almost the same as the one given by Eq. (4) that is depicted by the dashed curve. Since the difference of two values is given as desired output $y^*$, the obtained value $y^3$ (i.e., output of the hierarchical fuzzy model) is assumed to have an interval scale. Figure 4 shows computer graphics of three submodels of the hierarchical fuzzy model. In this way the structure of the model can be visualized if each submodel has two input variables. This is an advantage of the hierarchical fuzzy model as compared with the multi-layer neural network.

3. GUTTMAN'S METHOD OF QUANTIFYING PAIRWISE COMPARISONS

In this section, we briefly survey the Guttman's method and propose its fuzzified version. This method is applied in Section 5 and compared with the neuro-fuzzy approach.
Guttman's quantification method by pairwise comparisons [15] is a sort of data-analysis technique. The solution appears in the form of the latent vector associated with the maximum root of a matrix obtained from the pairwise comparisons. The judgment in Guttman's pairwise comparisons is to have the individuals decide which of the two objects should be given a higher rank. The value of the computational result from this approach seems to have an interval scale because any of the real values can be chosen as a mean value of the associated weights. In other words any origin of the scale can be chosen. The basic principle of Guttman's computational scheme is that a group of data given to a decision-maker that is judged to be higher should be as different as possible from another group of data that is judged to be lower. The respective weights of each of the \( n \) objects would be obtained by separating the distribution of the two groups as far as possible. Naturally, judgments vary from decision-maker to decision-maker. The principle is to determine a set of weights for the \( n \) objects to be compared pairwise in order to represent the whole population effectively. For example paired comparisons may come from \( N \) decision-makers. The basic principle is to quantify the judgments by obtaining the proper weights in preferences that would maximize the distance of the two groups such that these two groups would separate into two distinct groups. In other words, the weights could be determined in preferences by maximizing two groups' differences and minimizing the variance of the whole group simultaneously.

Figure 4. Three dimensional graphics of the hierarchical fuzzy model after learning.
Unfortunately, we found that the Guttman's method has two demerits.

1. There are only two types of responses (0 or 1) from the decision-maker. For example, decision-maker \( i \) is required to compare objects \( O_r \) and \( O_s \). If his judgment was \( O_r > O_s \), we only have the information of \( O_r \) is larger than \( O_s \) but we have no information on how large \( O_r \) is compared with \( O_s \). Therefore, if the number of decision-makers is small, it is difficult to measure an appropriate scale from analyzing the small amount of data.

2. All decision-makers are required to make pairwise comparisons of all \( n \) objects in a compulsory manner. Because an exhaustive comparison is required, the process of pairwise comparisons takes a very long time. Decision-makers undertake a heavy load to make all pairwise comparisons in such a process.

### 3.1. Quantification by fuzzy Frequencies and Fuzzy Variances

Only the pairwise comparisons method by Guttman provides the judgment of "greater than" or "smaller than" in which they are represented by two values \( \{1\} \), \( \{0\} \), respectively. The solution is obtained by the eigenvector method. However, in order to obtain much more precise weights from the small quantity of data, the information "how much greater than" or "how much smaller than" are necessary. Therefore, we introduce the fuzzy set theory proposed by Zadeh [20]. The fuzzy quantification method type I–IV was proposed by Watada and Tanaka [14]. It should be recalled that Guttman’s pairwise comparisons have a disadvantage because the method requires exhaustive pairwise comparisons. In other words, the decision-maker must answer \( n(n - 1)/2 \) questions in pairwise comparisons. By applying the fuzzy quantification method in [14], we can obtain a comparatively reliable complementary solution even when we have a non-exhaustive set of pairwise comparisons. First, let \( O_1, O_2, \ldots, O_n \) be the \( n \) objects to be compared by means of pairwise comparisons. Let each of the \( N \) decision-makers be asked to make judgments on all \( n \) objects. For instance, suppose a decision-maker \( i \) is asked to judge how large \( O_r \) is compared with \( O_s \). Let \( e_{irs} \in [0, 1] \) represent the degree with which a decision-maker \( i \) judges \( O_r \) to be greater or smaller in comparison to \( O_s \). Clearly in Guttman’s method, the answers to such preferences are not obtained.

We will first describe the case for an exhaustive set of comparisons. Thus each of the \( N \) decision-makers is asked to make the \( n(n - 1)/2 \) pairwise comparisons. Hence the total number of pairwise comparisons is \( Nn(n - 1)/2 \) in the whole process. A standard definition of \( e_{irs} \)'s are presented in Table 1. When a decision-maker \( i \) is asked to make a
Table 1. Definition of Intensity of Comparisons Corresponding to $e_{irs}$ (Fuzzy Version of Guttman's Method)

<table>
<thead>
<tr>
<th>Guttman</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_r$ is absolutely greater than $O_s$</td>
<td>0.9</td>
</tr>
<tr>
<td>$O_r$ is demonstrably greater than $O_s$</td>
<td>0.8</td>
</tr>
<tr>
<td>$O_r$ is essentially greater than $O_s$</td>
<td>0.7</td>
</tr>
<tr>
<td>$O_r$ is weakly greater than $O_s$</td>
<td>0.6</td>
</tr>
<tr>
<td>$O_r$ is equal to $O_s$</td>
<td>0.5</td>
</tr>
<tr>
<td>$O_r$ is weakly smaller than $O_s$</td>
<td>0.4</td>
</tr>
<tr>
<td>$O_r$ is essentially smaller than $O_s$</td>
<td>0.3</td>
</tr>
<tr>
<td>$O_r$ is demonstrably smaller than $O_s$</td>
<td>0.2</td>
</tr>
<tr>
<td>$O_r$ is absolutely smaller than $O_s$</td>
<td>0.1</td>
</tr>
</tbody>
</table>

(comparison of object $r$ to object $s$) \(\rightarrow (e_{irs})\)

judgment on how much greater (smaller) $O_r$ is than $O_s$, s/he refers to Table 1 and gives a response s/he assesses to be applicable. Thus judgments are represented by $e_{irs}$'s as an intensity of comparisons. If $O_r$ is equal to $O_s$ ($r \neq s$), then $e_{irs} = 0.5$. However, when $r = s$ we set $e_{irs} = 0$, $e_{isr} = 0$ ($i = 1, \ldots, N; r, s = 1, \ldots, n$).

As we see in the following, the computational scheme of the fuzzy version is almost the same as original Guttman's method [15]. We can observe that if $O_r$ is judged greater than $O_s$, then $O_s$ should be judged smaller than $O_r$, i.e., connectivity axiom of weak order. So, from the definition of intensity of comparisons corresponding to $e_{irs}$ in Table 1, we have:

$$e_{irs} + e_{isr} = 1, \quad (r \neq s)$$

Let $f_{ir}$ be the fuzzy frequency of an object $O_r$ that a decision-maker $i$ judges $O_r$ to be greater or smaller (as in Table 1) than the other objects ($O_s$),

$$f_{ir} = \sum_s e_{irs}$$

Conversely, let $g_{ir}$ be the fuzzy frequency of an object $O_r$ that a decision-maker judges the other objects ($O_s$) to be greater or smaller than $O_r$,

$$g_{ir} = \sum_s e_{isr}$$

With the assumption that complete pairwise comparisons are available, we have:

$$f_{ir} + g_{ir} = n - 1$$
for all $i$ and $r$. Let $F$ be the total fuzzy frequency of comparisons made by each decision-maker,

$$F = \sum_s f_{is} = \sum_s g_{is} = \frac{1}{2}n(n - 1)$$

for all $i$. Thus, let $c$ be the total fuzzy frequency such that $O_r$ was judged

$$c = \sum_i (f_{ir} + g_{ir}) = N(n - 1)$$

Let $C$ be the total fuzzy frequency of judgments

$$C = Nn(n - 1)$$

Here, the respective weights of each of the $n$ objects will be obtained by separating the distribution of the two groups as far as possible. The fuzzy means and fuzzy sum of squares to be considered are defined as follows. Let $w_r$ be the weight to be derived for $O_r$ on the basis of the comparisons. Let $t_i$ be the fuzzy mean of the weights of the objects that the decision-maker $i$ judges the object to be greater or smaller than all the other objects,

$$t_i = \frac{1}{F} \sum_s w_s f_{is}$$

Let $u_i$ be the fuzzy mean of the weights of the objects that the decision-maker $i$ judges all the other objects to be greater or smaller than the object,

$$u_i = \frac{1}{F} \sum_s w_s g_{is}$$

Let $M$ and $V^2$ be the total fuzzy mean and the total fuzzy sum of squares judged by all the decision-makers respectively.

$$M = \frac{1}{C} \sum_s w_s c = \frac{1}{n} \sum_s w_s$$

$$V^2 = \sum_s (w_s - M)^2 c$$

$$= c \sum (w_s)^2 - M^2 C$$

(5)
We define $R^2$ to be the fuzzy sum of squares between two groups,
\[
R^2 = \sum_i \left[ (t_i - M)^2 + (u_i - M)^2 \right] F
\]
\[
= \sum_i \left( (t_i)^2 + (u_i)^2 \right) F - M^2 C
\quad (6)
\]

The ratio $E^2$ is defined by the ratio of $V^2$ and $R^2$ as:
\[
E^2 = \frac{R^2}{V^2}
\]

According to the Guttman's method, the principle is to quantify the judgments by obtaining the weights that will minimize the sum of squares as a whole compared to the sum of squares between two groups. This means that the problem is to determine the weights that will maximize the ratio of sum of squares. In our proposed fuzzy extension of the Guttman's method, the weights can be obtained by quantifying the judgments that will maximize the ratio $E^2$. Since the ratio $E^2$ is invariant with respect to the translation of weights, we can, without loss of generality, set
\[
M = 0
\]

Thus by rewriting Eqs. (5) and (6), we have
\[
V^2 = \sum_s (w_s)^2
\quad (7)
\]
\[
R^2 = \sum_i \left( (t_i)^2 + (u_i)^2 \right) F
\quad (8)
\]

Furthermore, maximizing the ratio $E^2$ is the same as fixing the total fuzzy sum of squares $V^2$ and maximizing the fuzzy sum of squares between two groups i.e., $R^2$.

In order to find the maximum value of $E^2$, we differentiate $E^2$ with respect to the $w_s$ and set the differentiations equal to zero. Thus, we have
\[
\frac{\partial R^2}{\partial w_s} = E^2 \frac{\partial V^2}{\partial w_s}
\quad (9)
\]

Moreover, the maximum value of $E^2$ that satisfies Eq. (9) is also the same value that maximizes the ratio
\[
E^2 = \frac{R^2}{V^2}
\]
From Eqs. (7) and (8), we have

$$\frac{\partial V^2}{\partial w_r} = 2w_r c$$

$$\frac{\partial R^2}{\partial w_r} = 2 \sum_s w_s \sum_i \frac{1}{F} (f_{ir} f_{is} + g_{ir} g_{is})$$

Hence,

$$\sum_s w_s \sum_i \frac{1}{F} (f_{ir} f_{is} + g_{ir} g_{is}) = E^2 w_r c \quad (10)$$

Finally, we can obtain the respective weights $w_r$'s by determining the maximum $E^2$ that satisfies Eq. (10). Let

$$H_{rs} = \frac{1}{cF} \sum_i (f_{ir} f_{is} + g_{ir} g_{is})$$

We have

$$\sum_s w_s H_{rs} = E^2 w_r \quad (11)$$

We need to consider Eq. (11) for all $r$. That is Eq. (11) represents a system of simultaneous equations. Thus converting Eq. (11) into vector matrix form, we have:

$$Hw = E^2 w \quad (12)$$

Then solving this eigenvalue problem, weight $w$ can be obtained. It should be noted that Eq. (12) has a trivial solution and the maximal eigenvalue must be chosen from nontrivial solutions. Since weight $w$ is obtained only by maximizing the ratio $E^2$, it may happen that the sign of $w_r$ is reversed for all $r$.

The exhaustive set of pairwise comparisons is needed in Guttman's method. When the number of objects increases, the number of pairwise comparisons will increase rapidly. The time required to finish the whole process will also increase. An alternate of the proposed fuzzy approach provides a scheme for the reduction of the number of pairwise comparisons by curtailing parts of pairwise comparisons. An abridged pairwise comparisons can be determined by taking the average of all parts which pass from $r$ to $s$. Let $O_1, O_2, \ldots, O_n$ be the $n$ objects to be compared by means of pairwise comparisons. Decision-maker $i$ is asked to judge how large $O_r$ is compared with $O_s$ using Table 1. Suppose the degree of largeness is represented by $e_{irl}$. If $O_r > O_l$ and $O_s > O_l$ were also judged,
we can estimate $e_{irs}$ from $e_{irl}$ and $e_{isl}$ as

$$e_{irs} = (e_{irl} - e_{isl}) + 0.5$$

An estimate may be found by taking the average degree of all paths of length two.

$$e_{irs} = \frac{1}{q} \sum_{l} (e_{irl} - e_{isl}) + 0.5, \quad l \neq s$$

Where $q$ is the number of paths of length two.

However, a covering is the necessary condition when there are incomplete pairwise comparisons. Let $Q = \{1, 2, \ldots, n\}$ be the index set of $n$ objects. Let $\Omega = \{I_1, I_2, \ldots, I_n\}$ be the set of pairs, $I_s = (i, j), i \neq j$. Let $C = (c_1, c_2)$ and $D = (d_1, d_2)$ denote two elements of $\Omega$. $\Omega$ is said to be connected if and only if there exists a sequence $I_1, I_2, \ldots, I_s \in \Omega$ such that $I_s = C$ and $I_{s+1} = D$ with $C \neq D$ and $\{c_1, c_2\} \cap \{d_1, d_2\} \neq \emptyset$ for any $s \in \{1, \ldots, S - 1\}$. $\Omega$ is said to be a covering of $Q$ when $\Omega$ is connected and $Q$ is included by the union of the elements of $\Omega$. This condition of covering is necessary for the neuro-fuzzy too, when some parts of the pairwise comparisons are curtailed.

4. **SAATY'S AHP AND INCOMPLETE PAIRWISE COMPARISONS**

The analytic hierarchy process (AHP) was introduced by Saaty [16] as a tool for dealing with complex decision-making problems. AHP starts by decomposing a decision-making problem into a hierarchy. Each level consists of a few elements, and each element is, in turn, decomposed into another set of elements. The process continues down to the lowest level (alternative level). Structurally, the hierarchy is broken down into a series of pairwise comparison matrices, and the decision-makers are asked to evaluate all the matrices. In order to simplify the illustration, we will only review the pairwise comparisons and derivation of weights of objects. Let $O_1, O_2, \ldots, O_n$ be $n$ objects, and decision-makers are asked to compare object $O_i$ with object $O_j$ using the figure in Table 2. We suppose $d_{rr} = 1$, $d_{rs} = 1/d_{sr}$. The principal eigenvector of this matrix is then derived and weighted by the priority of the property with respect to which the comparison is made. Let the matrix so obtained by $D = [d_{rs}]$, then the eigenvalue problem is:

$$(D - \lambda E)w = 0$$

where $E$ is a unit matrix.
Table 2. Definition of Intensity in Comparisons Corresponding to $d_{rs}$ (Saaty’s AHP)

<table>
<thead>
<tr>
<th>Intensity</th>
<th>AHP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_r$ is absolutely greater than $O_s$</td>
<td>9</td>
</tr>
<tr>
<td>$O_r$ is demonstrably greater than $O_s$</td>
<td>7</td>
</tr>
<tr>
<td>$O_r$ is essentially greater than $O_s$</td>
<td>5</td>
</tr>
<tr>
<td>$O_r$ is weakly greater than $O_s$</td>
<td>3</td>
</tr>
<tr>
<td>$O_r$ is equal to $O_s$</td>
<td>1</td>
</tr>
<tr>
<td>$O_r$ is weakly smaller than $O_s$</td>
<td>$1/3$</td>
</tr>
<tr>
<td>$O_r$ is essentially smaller than $O_s$</td>
<td>$1/5$</td>
</tr>
<tr>
<td>$O_r$ is demonstrably smaller than $O_s$</td>
<td>$1/7$</td>
</tr>
<tr>
<td>$O_r$ is absolutely smaller than $O_s$</td>
<td>$1/9$</td>
</tr>
</tbody>
</table>

(comparison of object $r$ to object $s$) → ($d_{rs}$)

Eq. (13) has a positive solution $w$ with respect to the maximum eigenvalue $\lambda$. The AHP is a decision-analysis technique that uses judgments from a group of relevant decision-makers. We take the geometric mean of the data from different decision-makers for aggregating those responses. The problem in this method is that, with the standard mode of questioning in the AHP a decision-maker must determine a matrix by answering $n(n-1)/2$ questions. When the number of objects increases, then the number of pairwise comparison will increase rapidly.

From the principle of the eigenvector method, pairwise comparison of all objects is not an absolute condition. We can obtain the result even when part of the pairwise comparisons are abridged. The abridged method was proposed by Harker [21, 22].

5. AN APPLICATION TO ANALYSIS OF PSYCHOLOGICAL SENSATION LEVEL TO MINUTE VIBRATIONS

In the fields of ergonomics, psychology, etc., small-sized accelerometers (Figure 5) are used to investigate the perception and sensation to minute vibration. In this paper, we used this kind of accelerometer to analyze the psychological sensation responses to minute vibrations. In this experiment, we choose three types of frequencies (63 Hz, 125 Hz, 250 Hz) and five types of vibration acceleration levels (63 Hz: 7.5, 12.5, 17.5, 22.5, 27.5 dB) (125 Hz and 250 Hz: 20, 25, 30, 35, 40 dB). Generally human beings cannot feel vibration when the vibration acceleration level is below 7.5 dB of 63 Hz. Therefore we set the minimum value of vibration acceleration level to 7.5 dB. We have adopted three types of frequencies and five types of vibration acceleration levels, so we have 15 combinations of vibration
outputs. We chose ten healthy men all aged around 20-years-old to be our subjects. The subjects are asked to make pairwise comparisons in agreement with the alternative responses shown in Tables 1–3. The responses of all subjects are collected and analyzed. The linguistic expressions are the same among Tables 1–3, but the numerical values used for computation in each method are different from each other.

The proposed two neuro-fuzzy models, the fuzzified version of Guttman’s method, and the AHP are used to calculate the associated weights of 15 vibration outputs (i.e., three types of frequencies and five types of acceleration levels). First, five standard objects are selected from 15 combinations of vibration outputs. Then, these five standard objects are compared to other objects by pairwise comparisons. In order to make a covering in the process, the largest vibration acceleration level and frequency was chosen as a standard object. Before starting the experiment for the training, the sensation of the maximum stimulus and the sensation of the minimum stimulus are given to all subjects. It is necessary to determine some standard first, because the pairwise comparisons of some psychological levels are extremely difficult. Secondly, the standard objects are compared with the other objects. Originally, the total number of comparisons in complete pairwise comparisons is \( n(n - 1)/2 = 105 \), because the number of objects to be compared is \( n = 3 \times 5 = 15 \). But, we reduced the num-
ber of comparisons to 40 by choosing five standard objects. In general the number of comparisons can be decreased to about half of the number of complete comparisons. In AHP, we take the geometric mean of the data from the ten subjects for aggregating their responses, and we convert the maximum value of weight to 1. For the neuro-fuzzy, the definition of intensity of comparisons is given in Table 3. To have the associated weights in interval \([-1, +1]\), the value of \(y^*\) is chosen from \([-2, +2]\) for the case that \(y^*\) denotes the difference of two objects. The number of iteration is 1,000 and the learning rate \(\tau = 0.0001\). Two input variables are frequency (Hz) and vibration acceleration level (dB), which are normalized within the interval \([0, 1]\). The computational results shown in Figures 6 and 7 are approximately the same. As the value of vibration acceleration level becomes greater, the psychological value becomes larger. Figures 6 and 7 show the results in the cases of difference data and ratio data respectively. The results by the fuzzified version of Guttman’s method in Figure 6 show that the sign of \(w_r\) is reversed for all \(r\).

Table 3. Definition of Intensity in Comparisons Corresponding to \(y^*\) (Neuro-Fuzzy Model)

<table>
<thead>
<tr>
<th>Difference</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>(O_r) is absolutely greater than (O_s)</td>
<td>2.0</td>
</tr>
<tr>
<td>(O_r) is demonstrably greater than (O_s)</td>
<td>1.5</td>
</tr>
<tr>
<td>(O_r) is essentially greater than (O_s)</td>
<td>1.0</td>
</tr>
<tr>
<td>(O_r) is weakly greater than (O_s)</td>
<td>0.5</td>
</tr>
<tr>
<td>(O_r) is equal to (O_s)</td>
<td>0.0</td>
</tr>
<tr>
<td>(O_r) is weakly smaller than (O_s)</td>
<td>-0.5</td>
</tr>
<tr>
<td>(O_r) is essentially smaller than (O_s)</td>
<td>-1.0</td>
</tr>
<tr>
<td>(O_r) is demonstrably smaller than (O_s)</td>
<td>-1.5</td>
</tr>
<tr>
<td>(O_r) is absolutely smaller than (O_s)</td>
<td>-2.0</td>
</tr>
</tbody>
</table>

(comparison of object \(r\) to object \(s\) \(\rightarrow (y^*)\))

Though the vibration–acceleration level is the value transformed by logarithmic function from physical value in each frequency, the values obtained in Figures 6 and 7 present a common scale among all frequencies. The psychological values obtained from each human being are greatly different. Therefore the results represent average value of many subjects (or decision-makers). Hence, it is difficult to state analytically which method is better. In comparing the results, we find little difference from the results of both Guttman’s method and AHP. The AHP is a convenient method to structure a ratio scale for one subject (or one decision-maker), but the geometric mean used to have a reciprocal matrix from many subjects is not rational except that it maintains reciprocal property. In the Guttman’s method, to maximize the ratio of the sum of squares is indirect,
6. CONCLUSIONS

In this paper, a neuro-fuzzy approach to quantify pairwise comparisons is proposed. Fuzzy rules are obtained in the form of Gaussian radial basis

Figure 6. The psychological sensation level of magnitude in minute vibration calculated from the difference data.

and it is not clear that the scale so obtained is an interval scale. The proposed neuro-fuzzy model is easier to treat many subjects (or decision-makers) than the AHP. The proposed approach is a more direct way to reconstruct interval and ratio scales by minimizing the squared sum of errors. In this manner, psychological values are obtained by smooth functions of class $C^\infty$.

Figure 7. The psychological sensation level of magnitude in minute vibration calculated from the ratio data.
functions. The method can be applied even in cases of incomplete pairwise comparisons. Both the interval scale and the ratio scale results are widely used in actual problems. Due to the difference in the characteristic of problems, the most important point is to choose the right scale to satisfy the property of the problem.

ACKNOWLEDGMENT

We are grateful to Mr. K. K. So, Mr. K. Kin, and Mrs. I. Mu, students of U.O.P., for their assistance. We are especially indebted to Professor Y. Kume of U.O.P. for his discussion on an earlier draft of this paper and his advice on the experimental study.

References


