Distance measure for linguistic decision making

Yejun Xu*, Huimin Wang

Business School, Hohai University, Nanjing, 211100, China

Abstract

In this paper, we extend the distance measure to the linguistic fuzzy sets, and develop the linguistic distance operators, such as linguistic weighted distance (LWD) operator, linguistic ordered weighted distance (LOWD) operator, and study some of their desired properties. These aggregation operators are very useful for decision-making problems because they establish a comparison between an ideal alternative and available options in order to find the optimal choice. We also develop a procedure to the linguistic decision problem with the developed linguistic distance operators. Finally, a practical example is given to illustrate the multiple attribute group decision making process.

© 2011 Published by Elsevier B.V. Selection and/or peer-review under responsibility of the Organising Committee of The International Conference of Risk and Engineering Management.

Keywords: Linguistic decision making; distance measure; linguistic weighted distance (LWD) operator; linguistic ordered weighted distance (LOWD) operator; engineering investment.

1. Introduction

In day-to-day activities we have to solve different problems and depending on aspects presented by each problem we can deal with different type of precise numerical values, but in other cases, the problems present qualitative aspects that are complex to assess by means precise and exact values. In the latter case, the use of fuzzy linguistic approach has provided very good results. For example, when evaluating the “comfort” or “design” of a car, terms like “good”, “medium”, “bad”[1] are usually used, and evaluating a car’s speed, terms like “very fast”, “fast”, “slow” can be used instead of numeric values[2].

Distance measures are fundamentally important in a variety of scientific fields such as decision making, pattern recognition, machine learning and market prediction, etc. Distance measures are a common tool widely used for measuring the deviations of different arguments. In the existing literature, a variety of distance measures have been introduced and investigated, such as the Hamming distance[3], the Euclidean distance[4], Hausdorff metric[5], etc. And also these distance measures have been extended to the intuitionistic fuzzy sets (IFSs)[6], inter-valued intuitionistic fuzzy sets (IVIFSs) [7], hesitant fuzzy sets (HFs)[8], linguistic fuzzy sets [9], etc. In this paper, we develop the distance measure to the linguistic fuzzy sets. In order to do this, the reminder of the paper is organized

* Corresponding author. Tel.: +86-25-85427377; fax: +86-25-85427972.
E-mail address: xuyejohn@163.com.

© 2011 Published by Elsevier B.V. Selection and/or peer-review under responsibility of the Organising Committee of The International Conference of Risk and Engineering Management.
doi:10.1016/j.sepro.2011.08.066
as follows. Section 2 introduces some basic concepts of linguistic variables and their operational laws. Section 3, the distance measure is extended to the linguistic fuzzy sets, and developed some linguistic distance operators, such as linguistic weighted distance (LWD) operator, linguistic normalized distance (LND) operator, linguistic ordered weighted distance (LOWD) operator, and study some of their desired properties. Section 4 analyzes different families of LOWD operator. In Section 5, we develop an approach to decision making with linguistic distance operators. Section 6, we illustrate an example to show the application of the linguistic distance operators. Finally, concluding remarks and future research are pointed out in Section 7.

2. Basic notations and operational laws

The linguistic approach is an approximate technique which represents qualitative aspects as linguistic values by means of linguistic variables. Suppose that $S = \{s_i | i = -t, \ldots, t\}$ is a finite and totally ordered discrete term set, where $s_i$ represents a possible value for a linguistic variable. For example, a set of nine terms $S$ could be

$$S = \{s_{-4} = \text{extremely poor}, s_{-3} = \text{very poor}, s_{-2} = \text{poor}, s_{-1} = \text{slightly poor}, s_0 = \text{fair},$$
$$s_1 = \text{slightly good}, s_2 = \text{good}, s_3 = \text{very good}, s_4 = \text{extremely good}\}$$

Obviously, the mid linguistic label $s_0$ represents an assessment of “indifference”, and with the rest of the linguistic labels being placed symmetrically around it.

In these cases, it is usually required that there exist the following:

1. The set is ordered: $s_i \geq s_j$ if $i \geq j$;
2. There is the negation operator: $\text{neg}(s_i) = s_{-i}$;
3. Max operator: $\text{max}(s_i, s_j) = s_i$, if $s_i \geq s_j$;
4. Min operator: $\text{min}(s_i, s_j) = s_i$, if $s_i \leq s_j$.

In the process of linguistic information, however, some results may not exactly match any linguistic labels in $S$. To preserve all the given information, we extend the discrete term set $S$ to a continuous term set $\tilde{S} = \{s_\alpha | \alpha \in [-t, t]\}$. If $s_\alpha \in S$, then we call $s_\alpha$ an original linguistic term, otherwise, we call $s_\alpha$ a virtual linguistic term. In general, the decision maker used the original linguistic terms to evaluate alternatives, and the virtual linguistic terms can only appear in operation.

Consider any two linguistic terms $s_\alpha, s_\beta \in \tilde{S}$, and $\lambda, \lambda_1, \lambda_2, \in [0,1]$, their operational laws are given as follows [14]:

1. $s_\alpha \oplus s_\beta = s_{\alpha+\beta}$;
2. $s_\alpha \ominus s_\beta = s_\beta \ominus s_\alpha$;
3. $\lambda s_\alpha = s_{\lambda\alpha}$;
4. $(s_\alpha)^\gamma = s_{\gamma\alpha}$;
5. $\lambda(s_\alpha \oplus s_\beta) = \lambda s_\alpha \oplus \lambda s_\beta$;
6. $(\lambda_1 + \lambda_2) s_\alpha = \lambda_1 s_\alpha \oplus \lambda_2 s_\beta$.

3. Linguistic aggregation operators with distance measure

**Definition 1.** Let $s_\alpha, s_\beta \in \tilde{S}$ be two linguistic variables, then we call

$$|s_\alpha - s_\beta| = s_{|\alpha-\beta|}$$

the distance between $s_\alpha$ and $s_\beta$.


Definition 2. Let \( s_{a_j}, s_{\beta_j} \in \overline{S} \) be two collections of linguistic variables, a linguistic weighted distance operator of dimension \( n \) is a mapping \( \text{LWD}: \overline{S}^n \times \overline{S}^n \rightarrow \overline{S} \) that has an associated weighting vector \( w \) of dimension \( n \) with \( \sum_{j=1}^{n} w_j = 1 \) and \( w_j \in [0,1] \), such that:

\[
\text{LWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) = \bigoplus_{j=1}^{n} w_j \left| s_{a_j} - s_{\beta_j} \right|
\]

Especially, if \( w_j = 1/n \), for all \( j \), then the linguistic weighted distance operator becomes the linguistic normalized distance (LND) operator, that is

\[
\text{LND} \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) = \frac{1}{n} \bigoplus_{j=1}^{n} \left| s_{a_j} - s_{\beta_j} \right|
\]

Now, we discuss some properties of the LWD operator.

Theorem 1 (Monotonicity). Let \( s_{a_j}, s_{\beta_j}, s_{a_j'}, s_{\beta_j'} \in \overline{S} \) (\( j = 1, 2, \ldots, n \)) be four collections of linguistic variables, if \( \left| s_{a_j} - s_{\beta_j} \right| \geq \left| s_{a_j'} - s_{\beta_j'} \right| \), for all \( j \), then

\[
\text{LWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) \geq \text{LWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right)
\]

Proof. It is straightforward and thus omitted.

Theorem 2 (Idempotency). Let \( s_{a_j}, s_{\beta_j} \in \overline{S} \) (\( j = 1, 2, \ldots, n \)) be two collections of linguistic variables, if \( \left| s_{a_j} - s_{\beta_j} \right| = d \), for all \( j \), then

\[
\text{LWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) = d
\]

Proof. It is straightforward and thus omitted.

Theorem 3 (Bounded). Let \( s_{a_j}, s_{\beta_j} \in \overline{S} \) (\( j = 1, 2, \ldots, n \)) be two collections of linguistic variables, then

\[
\min_j \left| s_{a_j} - s_{\beta_j} \right| \leq \text{LWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) \leq \max_j \left| s_{a_j} - s_{\beta_j} \right|
\]

Proof. Since \( \min_j \left| s_{a_j} - s_{\beta_j} \right| \leq \left| s_{a_j} - s_{\beta_j} \right| \leq \max_j \left| s_{a_j} - s_{\beta_j} \right| \), then

\[
\bigoplus_{j=1}^{n} w_j \left( \min_j \left| s_{a_j} - s_{\beta_j} \right| \right) \leq \text{LWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) \leq \bigoplus_{j=1}^{n} w_j \left( \max_j \left| s_{a_j} - s_{\beta_j} \right| \right)
\]

that is

\[
\min_j \left| s_{a_j} - s_{\beta_j} \right| \leq \text{LWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) \leq \max_j \left| s_{a_j} - s_{\beta_j} \right|
\]

Based on the OWA[20] operator and LWD, we define linguistic ordered weighted distance (LOWD) operator.

Definition 3. Let \( s_{a_j}, s_{\beta_j} \in \overline{S} \) be two collections of linguistic variables, a linguistic ordered weighted distance operator of dimension \( n \) is a mapping \( \text{LOWD}: \overline{S}^n \times \overline{S}^n \rightarrow \overline{S} \) that has an associated weighting vector \( w \) of dimension \( n \) with \( \sum_{j=1}^{n} w_j = 1 \) and \( w_j \in [0,1] \), such that:

\[
\text{LOWD}_w \left( \left( s_{a_1}, s_{\beta_1} \right), \ldots, \left( s_{a_n}, s_{\beta_n} \right) \right) = \bigoplus_{j=1}^{n} w_j s_{a_{\sigma(j)}}
\]

where \( s_{a_{\sigma(j)}} \) is the \( j \)th largest of the \( \left| s_{a_j} - s_{\beta_j} \right| \).
Based on the reordering step, we can distinguish between the Descending (LDOWD) operator and the Ascending (LAOWD) operator. Normally, we call the LOWD as (LDOWD) operator. If $s_{\alpha(i)}$ of Eq.(7) is arranged in ascending order, then we call it LAOWD operator.

**Theorem 4 (Commutativity).** Let $s_{\alpha_i}, s_{\beta_i}, s_{\alpha_j}, s_{\beta_j} \in \bar{S}$ ($j = 1, 2, \ldots, n$) be four collections of linguistic variables, then

$$\text{LOWD}_w \left( \left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle, \ldots, \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle \right) = \text{LOWD}_w \left( \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle, \ldots, \left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle \right)$$

(8)

where $\left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle, \ldots, \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle$ is any permutation of $\left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle, \ldots, \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle$.

**Theorem 5 (Monotonicity).** Let $s_{\alpha_i}, s_{\beta_i}, s_{\alpha_j}, s_{\beta_j} \in \bar{S}$ ($j = 1, 2, \ldots, n$) be four collections of linguistic variables, if $s_{\alpha_j} - s_{\beta_j} \geq s_{\alpha_i} - s_{\beta_i}$, for all $j$, then

$$\text{LOWD}_w \left( \left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle, \ldots, \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle \right) \geq \text{LOWD}_w \left( \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle, \ldots, \left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle \right)$$

(9)

**Theorem 6 (Idempotency).** Let $s_{\alpha_i}, s_{\beta_i} \in \bar{S}$ ($j = 1, 2, \ldots, n$) be two collections of linguistic variables, if $|s_{\alpha_i} - s_{\beta_i}| = d$, for all $j$, then

$$\text{LOWD}_w \left( \left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle, \ldots, \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle \right) = d$$

(10)

**Theorem 7 (Bounded).** Let $s_{\alpha_i}, s_{\beta_i} \in \bar{S}$ ($j = 1, 2, \ldots, n$) be two collections of linguistic variables, then

$$\min_j |s_{\alpha_i} - s_{\beta_i}| \leq \text{LOWD}_w \left( \left\langle s_{\alpha_i}, s_{\beta_i} \right\rangle, \ldots, \left\langle s_{\alpha_j}, s_{\beta_j} \right\rangle \right) \leq \max_j |s_{\alpha_i} - s_{\beta_i}|$$

(11)

4. Families of LOWD operators

An interesting feature of the LOWD operator is that it provides a parameterized family of distance aggregation operators between the maximum and the minimum. These families use a methodology for establishing the weights similar to the OWA operator. In the literature, we find a lot of methods for determining the OWA weights which also can be implemented for LOWD operator. By choosing different manifestation of the weighting vector, we are able to obtain different types of distance aggregation operators. In the following, we present some of these families.

**Remark 1.** If $w_1=1$, and $w_j=0$ for all $j \neq 1$, then the LOWD is reduced to the maximum distance. If $w_v=1$, $w_j=0$ for all $j \neq n$, then the LOWD is reduced to the minimum distance.

**Remark 2.** The step-LOWD operator with $w_k=1$ and $w_j=0$ for all $j \neq k$. Note that if $k=1$, the step-LOWD is reduced to the maximum distance operator, and if $k=n$, the step-LOWAD becomes the minimum distance operator.

**Remark 3.** The linguistic normalized distance is obtained when $w_j=1/n$, for all $j$, and the linguistic weighted distance is obtained when then ordered position of $i$ is the same as the ordered position of $j$.

**Remark 4.** The Olympic-LOWD is obtained when $w_1=w_n=0$, and for all others $w_j=1/(n-2)$.

**Remark 5.** A very useful approach for obtaining the weights that is also applicable for the LOWD operator is the functional method introduced by Yager for the OWA aggregation operator. We can obtain the OWA weights by

$$w_j = Q(j/n) - Q((j-1)/n), \quad j = 1, \ldots, n$$

(12)

where $Q$ is a basic unit-interval monotonic (BUM) function $Q: [0,1] \rightarrow [0,1]$ with $Q(0)=0$ and $Q(1)=1$ and. It can be shown these weights satisfy the conditions $w_j \in [0,1]$, and $\sum_{j=1}^{n} w_j = 1$. 
5. Approach to decision making with linguistic distance operators

Multiple attribute decision making (MADM) problem is the process of finding the best alternative from all of the feasible alternatives where all the alternatives can be evaluated according to a number of attributes. In general, multiple attribute decision making problems include uncertain and imprecise data and information. In this paper, we consider the multiple attribute decision making problems based on linguistic preference information.

Step 1. Let $X=\{x_1, x_2, \ldots, x_m\}$ be a discrete set of alternatives, $C=\{c_1, c_2, \ldots, c_n\}$ be a set of attributes, and $w=(w_1, w_2, \ldots, w_n)^T$ be the weighting vector of attributes, where $w_j \in [0,1]$, and $\sum_{j=1}^{n} w_j = 1$, and, for each alternative $x_i \in X$, the decision maker gives his/her preference value $a_{ij}$ with respect to attribute $c_j \in C$, where $a_{ij}$ takes the form of linguistic variables, that is $a_{ij} \in \tilde{S}$, then all the preference values of the alternatives consist the decision matrix $A=(a_{ij})_{m \times n}$, the information is presented in Table 1.

Step 2. For each attribute, the decision maker gives his/her ideal preference value, which can be seen as the ideal alternative. This information is presented in Table 2.

Step 3. Compare the ideal alternative and the candidate alternative under consideration, and obtain the linguistic distance, then use the linguistic distance operators to derive the collective distance preference values for each alternative $x_i$ according to the ideal alternative.

Step 4. Rank all the alternatives and select the best one(s) according to the results obtained in the previous steps. Note that the smaller linguistic distance value, the better alternative. That is, we rank the alternatives in accordance with linguistic distance value in ascending order.

Step 5. End.

6. Numerical example

Let us suppose an engineering investment company, which wants to invest a sum of money in the best option (adapted from [21]). There is a panel with five possible alternatives in which to invest the money:

1. $x_1$ is a car industry;
2. $x_2$ is a food company;
3. $x_3$ is a computer company;
4. $x_4$ is an arms company;
5. $x_5$ is a TV company.

The engineering investment company must take a decision according to the following four attributes:

1. $c_1$ is the risk analysis;
2. $c_2$ is the growth analysis;
3. $c_3$ is the social-political impact analysis;
4. $c_4$ is the environmental impact analysis.

Table 1. The decision matrix

<table>
<thead>
<tr>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
<th>$c_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$a_{11}$</td>
<td>$a_{12}$</td>
<td>$a_{1j}$</td>
<td>$a_{1n}$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$a_{21}$</td>
<td>$a_{22}$</td>
<td>$a_{2j}$</td>
<td>$a_{2n}$</td>
</tr>
<tr>
<td>$x_3$</td>
<td>$a_{31}$</td>
<td>$a_{32}$</td>
<td>$a_{3j}$</td>
<td>$a_{3n}$</td>
</tr>
<tr>
<td>$x_4$</td>
<td>$a_{41}$</td>
<td>$a_{42}$</td>
<td>$a_{4j}$</td>
<td>$a_{4n}$</td>
</tr>
<tr>
<td>$x_5$</td>
<td>$a_{51}$</td>
<td>$a_{52}$</td>
<td>$a_{5j}$</td>
<td>$a_{5n}$</td>
</tr>
</tbody>
</table>

Table 2. The ideal alternative

<table>
<thead>
<tr>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
<th>$c_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x*$</td>
<td>$a_1$</td>
<td>$a_2$</td>
<td>$a_j$</td>
<td>$a_n$</td>
</tr>
</tbody>
</table>
Table 3. Linguistic decision matrix $A$

<table>
<thead>
<tr>
<th>$x_i$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>$s_1$</td>
<td>$s_1$</td>
<td>$s_0$</td>
<td>$s_1$</td>
</tr>
<tr>
<td>$s_2$</td>
<td>$s_3$</td>
<td>$s_2$</td>
<td>$s_4$</td>
<td>$s_2$</td>
</tr>
<tr>
<td>$s_3$</td>
<td>$s_1$</td>
<td>$s_1$</td>
<td>$s_2$</td>
<td>$s_0$</td>
</tr>
<tr>
<td>$s_4$</td>
<td>$s_1$</td>
<td>$s_0$</td>
<td>$s_1$</td>
<td>$s_2$</td>
</tr>
<tr>
<td>$s_5$</td>
<td>$s_2$</td>
<td>$s_3$</td>
<td>$s_2$</td>
<td>$s_1$</td>
</tr>
</tbody>
</table>

Table 4. The ideal alternative

<table>
<thead>
<tr>
<th>$x^*$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_2$</td>
<td>$s_3$</td>
<td>$s_4$</td>
<td>$s_3$</td>
<td>$s_1$</td>
</tr>
</tbody>
</table>

Table 5. Aggregated results by different linguistic distance operators

<table>
<thead>
<tr>
<th>$x_i$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
<th>Rankings</th>
</tr>
</thead>
<tbody>
<tr>
<td>LWD</td>
<td>$s_{2.1}$</td>
<td>$s_{1.9}$</td>
<td>$s_{1.8}$</td>
<td>$s_{2.4}$</td>
<td>$s_{0.6}$</td>
</tr>
<tr>
<td>LND</td>
<td>$s_{1.75}$</td>
<td>$s_{1.5}$</td>
<td>$s_{1.5}$</td>
<td>$s_2$</td>
<td>$s_{0.5}$</td>
</tr>
<tr>
<td>LOWD</td>
<td>$s_{2.3}$</td>
<td>$s_{2.2}$</td>
<td>$s_{1.6}$</td>
<td>$s_{2.3}$</td>
<td>$s_{0.7}$</td>
</tr>
<tr>
<td>LAOWD</td>
<td>$s_{1.9}$</td>
<td>$s_{1.2}$</td>
<td>$s_{1.2}$</td>
<td>$s_{1.5}$</td>
<td>$s_{0.3}$</td>
</tr>
</tbody>
</table>

The five possible alternatives $x_i$ ($i=1, 2, 3, 4, 5$) are evaluated using the linguistic term set:

$S = \{ s_{-4} = \text{extremely poor}, s_{-3} = \text{very poor}, s_{-2} = \text{poor}, s_{-1} = \text{slightly poor}, s_0 = \text{fair}, s_1 = \text{slightly good}, s_2 = \text{good}, s_3 = \text{very good}, s_4 = \text{extremely good} \}$

by the decision maker under the above four attributes, and construct the decision matrix $A=(a_{ij})_{5 \times 4}$ as listed in Table 3.

Comparing the ideal alternative and the candidates considered using the linguistic distance operators. We will consider the LWD, LND, LOWD, LAOWD operators, suppose that the weighting vector of four attributes is $w=(0.3, 0.4, 0.2, 0.1)^T$. Then, we get the ranking results, which are listed in Table 5. Note also that “$>$” means “preferred to” and “$\sim$” means “equal to”. We find that even though the rankings are different by different operators, but in all the rankings, $x_5$ is the best alternative, and $x_4$ is the worst one.

7. Concluding remarks

In this paper, we have developed some linguistic distance operators, such as linguistic weighted distance (LWD) operator, linguistic ordered weighted distance (LOWD) operator, and studies some of their desired properties, such as commutativity, monotonicity, ideempotency, bounded, etc. We also investigate some families of the LOWD operator. We develop a procedure to the linguistic decision problem with the developed linguistic distance operators. Finally, an engineering investment example is given to illustrate the multiple attribute group decision making process.

In the future, we will develop other extensions of the distance measures to the linguistic environment, such as the use generalized and quasi-arithmetic means. We will also investigate the potential applications of the developed linguistic distance operators to other fields, such as pattern recognition, supply chain management, image process, engineering evaluation, etc.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No.50979024, No. 90924027).
References