Weighted Rank Correlation Measures Based on Fuzzy Order Relations

Sascha Henzgen and Eyke Hüllermeier

Department of Mathematics and Computer Science,
Marburg University, Germany
{henzgen, eyke}@mathematik.uni-marburg.de

Abstract

In this paper, we propose a weighted rank correlation measures on the basis of fuzzy order relations. Our measure, called *scaled gamma*, is related to Goodman and Kruskal's gamma rank correlation and parametrized by a fuzzy equivalence relation on the rank positions. To specify this relation in a convenient way, we make use of so-called scaling functions. The usefulness of our approach is indicated in a first experimental study, in which we analyze a real-world ranking data set.

1 Introduction

Rank correlation measures such as Kendall's tau [1] and Spearman's rho [2], which have originally been developed in statistics, are becoming increasingly important in fields like information retrieval and preference learning [3]. In particular, such measures are commonly used as performance metrics to quantify the similarity between an ideal or "true" ranking of a set of items and a prediction produced by a ranking algorithm or learning-to-rank method. Formally, a rank correlation measure is a mapping $\mathbb{S}_N \times \mathbb{S}_N \to [-1, +1]$, with $\mathbb{S}_N$ the set of permutations of $[N] = \{1, 2, \ldots, N\}$ and N the number of items to be ranked.

In many applications, such as Internet search engines, one is not equally interested in all parts of a predicted ranking. Instead, accuracy in the top of the ranking, i.e., accurate predictions of the ranks of the top-items, is typically considered more important than accuracy in the middle and bottom part. Standard rank correlation measures, however, put equal emphasis on all positions and, therefore, are not able to make distinctions of that kind. This is why *weighted* variants have been proposed for some correlation measures, as well as alternative measures specifically focusing on the top of a ranking, such as the Discounted Cumulative Gain [4].

In this paper, we develop a formal framework for designing weighted rank correlation measures based on the notion of *fuzzy order relation*. More specifically, the key ingredients of our approach are as follows:

- *Fuzzy order relations* [5] are generalizations of the conventional order relations on the real numbers: `SMALLER`, `EQUAL` and `GREATER`. They enable a smooth transition between these predicates and allow for expressing, for instance, that a number x is smaller than y *to a certain degree*, while to some degree these numbers are also considered as being equal. Here, the `EQUAL` relation is understood as a kind of similarity relation that seeks to model the "perceived equality" (instead of the strict mathematical equality).

- *Fuzzy rank correlation* [6, 7] generalizes conventional rank correlation on the basis of fuzzy order relations, thereby combining properties of standard rank correlation (such as Kendall's tau) and conventional numerical correlation measures (such as Pearson correlation). Fuzzy rank correlation measures are especially useful in cases where computations with numerical differences are meaningful within a certain range but less reasonable if these differences are too small.

- *Scaling functions* for modeling fuzzy equivalence relations [8]. For each element x of a linearly ordered domain X, a scaling function $s(\cdot)$ essentially expresses the degree $s(x)$ to which x can be (or should be) distinguished from its neighboring values. A measure of distance (or, equivalently, of similarity) on X can then be derived via accumulation of local degrees of distinguishability.

The rest of the paper is organized as follows. In the next two sections, we briefly recall the basics of fuzzy order relations and fuzzy rank correlation, respectively. Our scaled gamma measure is then introduced in Section 5. A small experimental study is presented in Section 6, prior to concluding the paper in Section 7.

2 Fuzzy Order Relations

A fuzzy relation $E : \mathbb{X} \times \mathbb{X} \to [0, 1]$ is called *fuzzy equivalence* with respect to a t-norm $\top$, for brevity $\top$-equivalence, if it is

- reflexive: $E(x, x) = 1$,
- symmetric: $E(x, y) = E(y, x)$,
- and $\top$-transitive: $\top(E(x, y), E(y, z)) \leq E(x, z)$.

Moreover, a fuzzy relation $L : \mathbb{X} \times \mathbb{X} \to [0, 1]$ is called *fuzzy ordering* with respect to a t-norm $\top$ and a $\top$-equivalence E, for brevity $\top$-E-$ordering$, if it is

- E-reflexive: $E(x, y) \le L(x, y)$,
- $\top$-E-antisymmetric: $\top(L(x, y), L(y, x)) \le E(x, y)$,
- and $\top$-transitive: $\top(L(x, y), L(y, z)) \le L(x, z)$.

We call a $\top$-E-ordering L, which plays the role of a `LESS` or `EQUAL` relation, *strongly complete* if $\max(L(x, y), L(y, x)) = 1$ for all $x, y \in \mathbb{X}$.

Finally, let R denote a strict fuzzy ordering associated with a strongly complete $\top$-E-ordering L. As will be seen below, the well-known Łukasiewicz t-norm $\top(x, y) = \max(0, x + y - 1)$ will be most relevant for us. In this case, R can simply be taken as $R(x, y) = 1 - L(y, x)$ [5]. Obviously, R is playing the role of a (strictly) `GREATER` relation.

3 Fuzzy Rank Correlation

In this section, we briefly recall Goodman and Kruskal's gamma rank correlation measure [9] as well as its fuzzy extension as proposed by Bodenhofer and Klawonn [6] and further analyzed by Ruiz and Hüllermeier [7].

Consider $n \ge 2$ paired observations $\{(x_i, y_i)\}_{i=1}^N \subset \mathbb{X} \times \mathbb{Y}$ of two variables X and Y, where $\mathbb{X}$ and $\mathbb{Y}$ are two linearly ordered domains; we denote $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_N)$. The goal of a rank correlation measure is to capture the dependence between the two variables in terms of their tendency to increase and decrease in the same or the opposite direction. If an increase in X tends to come along with an increase in Y, then the (rank) correlation is positive. The other way around, the correlation is negative if an increase in X tends to come along with a decrease in Y. If there is no dependency of either kind, the correlation is (close to) 0.

Several rank correlation measures are defined in terms of the number C of *concordant*, the number D of *discordant*, and the number of *tied* data points. For a given index pair $(i, j) \in [N]^2$, we say that (i, j) is concordant, discordant or tied depending on whether $(x_i - x_j)(y_i - y_j)$ is positive, negative or 0, respectively. A well-known example is Goodman and Kruskal's *gamma rank correlation* [9], which is defined as

$$\gamma = \frac{C - D}{C + D} \quad .$$

Now, assume two $\top$-equivalences $E_{\mathbb{X}}$ and $E_{\mathbb{Y}}$ to be given on $\mathbb{X}$ and $\mathbb{Y}$, respectively, as well as strict fuzzy order relations $R_{\mathbb{X}}$ and $R_{\mathbb{Y}}$. Using these relations, the concepts of concordance and discordance of data points can be generalized as follows [6]: Given an index pair (i, j), the degree to which this pair is concordant, discordant, and tied is defined, respectively, as

$$\hat{C}(i,j) = \top(R_{\mathbb{X}}(x_i, x_j), R_{\mathbb{Y}}(y_i, y_j)), \tag{1}$$

$$\hat{D}(i,j) = \top(R_{\mathbb{X}}(x_i, x_j), R_{\mathbb{Y}}(y_j, y_i)), \tag{2}$$

$$\hat{T}(i,j) = \bot(E_{\mathbb{X}}(x_i, x_j), E_{\mathbb{Y}}(y_i, y_j)), \tag{3}$$

where $\top$ is a t-norm and $\bot$ is the dual t-conorm of $\top$ (i.e., $\bot(x, y) = 1 - \top(1 - x, 1 - y)$). With $\tilde{C}(i,j) = \hat{C}(i,j) + \hat{C}(j,i)$ the degree of concordance of the index pair (i, j) and $\tilde{D}(i,j) = \hat{D}(i,j) + \hat{D}(j,i)$ the degree of discordance, the following equality holds for all (i, j):

$$\tilde{C}(i,j) + \tilde{D}(i,j) + \hat{T}(i,j) = 1.$$

Adopting the simple sigma-count principle to measure the cardinality of a fuzzy set, the number of concordant and discordant pairs can be computed, respectively, as

$$\tilde{C} = \sum_{1 \le i < j \le N} \tilde{C}(i,j), \qquad \tilde{D} = \sum_{1 \le i < j \le N} \tilde{D}(i,j).$$

The *fuzzy ordering-based* gamma rank correlation measure $\tilde{\gamma}$, or simply "fuzzy gamma", is then defined as

$$\tilde{\gamma} = \frac{\tilde{C} - \tilde{D}}{\tilde{C} + \tilde{D}} \ . \tag{4}$$

From this definition, it is clear that the basic idea of the fuzzy gamma is to decrease the influence of "close-to-tie" pairs (x_i, y_i) and (x_j, y_j). Such pairs, whether concordant or discordant, are turned into a partial tie, and hence are ignored to some extent. Or, stated differently, there is a smooth transition between being concordant (discordant) and being tied.

4 The Scaled Gamma Measure

In the previous section, $\mathbb{X}$ and $\mathbb{Y}$ could be any linearly ordered domains. Coming back to our original goal of a (weighted) comparison of two rankings, we now take $\mathbb{X} = \mathbb{Y} = [N]$, i.e., both domains are given by the set of rank positions. Then, the relation $E = E_{\mathbb{X}} = E_{\mathbb{Y}}$ is simply expressing to what degree two ranks are considered as being equivalent, i.e., as not being distinguished. The basic idea is to modulate the influence of a concordance or discordance involving two rank positions according to the degree of indistinguishability of these positions.

The fuzzy equivalence E can be defined quite conveniently by means of a so-called *scaling function* [8]. For each element of a linearly ordered domain, a scaling function $s(\cdot)$ essentially expresses the degree to which this element can be (or should be) distinguished from its neighboring values. A measure of distance (or, equivalently, of similarity) on the whole domain can then be derived via accumulation (summation or integration) of local degrees of distinguishability.

More concretely, a scaling function $s'(\cdot)$ can be defined on $\mathbb{X} = [N]$ in our case, with $s'(x)$ indicating the degree to which rank x ought to be distinguished from its neighbor rank $x + 1$. These degrees of distinguishability between neighbors can then be extended to degrees of distinguishability between any pair of ranks x and y via

$$\Delta(x,y) = \sum_{r=\min(x,y)}^{\max(x,y)-1} s'(r) \quad . \tag{5}$$

As we find it more convenient to work with a continuous scaling function $s : [1, N] \rightarrow \mathbb{R}_+$, we define $s'(\cdot)$ via $s(\cdot)$ as follows:

$$s'(x) = \int_{x}^{x+1} s(z)\, dz \tag{6}$$

The function $s(\cdot)$ *scales* the positions of the ranks in the sense of replacing the constant distance 1 between two adjacent ranks x and $x+1$ by $s'(x)$ as defined in (6). Thus, (5) becomes

$$\Delta(x,y) = \int_{\min(x,y)}^{\max(x,y)} s'(z)\, dz \quad . \tag{7}$$

Based on this distinguishability relation, we define the relation E by

$$E(x,y) = \max\big(1 - \Delta(x,y), 0\big) \quad .$$

From the results of [8], it follows that E thus defined is indeed a proper fuzzy equivalence relation, namely a $\top$-equivalence with $\top$ the Łukasiewicz t-norm. Thus, the fuzzy gamma coefficient as introduced in the previous section can be derived on the basis of this fuzzy equivalence on ranks.

More concretely, consider two rankings specified, respectively, by permutations $\pi, \pi' : [N] \to [N]$, where $\pi(i)$ denotes the position of object i in the ranking defined by π. The pair of objects (i, j) is concordant if $(\pi(i) - \pi(j))(\pi'(i) - \pi'(j)) > 0$. However, instead of counting an increment of $+1$, the overall level of concordance D is increased by

$$m_{i,j} = \max\left(1 - E(\pi(i), \pi(j)) - E(\pi'(i), \pi'(j)), 0\right) \ ,$$

which can be seen as the *weight* attached to the pair (i, j). Thus, full concordance is counted only of the ranks $\pi(i)$ and $\pi(j)$ as well as the ranks $\pi'(i)$ and $\pi'(j)$ are sufficiently distinguished. Otherwise, the degree of concordance is reduced in correspondence with the respective degrees of indistinguishability (equivalence). Computing the overall level of discordance in the same way, we eventually obtain our weighted rank correlation measure, which we call *scaled gamma*, as the fuzzy gamma (4) with

$$\tilde{C}(i,j) = \begin{cases} m_{i,j} & \text{if } (\pi(i) - \pi(j))(\pi'(i) - \pi'(j)) > 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\tilde{D}(i,j) = \begin{cases} m_{i,j} & \text{if } (\pi(i) - \pi(j))(\pi'(i) - \pi'(j)) < 0 \\ 0 & \text{otherwise} \end{cases}$$

This measure has a number of desirable properties. In particular, the scaled gamma inherits all formal properties of $\tilde{\gamma}$ as shown in [7]. Moreover, several interesting measures can be recovered as special cases. For example, the original gamma correlation measure is obtained for the scaling function $s \equiv 1$ (all ranks are fully distinguished, i.e., $E(x, y) = 0$ for all $x \neq y$). Another interesting special case is obtained for the top-k scaling function $s(x) = 1$ for $x \leq k$ and $s(x) = 0$ for $x > k$. Then, $m_{i,j} = 1$ unless either $\pi(i)$ and $\pi(j)$ or $\pi'(i)$ and $\pi'(j)$ both exceed k, in which case $m_{i,j} = 0$. In other words, concordances and discordances in the bottom part of the ranking are completely ignored.

5 Illustration

To show the usefulness of our weighted rank correlation measure, we conducted experiments with a data set introduced in [10]. The data originates from a study in which 409 students aged 15–18 were asked to rank 16 goals such as “I want to feel calm, at ease” and “I want to get along with my parents” according to how important they personally find these goals. Thus, each student is characterized by a ranking of 16 goals. An obvious idea, then, is to define a degree of similarity between two students in terms of the correlation between the corresponding rankings.

We computed binary similarity relations (409×409 matrices) using Kendall's tau and Spearman's rho as correlation measures. Moreover, as it is plausible that higher ranks carry more information than lower ones (while reliably sorting the top 3 to 5 goals might be feasible, the remaining ones are likely to be put in a more or less random order), we also applied weighted rank correlation measures that put more emphasis on the top-ranks. More specifically, we used the Canberra distance [11], a weighted variant of Spearman's footrule, and a weighted version of Kendall's tau introduced by Kumar and Vassilvitskii [12].

Finally, we also applied our scaled γ with a parametrized scaling function

$$s_M(x) = \frac{1-c}{1+\exp(a(x-b))} + c \ . \tag{8}$$

Figure 1 illustrates the influence of the three parameters a, b, and c on the shape of the function. The weighted Kendall measure has (weight) parameters δ_i, too. Since these parameters are serving a purpose very similar to the scaled distance between adjacent ranks in our case, we set them to $\delta_i = \int_i^{i+1} s_M(x)\, dx$.

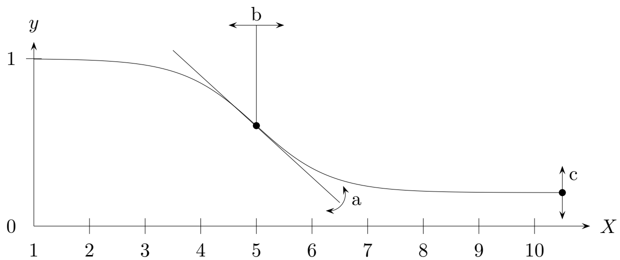

Figure 1: Parametrizable scaling function.

We embedded the 409 data objects (students) in $\mathbb{R}^2$ by applying kernel principle component analysis (kPCA) [13] to the binary similarity relations. The results are shown in Figure 2 in the form of a scatter plot. As can be seen, neither Kendall and Spearman nor Canberra are able to produce any structure in the data. More structure does clearly become visible by choosing appropriate parameters for the two parametrized measures, the weighted Kendall and our scaled gamma—the increased flexibility of these measures is obviously an advantage. The results of our scaled gamma are perhaps the most interesting ones. In fact, the scaled gamma is able to produce well-defined clusters that arguably correspond to subgroups of students with similar moral concepts.

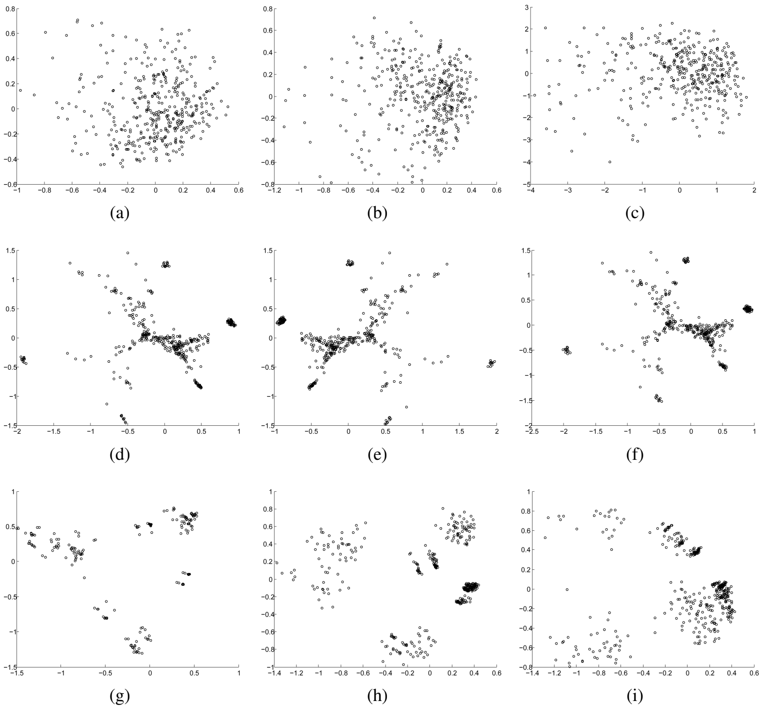

Figure 2: Two-dimensional embedding of similarity relations by means of kernel PCA. Similarities were produced using Kendall's tau (a), Spearman's rho (b), Canberra (c), weighted Kendall (d, e, f) and scaled gamma (g, h, i). The three versions of the last two measures were produced using the scaling function s_M with the parameters $a = 10$, $b = 3$, $c = 0$ in (d) and (g), $c = 0.0125$ in (e) and (h), $c = 0.025$ in (f) and (i).

6 Conclusion

We introduced a weighted rank correlation measure called *scaled gamma*. This measure is a special case of a recently proposed generalization of the gamma rank correlation based on fuzzy order relations. The scaled gamma allows for specifying the importance of rank positions in a quite convenient way by means of a scaling function. Thanks to the underlying formal foundation, such a scaling function immediately translates into a concrete version of our measure, in which the rank positions are processed within an appropriate weighting scheme.

The usefulness of the scaled gamma was indicated in a first experimental study, in which we analyzed a real-world ranking data set. Here, interesting structure in the data could be revealed through kernel PCA when using the scaled gamma as a correlation measure, whereas no such structure is uncovered by standard rank correlation measures.

More applications of that kind will be considered in future work. Apart from that, we are planning a detailed formal comparison of the scaled gamma and other weighted rank correlation measures.

References

[1] Kendall, M.: *Rank correlation methods*. London: Charles Griffin. 1955.

[2] Spearman, C.: The proof and measurement for association between two things. *Amer. Journal of Psychology* 15 (1904), S. 72–101.

[3] Fürnkranz, J.; Hüllermeier, E. (Hg.): *Preference Learning*. Springer-Verlag. 2011.

[4] Jarvelin, K.; Kekalainen, J.: Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems* 20 (2002) 4, S. 422–446.

[5] Bodenhofer, U.; Demirci, M.: Strict Fuzzy Orderings with a Given Context of Similarity. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems* 16 (2008) 2, S. 147–178.

[6] Bodenhofer, U.; Klawonn, F.: Robust rank correlation coefficients on the basis of fuzzy orderings: Initial steps. *Mathware & Soft Computing* 15 (2008), S. 5–20.

[7] Ruiz, M. D.; Hüllermeier, E.: A Formal and Empirical Analysis of the Fuzzy Gamma Rank Correlation Coefficient. *Information Sciences* 206 (2012), S. 1–17.

[8] Klawonn, F.: Fuzzy Sets and Vague Environment. *Fuzzy Sets and Systems* 66 (1994), S. 207–221.

[9] Goodman, L.; Kruskal, W.: *Measures of Association for Cross Classifications*. New York: Springer-Verlag. 1979.

[10] Boekaerts, M.; Smit, K.; Busing, F.: Salient goals direct and energise students' actions in the classroom. *Applied Psychology: An International Review* 4 (2012) S1, S. 520–539.

[11] Lance, G.; Williams, W.: Computer programs for hierarchical polythetic classification ("similarity analyses"). *The Computer Journal* 9 (1966), S. 60–64.

[12] Kumar, R.; Vassilvitskii, S.: Generalized distances between rankings. In: *Proc. WWW, 19. International Conference on World Wide Web*, S. 571–580. 2010.

[13] Schölkopf, B.; Smola, A.; Müller, K.: Kernel principal component analysis. In: *Advances in Kernel Methods: Support Vector Learning*, S. 327–352. MIT Press. 1999.