Evaluating the Quality Level of Projects, Authors and Experts

Vorobev Alexandr

Moscow State University, Faculty of Mathematics and Mechanics, Department of Probability Theory
Witology
alexandr.vorobyev@witology.com
http://mech.math.msu.su/probab/index-e.html
http://witology.com/

Abstract. There are some situations when we need to compare work results of different people. For example, it may be an announced tender, a grants' distribution or a competition in poetry. And sometimes it is allowed for one participant to present more than one work. It is extremely actual when we want not only to identify the winner but also to get a lot of effective projects. We don't want to bound the author's creativity therefore. Then having expert evaluations we must obtain integral evaluation for each work. How to do this? Also it may be useful to evaluate author's level on the basis of his works' evaluations. Moreover, we may want to evaluate expert's level on the basis of his evaluations. In this paper we solve these problems by constructing a Bayesian network model for such a competition and applying a maximum a posteriori estimation method to it.

Keywords: Bayesian network, maximum a posteriori estimation, maximal likelihood method, expert evaluations, expert competency, parameter evaluation.

1 Model description

Let us take a competition between projects, which allows an author to present more than one project. The participants present their projects and experts evaluate these projects, not more than one evaluation given by one expert for one project.

The aim is to estimate level of each project, level of each expert and level of each author on the basis of expert evaluations.

We introduce the following model of the competition.

1) Each author has his level C(Creator) - degree of his ability to create the effective projects.

2) Each expert has his level E (Evaluator) - degree of his ability to give for a project the proper evaluation.

3) Each project has its level L(Level) - degree of its effectiveness. This level is random value and appears when its author creates this project. We also assume that L has a distribution depended on its author level C - $Rand_L(C)$.

4) Each project evaluation e given by an expert is also a random value and its distribution depends on the project level and on the expert level - $Rand_e(L, E)$.

5) The number of projects presented by author is independent from his level.

Distributions $Rand_L(C)$ and $Rand_e(L, E)$ may be fixed if appropriate parameters are fixed or may have another parameters for which we don't have prior exact values. Then in general case we have distributions $Rand_L(C, P)$ and $Rand_e(L, E, Q)$ where P and Q are vectors of parameters.

Thus we can formulate the following problem:
Given 3-dimensional vector of project evaluations

$$(e_{i,j,k})_{i=1..n, j=1..n, k=1..m_j},$$

where $e_{i,j,k}$ - the evaluation given by the expert i to the k^{th} project of the author j (we number projects of one author beginning from the one);

m_j - number of the projects of the author j.

The evaluations of the following values have to be found:

$(C_j)_{j=1..n}$ - vector of the author levels;

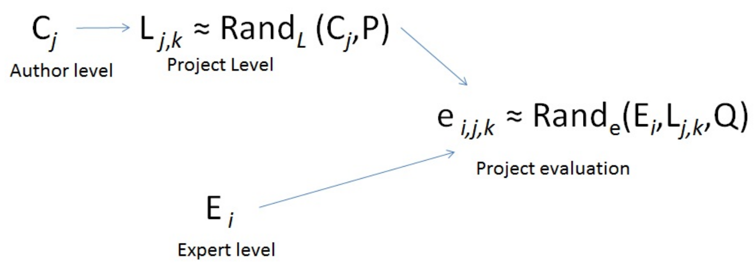
$(E_j)_{j=1..n}$ - vector of the expert levels;

$(L_{j,k})_{j=1..n, k=1..m_j}$ - 2-dimensional vector of the project levels;

P - vector of the distribution parameters $Rand_L$;

Q - vector of the distribution parameters $Rand_e$.

Thus we have constructed a Bayesian network as a model for our competition.



2 Method of problem solving

It may be interesting to find posterior distribution for unknown parameters in this model too, but here we are seeking for parameters' estimates only. The standard method for this problem is the maximum a posteriori estimation method. We do not have any special suggestions about the parameter distribution. Therefore we consider the prior parameter distribution to be uniform. In this case maximum a posteriori estimates coincide with maximal likelihood estimates.

It means that the estimate of a parameters' vector equals such a value, that conditional probability of the fact, that the expert evaluations are equal to their observed values

$$(e_{i,j,k})_{i=1..n,j=1..n,k=1..m_j},$$

is maximal.

Up to this moment all our values could lay in any set. But we need to specify the kind of these sets to write the formulae. We consider the finite sets due to the following reason. It is well known (see [1]) that it is natural for an expert to say which of two projects is better relative to a property or group of properties but it is not natural for him to give quantitative evaluation of project. A finite set for evaluations is closer to the theory than a continuous or discrete infinite set.

Then it is natural for evaluations $e_{i,j,k}$ and project levels $L_{j,k}$ to lay in the same set. Then evaluations can be considered as project level evaluations.

Let us also consider the values C, E, P, Q to be discrete. It makes the results more simple to be interpreted (values C, E) and makes the narrative more simple.

To simplify our figures we drop indexes in the names of vectors if they may have all the possible values. So we write

$(A_{i,j,k})_{k=1..m_j}$ instead of $(A_{i,j,k})_{i=1..n,j=1..n,k=1..m_j}$

and $(A_{i,j})$ instead of $(A_{i,j})_{i=1..n,j=1..n}$.

Also we mean that the index under the operator of sum or multiplication runs through the whole set of its possible values.

Then maximum a posteriori estimates have the following form:

$$\begin{aligned}
&((\widetilde{L_j}), (\widetilde{C_j}), (\widetilde{E_j}), \widetilde{P}, \widetilde{Q}) = \\
&= arg \max_{(L_{j,k})_{k=1..m_j},(C_j),(E_j),P,Q} P((e_{i,j,k})_{k=1..m_j} | (L_{j,k})_{k=1..m_j}, (C_j), (E_j), P, Q) = \\
&= arg \max_{(L_{j,k})_{k=1..m_j},(C_j),(E_j),P,Q} \prod_{i,j,k} (P(e_{i,j,k}|E_i, Q, L_{j,k}) P(L_{j,k}|C_j, P)) = \\
&= arg \max_{(L_{j,k})_{k=1..m_j},(C_j),(E_j),P,Q} \sum_{i,j,k} (lnP(e_{i,j,k}|E_i, Q, L_{j,k}) + lnP(L_{j,k}|C_j, P))
\end{aligned} \tag{1}$$

It is a problem of discrete optimization which is NP-complete. It means that the complexity of computations of an exact answer is huge for the big number of parameters. Hence we have to use optimization methods which give an approximate answer.

There are different optimization algorithms for this problem, for example the expectation–maximization algorithm, the gradient method, the analogue of belief propogation algorithm. The choice of such a method depends on the model. Probably we should try different methods to understand which is better in our case.

3 Model specification

Here we specify the distribution functions $Rand_L(C, P)$ and $Rand_e(L, E, Q)$. As we decided random values $L_{j,k}$ and $e_{i,j,k}$ are distributed on the same finite set. This set has to be small because of the big complexity of the optimization problem solving. In our data this set is the set if integers from the interval $[-3, 3]$.

Then we consider the following assumptions to be natural:

1) The mathematical expectation of the expert evaluation $e_{i,j,k}$ equals the true level of this project $L_{j,k}$.

2) Its volatility depends on the expert level: the higher the expert level, the less the volatility. It means if we give the big set of projects of the same level to the expert (even bad) the average of his evaluations will be close to this true level. But the better the expert the less standard deviation of his evaluations will be.

It is the best practice to use normal distribution in the models due to the existence of the central limit theorem. But we have discrete sets of variables. Hence we use the value of normal density divided by normalizing constant as the probability function value for each discrete point. Then the probability function has the bell curve shape.

Now we have to define the volatility as function of the expert level. Let us take volatility to be equal to $\frac{q}{E_i}$. We need constant q because E_i is bounded (it lays in the finite set), but we don't know the bound for volatility - it depends on the data. And we suppose that E_i is integer from the interval [0,3]. The vector Q consists of the only value q in this case.

Thus value $e_{i,j,k}$ is distributed according to the law

$$P(e_{i,j,k} = a | L_{j,k}, E_i) = \frac{e^{-\frac{(a-L_{j,k})^2 E_i^2}{2q^2}}}{\sum\limits_{b=-3}^{3} e^{-\frac{(b-L_{j,k})^2 E_i^2}{2q^2}}}, a = -3..3 \quad (2)$$

We define the same discretized normal distribution for the project level $L_{j,k}$ where the mathematical expectation equals the author level C_j and the volatility equals the constant parameter p. It means the higher the author level the higher the average level of his projects. Volatility is constant because we don't know any facts about the dependence between it and the author level.

Then C_j is the integer from the interval [-3,3] and the vector P consists of the only value p in this case.

Thus value $L_{j,k}$ is distributed according to the law:

$$P(L_{j,k} = a | C_j) = \frac{e^{-\frac{(a-C_j)^2}{2p^2}}}{\sum\limits_{b=-3}^{3} e^{-\frac{(b-C_j)^2}{2p^2}}}, a = -3..3 \quad (3)$$

4 Relation with other problems

There is a standard problem which can be formulated in the terms of this paper as the following: it is necessary to evaluate the project levels and the expert levels on the basis of the evaluations given by these experts to these projects.

Except our attempt to evaluate the author level there is the principal difference between our approach to this problem and standard approach. In our problem we have authors of the projects. And we think that the projects of one author will more probably be similar in their levels than the projects of two different authors. This fact underlies our model.

The standard approach (see [4]) consists in applying an iteration procedure. On each step for each expert we measure the closeness between his project evaluations and the current weighted-average project evaluations where the weight of the expert evaluation equals the expert level. Then the expert levels are re-obtained on the basis of this closeness and so on.

5 Conclusions

The paper presents a new model of a competition. Within this model a maximum a posteriori estimation can be applied to estimate the project levels, the project author levels and the expert levels on the basis of the evaluations given by these experts to these projects.

Notice that the presented algorithm does not have any requirements for the existence of the evaluation given by the definite expert to the definite project or for the number of evaluations per one project. Another question is that the more evaluations we have the more exact the results of the algorithm are.

We are going to apply this method to the data in the nearest future. Then we will be able to compare this algorithm with the standard approach in estimation of $((L_{j,k})_{k=1..m_j}, (E_j))$.

Also we are going to consider this model when the authors are the experts at the same moment. Then in our model we can take into account a correlation between participant level as an author and his level as an expert. Or we can evaluate this correlation on the basis of the results of this method if we don't take it into account in the model.

References

1. Orlov, A.: Expert Evaluations. Moscow (2002)
2. Lemeshko, B., Denisov, V., Postovalov, S.: Applied Statistics. Rules of Testing of Consensus between Sampling and Theoretical Distributions. Methodical recommendations. Part I. Chi-squared-type Tests (in Russian). NSTU, Novosibirsk (1998)
3. Orlov, A.: About Tests of Consensus with Parametrical Distribution Family. Zavodskaya laboratoriya. Diagnostika materialov/ 63,№5, 49-50 (1997)
4. Vidyapin, V.: Bachelor of Economics. Vol. 2. Triada, Moscow (1999)
5. Heckerman D.: A Tutorial on Learning With Bayesian Networks. Technical Report MSR-TR-95-06, Microsoft Research, Advanced Technology Division (1996)