ON COORDINATION OF EXPERTS' ESTIMATIONS OF QUANTITATIVE VARIABLE*

Gennadiy Lbov, Maxim Gerasimov

Abstract: *In this paper, we consider some problems related to forecasting of quantitative feature. We assume that decision rule is constructed on the base of analysis of empirical information represented in the form of statements from several experts. The criterion of a quality of experts' statements is suggested. The method of forming of united expert decision rule is considered.*

Keywords: *expert statements, coordination.*

ACM Classification Keywords: *I.2.6. Artificial Intelligence - knowledge acquisition.*

Conference: *The paper is selected from International Conference "Classification, Forecasting, Data Mining" CFDM 2009, Varna, Bulgaria, June-July 2009*

Introduction

In this work we assume that objects under investigation are described by some set of qualitative and quantitative features, and some independent experts give predictions of estimated quantitative feature. Their statements may be partially or completely identical, supplementary, and/or contradictory. Also, experts' statements may vary from time to time as well as new "knowledge" from new experts may be obtained. Hence, decision rule is constructed on the base of analysis of empirical information, represented in the form of several experts' statements. Obtained decision rule must be free from anomalies as conflict and redundancy.

Setting of a Problem

Let Γ be a population of elements or objects under investigation. By assumption, L experts give estimations of values of unknown quantitative feature Y for objects $a \in \Gamma$, being already aware of their description $X(a)$. We assume that $X(a) = (X_1(a),..., X_j(a),..., X_n(a))$, where the set X may simultaneously contain qualitative and quantitative features X_j, $j = \overline{1,n}$. Let D_j be the domain of the feature X_j, $j = \overline{1,n}$; D_Y be the domain of the quantitative feature Y, $D_Y = [\alpha, \beta] \subset R$. In this paper we assume that the feature space D is a subset of the product set $\prod_{j=1}^{n} D_j$.

Note that D may be not equal to $\prod_{j=1}^{n} D_j$.

<u>Example.</u> $D_1 = \{a,b,c,d\}$, $D_2 = [10,20]$, $D = [a,c] \times [10,15] \bigcup [b,d] \times [12,20]$.

We shall say that a set E is a *rectangular set* in D if $E = \prod_{j=1}^{n} E_j$, $E_j \subseteq D_j$, $E_j = [\alpha_j, \beta_j]$ if X_j is a quantitative feature, E_j is a finite subset of feature values if X_j is a nominal feature.

In this paper, we consider statements S^i, $i = \overline{1,M}$; represented as sentences of type "if $X(a) \in E^i$, then $Y(a) = y^i$", where E^i is a rectangular set in D. By assumption, each statement S^i has its own weight w^i

* The work was supported by the RFBR under Grants N07-01-00331a, 08-07-00136a.

($0 < w^i \le 1$ for individual statements). Such a value is like a measure of “confidence”. Each statement S^i corresponds to $\langle l^i, E^i, y^i, w^i \rangle$, where l^i is a code of expert from whom statement is obtained.

Without loss of generality we may assume that experts themselves have equal “weights”.

Denote the initial sets of statements obtained from l-th expert by Ω^l, the set of initial statements from all experts by Ω, $\Omega = \bigcup_{l=1}^{L} \Omega^l$.

The problem consists in constructing decision rule that reflects information synthesized from an organized group of expert opinions.

On Criterion of a Quality of Experts’ Statements

Let $y_0(x)$ be the value of the feature Y at the point $x \in D$, i.e. $y_0(x) = Y(a)$ if $X(a) = x$. Let $y_l(x)$ be the estimation of the $y_0(x)$ made by l-th expert.

We shall say that the set of the values $y_0(x)$ on D is a *strategy of nature* (denote it by c), and the set of the values $y_l(x)$ on D is a *strategy of* l *-th expert* (denote it by g_l).

In this paper we assume for simplicity that there exists rectangular sets $V^1, \ldots, V^{T_l} \subseteq D$ such that $D = \bigcup_{t=1}^{T_l} V^t$, $V^{t_i} \bigcap V^{t_j} = \varnothing$ if $i \neq j$, $y_l(x) \equiv \beta^t \ \ \forall x \in V^t$, where β^t is a constant.

Thus, we assume that the strategies g_l are piecewise constant in D.

Consider value h such that $0 \le h \le 1$. We shall say that l-th expert (a strategy g_l) has a *competence* h if $\frac{|y_0(x) - y_l(x)|}{\beta - \alpha} \le 1 - h \ \ \forall x \in D$.

Define the criterion of a quality of strategy g_l as the integral

$$\eta(g_l) = \frac{\int_D (y_0(x) - y_l(x))^2 dx}{(\beta - \alpha)^2 \mu(D)},$$

where $\mu(D)$ is a measure of the set D.

Consider strategies $g_1, \ldots, g_m$. Let A be some algorithm of constructing decision rule on the base of these strategies. Denote the resulted strategy by g^A, $g^A = A(g_1, \ldots, g_m)$.

We shall say that an algorithm A is a *linear combination* of strategies $g_1, \ldots, g_m$ if $\exists \lambda_1, \ldots, \lambda_m \ge 0$ such that $\sum_{l=1}^{m} \lambda_l = 1$, $y^A(x) = \sum_{l=1}^{m} \lambda_l y_l(x) \ \ \forall x \in D$.

Proposition 1. If strategies $g_1, \ldots, g_m$ have a competence h, then their linear combination has a competence at least equal to h.

Proof. Take any point $x \in D$. Then

$$\left| y_0(x) - y^A(x) \right| = \left| y_0(x) - \sum_{l=1}^{m} \lambda_l y_l(x) \right| = \left| \sum_{l=1}^{m} \lambda_l y_0(x) - \sum_{l=1}^{m} \lambda_l y_l(x) \right| \le \sum_{l=1}^{m} \lambda_l \left| y_0(x) - y_l(x) \right| \le 1 - h.$$

■

Proposition 2. There exists an algorithm A such that for any strategies g_1 and g_2 we have

$$\eta(A(g_1, g_2)) \le \frac{\eta(g_1) + \eta(g_2)}{2}.$$

Proof. Consider algorithm A such that $y^A(x) = \dfrac{y_1(x) + y_2(x)}{2} \quad \forall x \in D$.

Since strategies g_l are piecewise constant in D, strategy g^A is piecewise constant in D.

Take any point $x \in D$. Then

$$\left(y_0(x) - \frac{y_1(x) + y_2(x)}{2} \right)^2 = \frac{1}{4}\left(y_0(x) - y_1(x) + y_0(x) - y_2(x) \right)^2 =$$

$$= \frac{1}{2}\left(y_0(x) - y_1(x) \right)\left(y_0(x) - y_2(x) \right) + \frac{1}{4}\left(y_0(x) - y_1(x) \right)^2 + \frac{1}{4}\left(y_0(x) - y_2(x) \right)^2 \le$$

$$\le \frac{\left(y_0(x) - y_1(x) \right)^2 + \left(y_0(x) - y_2(x) \right)^2}{2}.$$

■

Proposition 3. There exists an algorithm A such that for any strategies $g_1, \ldots, g_m$ we have

$$\eta(A(g_1, \ldots, g_m)) \le \frac{\eta(g_1) + \ldots + \eta(g_m)}{m}.$$

Proof. Consider algorithm A such that $y^A(x) = \dfrac{y_1(x) + \ldots + y_m(x)}{m} \quad \forall x \in D$.

Further proof is similar to the proof of Proposition 2.

■

Note that equality in Proposition 3 is obtained if and only if $y_1(x) \equiv \ldots \equiv y_m(x) \quad \forall x \in D$.

Suppose that strategy of nature c is unknown and there are independent experts with the same competence. From propositions 1 and 3 it follows that the decision rule obtained by the considered algorithm A has at least the same competence and the quality better than average experts' quality.

Proposition 4. Let A be the linear combination of independent strategies $g_1, \ldots, g_m$; then the minimum of the value $E\eta(g^A) = E\eta(A(g_1, \ldots, g_m))$ is obtained if $\lambda_1 = \ldots = \lambda_m = \dfrac{1}{m}$.

Proof. Consider values $\varepsilon_l = \lambda_l - \dfrac{1}{m}$. Note that $\sum_{l=1}^{m} \varepsilon_l = 0$.

Using $E\left(\sum_{l=1}^{m} \varepsilon_l y_l \right)^2 \ge 0$, $E\left(\sum_{l=1}^{m} \varepsilon_l y_l \right) = 0$, $E\left(\sum_{i=1}^{m} y_i \sum_{l=1}^{m} \varepsilon_l y_l \right) = 0$, we get

$$E\left(y_0 - \sum_{l=1}^{m} \lambda_l y_l \right)^2 = E\left(y_0 - \frac{1}{m}\sum_{l=1}^{m} y_l - \sum_{l=1}^{m} \varepsilon_l y_l \right)^2 = E\left(y_0 - \frac{1}{m}\sum_{l=1}^{m} y_l \right)^2 -$$

$$-2E\left(\left(y_0-\frac{1}{m}\sum_{l=1}^{m}y_l\right)\left(\sum_{l=1}^{m}\varepsilon_l y_l\right)\right)+E\left(\sum_{l=1}^{m}\varepsilon_l y_l\right)^2 \geq E\left(y_0-\frac{1}{m}\sum_{l=1}^{m}y_l\right)^2.$$

■

A “Default” Algorithm of Constructing of a Consensus of Several Experts

Further on, we assume that strategy of nature c is unknown.

Let for some point $x \in D$ we have statements from several experts. Consider some "reasonable" algorithm of forming a consensus of experts' statements (denote it by A).

Firstly, the algorithm A coordinates each l-th expert's statements separately. Suppose that $S^1,\ldots,S^m \in \Omega^l$, $y^i(x)$ be the corresponding estimations of $y_0(x)$ made by l-th expert, $i=\overline{1,m}$.

Minimizing value $\sum_{i=1}^{m} w^i (y^i(x)-y)^2$, we get equation $\sum_{i=1}^{m} w^i (y^i(x)-y)=0$. Therefore, put

$$y_l(x)=\frac{\sum_{i=1}^{m} w^i y^i(x)}{\sum_{i=1}^{m} w^i};$$

here $y_l(x)$ is the coordinated opinion of l-th expert at the point $x \in D$.

Put $w_l = \max_i \left(1-\frac{2\Delta y^i}{\beta-\alpha}\right) w^i$, where $\Delta y^i = \left|y^i(x)-y_l(x)\right|$.

Secondly, the algorithm A coordinates all experts’ statements at the point $x \in D$. Suppose that we have statements from k experts, coordinated as above.

Minimizing value $\sum_{l=1}^{k} w_l (y_l(x)-y)^2$, we get equation $\sum_{l=1}^{k} w_l (y_l(x)-y)=0$. Therefore, put

$$y^A(x)=\frac{\sum_{l=1}^{k} w_l y_l(x)}{\sum_{l=1}^{k} w^l};$$

here $y^A(x)$ is the experts’ opinions at the point $x \in D$, coordinated by the algorithm A.

After coordination by the algorithm A for all $x \in D$ we have sets in the form of $\tilde{E}^1$ or $\tilde{E}^1 \setminus (\tilde{E}^2 \cup \tilde{E}^3 \cup \ldots)$ with different predictions, where $\tilde{E}^i$ are rectangular sets in D.

Let us notice that resulted decision rule may suffer from redundancy. Since there are M initial statements, we have up to 2^M sets in D with different predictions.

Consider algorithms B of forming a consensus of experts' statements under restrictions on amount of resulted statements. The value

$$F(B) = \frac{\int_D \left(y^A(x) - y^B(x)\right)^2 dx}{\mu(D)}$$

estimates a quality of the algorithm B. Here $y^A(x)$ and $y^B(x)$ are the estimations of the $y_0(x)$ prescribed to the point $x \in D$ by the algorithms A and B, respectively.

In the general case, the best algorithm $B^* = \arg\min_B F(B)$ is unknown. In the work [1], the heuristic algorithm of forming a consensus of experts' statements for the case of interval prediction is suggested. This algorithm uses distances / similarities between multidimensional sets in heterogeneous feature space [2, 3].

Conclusion

Suggested method of forming of united decision rule (as the method in [1]) can be used for coordination of several experts statements and different decision rules obtained from learning samples and/or time series. Applications of these methods are relevant to many areas, such as medicine, economics and management.

Acknowledgements

The work was supported by the RFBR under Grants N07-01-00331a, 08-07-00136a.

Bibliography

[1] G.Lbov, M.Gerasimov. Interval Prediction Based on Experts' Statements. In: Proc. of XIII Int. Conf. "Knowledge-Dialogue-Solution", 2007, Vol. 2, pp. 474-478.

[2] G.S.Lbov, M.K.Gerasimov. Determining of Distance Between Logical Statements in Forecasting Problems [in Russian]. In: Artificial Intelligence, Ukraine. 2004, Vol.2, pp. 105-108.

[3] A.Vikent'ev. Measure of Refutation and Metrics on Statements of Experts (Logical Formulas) in the Models for Some Theory. In: Int. Journal "Information Theories & Applications", 2007, Vol. 14, No.1, pp. 92-95.

Authors' Information

Gennadiy Lbov - Institute of Mathematics, SB RAS, Koptyug St., bl.4, Novosibirsk, Russia;
e-mail: lbov@math.nsc.ru

Maxim Gerasimov - Institute of Mathematics, SB RAS, Koptyug St., bl.4, Novosibirsk, Russia,
e-mail: max_post@ngs.ru