J. Geophys. 43, 215–225, 1977

Journal of Geophysics

Data Seizing and Information Processing

Application to Recognition of High Seismicity Earthquake Areas

C.F. Picard and J. Sallantin

Structures de l'Information, Groupe de Recherche du CNRS, Associé à l'Université Paris VI, Tour 45, 4, Place Jussieu, F-75230 Paris Cedex 05, France

Abstract. Different methods of information processing for discriminating types of events, and some results allowing to build a projective representation, are presented. The results are applied to the recognition of possible locations of strong earthquakes in Pamir and Tian Shan.

Key words: Information theory – Projective representation – Pattern recognition – Eearthquake areas – Questionnaire.

Introduction

When we study a phenomenon, we use data resulting from experiments, and these data contain all the information which can be obtained. We do not choose the same method for information processing if we want to classify events by data or to recognize one event. Obviously the choice of the method of information processing depends on the expected result. We present here three of these methods which are efficient when we want to recognize a class of events: each of them supposes a mathematical representation for the type of events and for the operations or questions separating the events.

In the first two methods, we can assume the possibility of a representation of events and tests by a probabilistic state, such that we may identify all the events with more or less accuracy by the subsets of this probabilistic state; in the third method we show how to build a representation of the events by a set of subspaces of a Hilbert space. The major part of this note is concerned with this so-called projective representation method; the others, questionnaires and pseudoquestionnaires, are briefly given here for a better understanding of the links between classification methods and information theory.

1. In a number of cases an experiment is made with the idea of repetition in mind. It is of course a wellknown aspect of computer science to work only with highly repetitive programs. Then, the experiments have a statistical aspect and it is often possible to build probabilistic spaces upon events known only by the frequencies of their realization. A questionnaire is a mathematical object which

can be used as a model of experiment or as a scheme of a computer program acting on a lot of data. Without introducing the exact definition of a questionnaire, we will give some of its properties (Picard, 1972, 1974).

Let us suppose we must realize an experiment with a finite number N of outcomes which are the events of a given space Ω, whose set of subsets is $\mathscr{A}$ (finite case). These events are called the answers; if we have at our disposal a mechanism such that we can identify the N events with only one elementary experiment or question, we will say that only one question was needed for the separation of these events. If the mechanism needs more than one question to do the separation, then we will speak of a questionnaire with several questions. The set-theoretical aspect of the questionnaire is given by the next remark: in many cases each questions is acting exactly as a partition operator, and Ω is cut into smaller and smaller subsets until every subset is atomic; the process is then called an "arborescent" one. But in other cases, a question does not realize a partition because two outcomes of an elementary question are subsets of Ω with a non empty intersection: there is a covering of preceeding subsets, and the process can be called a "latticoid" one. A graph obviously associated to a questionnaire has the property: from the first question called root to every other vertex (question or answer) there is only one path in the arborescent case, but not in the latticoid case. In both cases, the questionnaire allows to find the answer to the problem under study or what item would be selected. The questions and final answers are weighted by frequency coefficients – or, in the best case, by probabilities – so that it is possible to speak of weighted paths. The length of a path is the number of its arcs or of its questions: there is one incoming arc for every question, the root excepted. The routine length is defined as the mathematical expectation of length of all the paths, linking the root to the N answers. This routine length is fundamental to classify questionnaires built over the same set Ω.

A given questionnaire, with a unique set of questions, can be used for the processing of a lot of items coming from a collection of data. One of the main problems is to build a questionnaire which minimizes the routine length: it is a global optimization which cannot be done easily when some constraints forbid the use of certain partitions. In fact, every question is associated to an operator of interrogation. A simple one is the comparison between two distinct numbers a and b, the two outcomes being $a>b$ and $b>a$. But this operator leads to an arborescence with 4 answers and not only 3, when it is applied to a set of 3 numbers $\{a, b, c\}$ for finding their maximum.

Firstly, we put "$a>b$?" and at the second level "$a>c$?" after a positive answer, and "$c>b$?" after a negative answer. Then we get the 4 answers $\{a>(b, c), c>a>b, c>b>a, b>(a, c)\}$. It is obviously possible to keep only three answers (a or b or c is the maximum) with a loss of information so that the only feasible questionnaire with exactly 3 answers and built upon the operator "$\bullet>\bullet$?" is a latticoid, and the so-called processed information (with 4 events) is greater than the so-called transmitted information (with only 3 events).

The theory of information is a guide for the prediction of a lower bound of the routine length, and it gives a possibility to evaluate this length with the help of only the probabilities of answers; the restrictions imposed by the set of interrogation operators create a gap between the lower bound and the practical routine length. Other weights or parameters can be used to take into account some other economical or practical concepts: cost of every question, power of the questions, utility of the answers ...

We must point out the difference between these questionnaires and the questionnaires as they are known in socio-economics. The information obtained by such procedures allows one to do exclusively unions or intersections of sets, but does not allow identification of items of a given class in a probabilistic environment. The main applications of the questionnaires theory concern searching and sorting, pattern recognition and the organization of data files in computers by efficient procedures for classifying or interrogation problems.

2. Let us consider decision processes concerning a random variable T taking a finite number of values T_1, T_j, T_N, when having the observations X_i of a discrete random variable X. We denote by $(\Omega, \mathscr{A}, P)$ the probability space over which the random variable T is defined. Let us suppose that a set Q of finite random variables on $(\Omega, \mathscr{A}, P)$ is given. An element $q \in Q$ is associated with every internal vertex of an arborescence whose paths represent the interrogational procedure. In the case of a questionnaire, Q is the set of authorized operations to apply to every question, and there is a bijection between T and X. We cannot put a question on T itself but only on X, and the answer is reliable and sure: the answer X_i will mean that the real phenomenon occurs with the event $T_i \in T$. In the case where $p(T_i | X_i) < 1$, the process is unreliable and there is a doubt: it is possible that the real phenomenon corresponding to X_i is the event T_j with a probability $p(T_j | X_i) \leqq 1 - p(T_i | X_i)$. There is no bijection from X to T, and we cannot say it is such or such $t \in T$ which is actually concerned; such situations with indirect interrogation are not modeled by questionnaires with reliable answers but by so-called pseudoquestionnaires with probabilistic vectorial answers (Terrenoire, 1973).

3. Now we cannot assume the possibility to represent events and tests on a probabilistic space. It may be possible with the help of correlation on the data to find a vectorial space H, and to define the interesting events by subspaces of this space; this is the case if it is impossible to determine the features in the data that allow to classify the events.

These methods are called projective representations and have been studied by Watanabe (1969) and Watanabe and Pakvassa (1973). Some actual developments will be represented in this paper with pattern recognition of earthquakes areas as example.

Bongard and Gelfand's problem of pattern recognition is in the third category because, as we know, there are no discriminant features to distinguish dangerous areas from others (Gelfand et al., 1972, 1973). Nevertheless a strong earthquake as a geomorphological phenomenon has to mark the surface of the earth; and it is possible to suppose a difference between the geophysical descriptions of areas where earthquakes happen or not.

The pattern or class of events that we want to recognize is defined by an issue:

a: epicenters of strong earthquakes may be situated in the area,

b: opposite (may not).

If the data allow a good representation of the class of events, they do not allow always an inductive analysis of the phenomenon. For example, it is possible to describe only the secondary effects of the phenomenon in the code, and these are enough to obtain a good projective representation of some patterns.

Now we present some aspects of projective representation (Sallantin, 1976) which allow a practical approach to this problem. Afterwards we show how Bongard and Gelfand's method can be justified by these results. Then we present another method and compare the results.

Construction of a Projective Representation

From a mathematical point of view, we have three steps in the study of this representation:

- To find the conditions of existence.
- To obtain the practical formulation allowing projective information processing.
- To practise an algorithm.

1. For the first step, conditions are defined according to a logical algebraic formulation. It is possible to show that projective representations do not correspond to Boolean logic but to quantum logic as they are defined by J. von Neuman and G. Birkhoff to justify the basis of quantum mechanics.

This aspect is studied in von Neuman and G. Birkhoff (1973). We suppose here that it is possible to have a projective representation.

2. Let Ω be a set of objects about which we put a set of questions having a proposition as a semantic value: the assertion of each proposition his called an issue, and the non assertion, negation, out of the context, is not considered, $\mathscr{L}$ is the set of issues. Information on the objects is accessible by a code $x \in X$, $X \subset \mathbb{R}^n$.

The quest of the issue a transforms the code $x \in X$, we shall define that function:

$$\varphi_a\colon \mathbb{R}^n \to \mathbb{R}^n.$$

We have also to consider an application f:

$$f\colon \mathscr{L} \times \mathbb{R}^n \to [0, 1],$$

where $f(a, x)$ is the frequency of verification of the issue a for the class of objects coded by x.

The analysis describes an interrogational process

$$\{X, \mathscr{L}, f, \varphi_a | a \in \mathscr{L}\}.$$

To have a projective representation, it is necessary to suppose on $\mathscr{L}$ a monoid structure for an operation of concatenation $\square$, corresponding to the succession of 2 assertions, and to have for each code a morphism of $\mathscr{L}$ on $[0, 1]$.

This is verified for the next set operation

1) $f(a \square b, x) = f(a, x) \cdot f(b, \varphi_a(x))$,

and a pseudo complementation verifying:

2) $f(\neg a, x) = 1 - f(a, x)$.

These operations are proposed by S. Watanabe (1969). Now to obtain a practical projective representation we have to find

$$\{X, \mathscr{L}, f, \varphi_a | a \in \mathscr{L}\}$$

that verifies properties 1 and 2. We shall give a procedure to obtain it.

Let us consider $\mathbb{R}^n$, and $\mathscr{L}(\mathbb{R}^n)$ the set of vectorial subspaces of $\mathbb{R}^n$. The scalar product is noted $\langle | \rangle$; $\mathscr{L}(\mathbb{R}^n)$ is a lattice, ortho-complemented with

1) $\neg \cdot \neg a = a \qquad a \in \mathscr{L}(\mathbb{R}^n)$

2) $\neg a \wedge a = \{0\}$ with $\{0\} = \neg I, \qquad I = \mathbb{R}^n \in \mathscr{L}(\mathbb{R}^n)$

3) $a \wedge b = a \Rightarrow \neg b \wedge \neg a' = \neg b$

and orthomodular with

$$a \wedge b = b \Rightarrow b = a \wedge (b \vee \neg a).$$

On $\mathscr{L}(H)$ we have functions verifying

$$m(0) = 0, \qquad m(I) = 1$$

$$a \wedge b = a \Rightarrow m(a) \leqq m(b), \qquad a, b \in \mathscr{L}(H)$$

$$a \wedge b = 0 \Rightarrow m(a \vee b) = m(a) + m(b).$$

This functions are denoted generally state in literature; we can now use a Gleason result (Parthassarathy, 1970).

Theorem. *Let us have* $\mathbb{R}^n$, $n \geqq 3$, *and* $\mathscr{L}(\mathbb{R}^N)$ *the lattice of all the subspaces of* $\mathbb{R}^n$. *We have a one to one mapping on the positive self adjoint operator* $T\mu$ *with a unity trace to the states of* $\mathscr{L}(\mathbb{R}^n)$. *If* $(a_i, i = 1, n)$ *is a basis of* $a \in \mathscr{L}(\mathbb{R}^n)$, *then:*

$$\mu(a) = \sum_i \langle a_i | T\mu\, a_i \rangle.$$

With the help of this result, it is possible to obtain all the states of $\mathscr{L}(\mathbb{R}^n)$. We suppose now that $f(a, x) = \mu(\{x\})$, where $\{x\}$ is the subspace generated by x, for a state μ of $\mathscr{L}(\mathbb{R}^n)$, and we try to obtain subspaces that optimize the state μ. For this, we need also that they should minimize the entropy function.

Definition. Let $(l_1, \ldots, l_p)$ be a basis of the subspace l. Define

$$H^1(l) = \sum_1^p m(l_i) \ln \left(\frac{1}{m(l_i)} \right)$$

and if $\beta \neq 1$, $\beta \geqq 0$

$$H^{\beta}(l) = \left[\sum_{1}^{p} m(l_i)(1 - m(l_i)^{\beta-1})\right] \cdot [1 - 2^{1-\beta}]^{-1}.$$

These functions are known as β type informations:

Now it is possible to show

Proposition 1. *Let A be the self adjoint operator associated to the state m. Let $K = (K_1, \ldots, K_n)$ be a basis of eigenvectors of A; then any other orthonormal basis $l = (l_1, \ldots, l_n)$ of $\mathbb{R}^n$ verifies:*

$$H^{\beta}(l) \geqq H^{\beta}(K).$$

Now, we define the entropy of a subspace S of $\mathbb{R}^n$ by

$$H^{\beta}(S) = \min_{s \in \mathcal{O}} H^{\beta}(s)$$

where $\mathcal{O}$ is the set of orthonormal basis of S.

The next result can help us to choose the subspace that maximizes the state and minimizes the entropy. Let us call Karhunen Loève expansion the basis of eigenvectors of A ordered by decreasing eigenvalues (S. Watanabe, 1969).

Proposition 2. *For a state m, between all the subspaces S of $\mathscr{L}(\mathbb{R}^n)$ with* $\dim(S) = d$, *$d > 1$, there is one S_d, built with the d first vectors of the expansion of Karhunen Loève of the autoadjoint unity trace operator associated to m, which verifies*

$$H^{\beta}(S_d) \leqq H^{\beta}(S), \quad d \geqq \beta^{\frac{1}{\beta - 1}}.$$

For experimental results, we look for a state and choose a subspace corresponding to the solution of Proposition 2. It is possible to know this subspace; its basis is the eigenvectors basis of A.

3. To obtain the algorithm, we use the classical steps of pattern recognition

1) to describe the feature of the objects (code)

2) to learn the issue of $\mathscr{L}$ we want to recognize with a set of issues verifying them (learning)

3) to verify validity of learning on other objects (generalization), and to apply the program.

For each idempotent issue for $\square$, i.e., $a \square a = a$, we are looking for a representation. Let Ω_a be the learning set of a, $a \in \mathscr{L}$; Ω_a has σ elements; let $l_1, \ldots, l_n$ be the basis of the code, $\langle | \rangle$ be the scalar product. Every event ω_α has a code

$$c(\omega_\alpha) = (c_1(\omega_\alpha), \ldots, c_n(\omega_\alpha))$$

with a weight $p(\omega_\alpha)$ such that

$$\sum_{\alpha=1}^{\sigma} p(\omega_\alpha) = 1$$

let k be a symmetrical application of $\mathbb{R}^2$ to $\mathbb{R}$ verifying

$$k(x,x) \geqq 0.$$

A matrix G of self correlation will be:

$$G(l_i, l_j) = \left[\sum_{\sigma=1}^{\sigma} p(\omega_\alpha)\, k(c_i(\omega_\alpha), c_j(\omega_\alpha)) \right] \times \left[\sum_{\alpha=1}^{\sigma} \sum_{i=1}^{n} p(\omega_\alpha)\, k(c_i(\omega_\alpha), c_i(\omega_\alpha)) \right]^{-1}$$

with the associated state for $L \in \mathscr{L}(\mathbb{R}^n)$ with a basis $l_1, \ldots, l_p$

$$\rho(L) = \sum_{i=1}^{p} \sum_{j,j'=1}^{n} \langle l_i | l'_j \rangle\, G(l'_j, l'_j) \langle l'_j | l_i \rangle.$$

According to the preceding result, if we diagonalize $G(l_i, l_j)$, we find the subspace that maximizes the state corresponding to the learning set and minimizes the entropy.

We define:

$$f(a,x) = \frac{\|\varphi_a(x)\|^2}{\|(x)\|^2}$$

where

$$\|l\| = \langle l | l \rangle^{1/2}$$

and φ_a is the projector corresponding to the subspace chosen to represent a. For this choice we verify:

1) $f(a \Box a, x) = f(a,x)$

2) $f(\neg a, x) = 1 - f(a,x)$.

The projective information processing is here defined by

$$\left\{ \Omega, \mathbb{R}^n, \mathscr{L}, \frac{\|\varphi_a(x)\|^2}{\|x\|^2}, \varphi_a | a \in \mathscr{L} \right\}$$

and φ_a is the projector associated to a subspace chosen with Proposition 2 to represent a.

3.1. Now we can apply this to the study of earthquakes areas. We have a $\mathbb{R}^{32}$ code of events in Pamir, and for an event $c_i(\omega) = 0, 1, 2$ according to the feature i being false, true, or unknown. Between the functions k that it is possible to choose, we put

$$k(c_i(\omega_\alpha), c_j(\omega_\alpha)) = a_{c_i(\omega_\alpha) c_j(\omega_\alpha)}$$

where

$$(a_{ij})=\begin{pmatrix} 1 & 0 & 0.2 \\ 0 & 1 & 0.2 \\ 0.2 & 0.2 & 0.2 \end{pmatrix} \qquad \begin{matrix} i=0,1,2 \\ j=0,1,2 \end{matrix}$$

is generated by two issues:

a: epicenter of strong earthquake may be situated in this area
a': opposite (may not).

These issues are considered as simple: $a\Box a=a$, $a'\Box a'=a'$.

For the calculus we use the second part of Proposition 2 and choose a d-dimensional representation. To help this choice we use an informational dimension, for an issue a and its subspace A:

$$D(a)=e^{H^1(A)}.$$

This function measures how A represents the events, used to learn A; if the geometrical dimension of A is d:

$$1\leqq D(a)\leqq d.$$

The decision function that we use is the least sophisticated possible.

φ_a and $\varphi_{a'}$ can be chosen so that the geometrical dimension of the subspace corresponding to $\varphi_a+\varphi_{a'}$ is

$$\dim(\varphi_a+\varphi_{a'})=\dim\varphi_a+\dim\varphi_{a'}$$

and we have also:

$$\varphi_{a'}\circ\varphi_{\neg a}=\varphi_{a'}, \qquad a'\Box\neg a=\neg a\Box a'=a'.$$

$f(a,x)\geqq f(a',x)+\varepsilon$	decision $=+1$	the area is dangerous
$\varepsilon+f(a,x)\leqq f(a',x)$	decision $=-1$	the area is non dangerous
$\lvert f(a,x)-f(a',x)\rvert<\varepsilon$	decision $=\;\;0$	no decision

3.2. We can describe the Bongard-Gelfand algorithm (Gelfand et al., 1972; 1973) in the same formal way. From the first same code X in $\mathbb{R}^{32}$ they extract another one X' in $\mathbb{R}^{21}$. We have not to justify this change of code; on this new code, the next function k has been chosen:

$$k(c_i(\omega_\alpha), c_j(\omega_\alpha))=\delta_{ij}$$

where

$$\delta_{ij}=0 \quad \text{if } i\neq j, \qquad \delta_{ii}=1.$$

$G(l_i, l_j)$ is diagonalized for a and a', and we take the ten largest eigenvalues and these eigenvectors to represent a, and the 11 eigenvectors, with the lowest eigenvalues for a but largest for a', to represent a'.

$\left\{\Omega, \mathbb{R}^{21}, \mathscr{L}, \dfrac{\|\varphi_a(x)\|^2}{\|x\|^2}, \varphi_a \mid a\in\mathscr{L}\right\}$ are defined, and there we remark that a' is $\neg a$.

The decision function is here: an area is dangerous if

$$\|\varphi_a(x)\| - \|\varphi_{\neg a}(x)\| \geqq \inf \|\varphi_a(x_i) - \varphi_{\neg a}(x_i)\|, \qquad x_i \in c(\Omega_a).$$

Now we present the results, with the help of 2 data. On this example we can see how it is possible to have projective representations: but we do not know how to find the best one.

The Table 1 is a comparison of results when the learning is on an increasing

Table 1

		5 1896	8 1907	9 1911	10 1911	11 1934	12 1938	13 1941	14 1948	15 1955	16 1970
1	6	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1
3	3	1	1	1	1	1	1	1	1	1	1
4	35	1	1	1	1	1	1	1	1	1	1
5	13	→ 1	1	1	1	1	1	1	1	1	1
6	38	−1 • *	1	1	1	1	1	1	1	1	1
7	20	−1 • *	1	1	1	1	1	1	1	1	1
8	27	1	→ 1	1	1	1	1	1	1	1	1
9	4	1 •	1 •	→ 1	1	1	1	1	1	1	1
10	36	1 •	−1 • *	1	→ 1	1	1	1	1	1	1
11	25	−1 • *	1 •	1	1	→ 1	1	1	1	1	1
12	5	1	1	1	1	1	→ 1	1	1	1	1
13	24	−1 *	−1 *	1	1	1	1	→ 1	1	1	1
14	15	1	1 •	1	1	1	1	1	→ 1	1	1
15	22	1	1	1	1	1	1	1	1	→ 1	1
16	7	↔ 1	1	1	1	1	1 •	1	1	1	→ 1
17	37	1	1	1	1	1	1	1	1	1	1
18	23	−1	1	1	1	1	0	1	1	1	1
19	31	1	1	1	1	1	1	1	1	1	1
20	26	1	1	1	−1	−1	−1	−1	−1	−1	−1
21	41	1	−1	1	1	1	1	1	0	0	0
22	34	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
23	8	↔ −1	−1	−1 •	−1	−1	−1	−1	−1	−1	−1 •
24	32	−1	1 *	1 *	1 • *	1 *	1 *	1 *	1 *	1 *	1 *
25	10	−1 •	−1 •	−1 •	−1	−1	−1	−1	−1 •	−1	−1
26	14	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
27	39	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
28	40	−1	−1	−1 •	−1	−1	−1	−1	−1	−1	−1
29	16	−1	1 • *	1 • *	1 *	−1	1 *	1 • *	1 • *	1 *	1 *
30	17	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
31	19	−1	−1 •	−1	−1	−1	−1	−1	−1	−1	−1
32	21	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
33	33	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
34	9	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
35	11	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
36	12	−1	−1 •	−1	−1	−1	−1	−1	−1	−1	−1
37	29	−1	−1 •	−1	−1	−1	−1	−1	−1	−1	−1
38	30	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
39	2	−1	−1 •	−1	−1	−1	−1	−1	−1	−1	−1
40	18	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
41	28	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1

Table 2

1	6	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	3	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	35	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	13	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	38	−1	−1	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	20	−1	−1	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	27	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	4	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	36	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	25	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	5	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	24	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	15	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	22	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	7	−1	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	37	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	23	−1	−1	−1	1	1	1	1	1	−1	−1	−1	0	−1	−1	−1	−1	−1
19	31	−1	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	26	−1	1	1	−1	−1	1	1	1	−1	−1	−1	−1	−1	−1	−1	1	1
21	41	−1	−1	−1	0	0	0	1	1	0	−1	0	1	1	1	1	1	1
22	34	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	1	1	1	1	1	−1	0
23	8	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
24	32	−1	−1	−1	−1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
25	10	−1	1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
26	14	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
27	39	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
28	40	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
29	16	−1	−1	−1	−1	1	1	0	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
30	17	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
31	19	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
32	21	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
33	33	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
34	9	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
35	11	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
36	12	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
37	29	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
38	30	−1	1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
39	2	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
40	18	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
41	28	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1

number of events. The 16 first lines correspond to decision functions about effective events. These events are in a chronological order and decision function must be +1. From 17 to 22 we have the analysis of suspect areas. From 23 the lines correspond to the decision for areas considered as non dangerous, the decision must be −1, unless a wrong appreciation of the area. To learn the dangerous area we take the events that happen before the date, 5, 8 etc. The second column corresponds to the Gelfand et al. number to determine the area.

- • Decision is wrong in the analysis of Gelfand et al. (1973).
- * Decision is wrong in our analysis

Dimension of A and A' is 5.

In the Table 2 we have 16 events to learn A and each line corresponds to the decision function when dimension of A and A' vary from 1–17. Evidently the first 16 lines must have decision $+1$ and from the line 23 the decision must be -1. The dimension choice is empirically made because no relations have to be taken into account between issues.

Conclusion

We gave 3 methods useful for classification and pattern recognition. With the projective analysis on geomorphological data, we can give representations of the issues formulating the risk of strong earthquakes on an area.

In this example we had not enough issues to obtain relations on the set of the representations of issues. These relations allow a very interesting study of the part that each feature of the geomorphological data plays to represent the issues.

Acknowledgements. We thank Mr. Jobert and Mr. Keilis Borok for their encouragements and interesting discussions about this work.

References

Gelfand, I.M., Guberman, Sh.I., Izverova, M.L., Keilis Borok, V.I., Ranzman, E.J.: Recognition of possible location of strong earthquakes in Pamir, Tian Shan. Computational seismology, Vol. VI, pp. 107–133. Moskow: Nauka 1973 (in Russian)

Gelfand, I.M., Guberman, Sh.I., Keilis Borok, V.I., Ranzman, E.J.: Criteria of high seismicity, determined by pattern recognition. Tectonophysics **13**, 415–422, 1972

Neuman, J.v., Birkhoff, G.: The logic of quantum theory. Ann. Math. **37**, 823, 1936

Parthassarathy, K.R.: Probability theory on the closed subspaces of a Hilbert space. In: Les probabilités sur les structures algébriques, Colloques Internationaux du C.N.R.S., n° 186, Paris, 1970

Picard, C.F.: Graphes et questionnaires. Paris: Gauthier-Villars 1972

Picard, C.F., Petolla, S.: Théories de l'information. Lecture Notes in Math. 398, pp. 126–151. Berlin-Heidelberg-New York: Springer 1974

Sallantin, J.: Sur la représentation de processus d'interrogation. Comptes-Rendus à l'Académie des Sciences, Paris, Tome 283, Série A, n° 4, pp. 219–222, 1976

Terrenoire, M.: Convergence of heuristics for some interrogation processes. First International Joint Conference on Pattern Recognition, Washington, 1973

Watanabe, S.: Knowing and guessing. New York: Wiley 1969

Watanabe, S., Pakvassa, N.: Subspace method in pattern recognition. First International Joint Conference on Pattern Recognition, Washington, 1973

Received October 15, 1976; Revised Version February 21, 1977