DOI: 10.2478/auom-2024-0027
An. Şt. Univ. Ovidius Constanţa **Vol. 32(3),2024, 33–55**

Using the Choquet integral for the determination of the anxiety degree

Ion Chiţescu, Mădălina Giurgescu (Manea) and Titi Paraschiv

Abstract

One considers a group S of subjects and one uses the measured values of EEG waves and the levels of anxiety given by the psychologists for the members of S in order to construct a monotone measure μ. This measure μ is such that the levels of anxiety of the members of S are Choquet integrals with respect to μ. One infers that this is true for any other subject. This inference is validated by comparing the values obtained via the above mentioned method with the results given by the psychologists for another group of subjects S'. We think that this procedure is completely new for the computation of the anxiety degree.

1 Introduction

The present paper is in the domain of Mathematical Psychology, being concerned with the study of anxiety. Namely, we try to compute the level of anxiety of different subjects using a new pattern. To be more precise, we use the Choquet integral as aggregation tool, as follows. First we consider a fixed set S of subjects (here S has 70 members) and we collect the measured values for the EEG waves and the values of anxiety levels (given by psychologists) of all members in S. Using the aforementioned values as coefficients, we create a monotone measure μ having the property that for any member t of S its level of anxiety is given by the Choquet integral $(C)\int f d\mu$, where f is a function depending on t.

Key Words: Monotone measure, Nonlinear integral, Choquet integral, Instrument of aggregation, Inverse problem of information fusion, Anxiety, BigFive, EEG, NeuroSky.
2010 Mathematics Subject Classification: 26E50, 28A25, 28E10, 92C50.
Received: 10.05.2023
Accepted: 15.11.2023

The idea is to infer that for any subject the level of anxiety is obtained as above. This inference is checked using another set S' (here S' has 10 members) comparing the real levels of anxiety of members in S' (given by psychologists) with the levels of anxiety given by the above formula (the result should be approximately the same).

To summarize:

1. We use EEG waves and real levels of anxiety in order to construct a monotone measure μ.

2. The levels of anxiety should be Choquet integrals with respect to μ.

3. Last but not least, we work with Fuzzy Analysis.

Our opinion is that 1, 2 and 3 are new ideas in Mathematical Psychology.

We lay stress upon the following facts which are new. First, the idea to work with Fuzzy Analysis is new: Classical Analysis and Statistics were used up to now. Secondly, the idea to use psychological data to construct a measure which will be used afterwards for computing a Choquet integral is new.

We think that some brief historical considerations concerning Mathematical Psychology can be useful. Let us begin with R. Hooke who worked on "modelling human memory". The real history of Mathematical Psychology began in the 19th century. During this century, the main schools in the domain were the German school and the English school. The German school was mainly concerned with experimental psychology and psychophysics. Prominent researchers of this school were E. Weber (first mathematical law of the mind - Weber's law), G. Fechner (theories of sensations and perceptions - modified law of Weber). Combined, we have the famous Weber-Fechner law. H. von Helmholtz (nerve conduction speed theory of hearing, color vision) and J. Herbart (mathematical theories of cognitive area). The English school was mainly concerned with statistical considerations. In the 19th century, the main representative of the English school of Mathematical Psychology was F. Galton, who is credited as the founder of psychometrics. He also initiated the study of anthropometric development.

In the 20th century, the center of researches in Mathematical Psychology moved to USA. Ch. Spearman invented the factor analysis. Bush, Mosteller and Estes had remarkable achievements in learning theory. The American school of Mathematical Psychology uses now very much Information Theory, Computation Theory, Mathematical Linguistics, with the permanent support of Statistics and study of measurement error. The Journal of Mathematical Psychology was founded in 1964, being the most influential journal in the domain.

Importat, fundamental results in Mathematical Psychology emerged from this history. First, the following deterministic laws: Weber-Fechner law, Ekman's law, Stevens's power law. Next, the following stochastic laws: Thurstone's law, the matching law, Rescorla-Wagner rule. It is worth mentioning that, in mathematical modelling of psychological processes, we have the following trend: from deterministic relations (as classical physics) to stochastic relations (see [11]).

There is another trend, what concerns the mathematics used by various branches of science and in particular, in Mathematical Psychology. One passes from classical measure and integral theory to generalized measure and integral theory (called by many people fuzzy theory). These new theories become more and more used. Fuzzy integrals became very powerful aggregation tools in Data Mining. This was our main motivation to use the Choquet integral in the present paper.

Here are some examples of recent papers of mathematical psychology, illustrating various tendencies and using fuzzy mathematics.

In [3] measurements of psychological data from the Likert scale and genetic processes are studied.

In [6] a purely ordinal model for aggregation functionals for lattice-valued functions, using the Ky Fan metric and the Sugeno integral is introduced. Reflection lattices are used for modelling psychological experiments.

In [13] a neuron model using a fuzzy integral in a multiclass decision making is introduced. The Sugeno integral with respect to a Sugeno measure is used.

In [19] the compatibility between psychology and linguistics terms is used. The authors work with a generalized Choquet integral.

Let us pass to the content of our paper. After the present **Introduction**, section 2 entitled **Preliminary facts** comes, being divided into two subsections: subsection 2.1 entitled **Mathematical preliminary facts** and subsection 2.2 entitled **Psychological preliminary facts**. Notions, notations and results used in the remainder of the paper are introduced. Section 3 is entitled **Using the Choquet integral to solve the inverse problem of information fusion** and is divided into three subsections. Subsection 3.1 entitled **Linearization formula for the computation of the finite discrete Choquet integral** (a formula which is similar to the linear formula of the integral of simple functions). Subsection 3.2 entitled **Canonical enumeration of** $P^*(T)$ (a lexicographical ordering of the set $P(\{1, 2, \cdots, n\})$. Subsection 3.3 entitled **Solving the inverse problem of information fusion in this case (Identification of the monotone measure used to generate the aggregation tool)**. Using the input data as coefficients, one constructs a monotone measure (like in subsection 3.1) as the (possibly approximate) solution of a linear system. For section 3 one can consult [20] and [21]. Section 4 is entitled **Determination of the anxiety degree** and is divided into two subsections. Subsection 4.1 is entitled **Creating the instrument (Measure). The inverse problem**. Using the input data given by psychologists and measurements for a fixed group of subjects, a monotone measure is constructed, using the theory from section 3. Subsection 4.2 is entitled **Using the obtained instrument (Measure) to determine the anxiety level of other subjects. The direct problem**. One computes the anxiety level of 10 new subjects using the integral formula of our method. The results are compared with the real anxiety levels of the 10 new subjects (given by psychologists) and one notices that the two results are very close. One considers that this validates the formula. Section 5 is entitled **Conclusions**.

2 Preliminary facts

2.1 Mathematical preliminary facts

Throughout the paper, the positive integer numbers will be $\mathbb{N} = \{0, 1, 2, \cdots\}$, the real numbers will be $\mathbb{R}$ and the positive real numbers will be $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$. As usual, $\bar{\mathbb{R}}_+ = \mathbb{R}_+ \bigcup \{\infty\}$.

For any set T, the Boolean of T is $P(T) = \{A | A \subset T\}$. A measurable space is a couple (T, τ), where T is a non-empty set and $\tau \subset P(T)$ is a σ-algebra.

If (T, τ) is a measurable space, a monotone measure (or a fuzzy measure) is a function $\mu : \tau \to \bar{\mathbb{R}}_+$ having the properties:

i) $\mu(\emptyset) = 0$

ii) $\mu(A) \leq \mu(B)$ for any A, B in τ such that $A \subset B$

Now, let us consider a measurable space (T, τ), a monotone measure $\mu : \tau \to \bar{\mathbb{R}}_+$ and a positive τ-measurable function $f : T \to \mathbb{R}_+$.

For any $a \in \mathbb{R}_+$, we consider the inferior level set $F_a = \{t \in T | f(t) \geq a\} = f^{-1}([a, \infty)) \in \tau$. Because $F_b \subset F_a$, whenever $0 \leq a < b < \infty$, it is seen that the function $\phi : \mathbb{R}_+ \to \bar{\mathbb{R}}_+$, given via $\phi(a) \stackrel{\text{def}}{=} \mu(F_a)$, is decreasing. Considering the Lebesgue measurable sets of $\mathbb{R}_+$ which we denote by $\mathcal{L}$ and the Lebesgue measure on $\mathcal{L}$ which we denote by $L : \mathcal{L} \to \bar{\mathbb{R}}_+$, we can compute $\int \phi dL$, because ϕ is $\mathcal{L}$-measurable. For the sake of concreteness and respecting the traditional notations, we shall write $\int_0^\infty \mu(F_a)da$ instead of $\int \phi dL$.

Definition 1. *The Choquet integral of the function f with respect to the measure μ is the element $\int_0^\infty \mu(F_a)da \in \bar{\mathbb{R}}_+$. We shall write*

$$(C) \int f d\mu = \int_0^\infty \mu(F_a)da$$

We shall say that f is Choquet integrable with respect to μ in case $(C) \int f d\mu < \infty$. (see [5], [8], [16], [20] and [21])

The Choquet integral is a generalization of the abstract Lebesgue integral. Namely, in case μ is a classic measure (i.e. μ is σ-additive), we can see that the Choquet integral $(C)\int f d\mu$ coincides with the abstract Lebesgue integral $\int f d\mu$.

We have special formulae for the computation of the Choquet integral in case the function f is simple, i.e. f has the form

$$f = \sum_{i=1}^{n} a_i \phi_{A_i}$$

where $a_i \in \mathbb{R}_+$, $A_i \in \tau$ are mutually disjoint and $\cup_{i=1}^{n} A_i = T$. Here ϕ_A is the characteristic (indicator) function of the set $A \subset T$.

Namely, for such f, one can put in order the numbers a_i such that $a_1 \leq a_2 \leq \cdots \leq a_n$. Considering that f is in this situation, one has the formula

$$(C)\int f d\mu = \sum_{i=1}^{n} (a_i - a_{i-1})\mu(A_i \cup A_{i+1} \cup \cdots \cup A_n) \tag{1}$$

with the convention $a_0 \stackrel{\text{def}}{=} 0$.

Caution: The reordering of the values a_i is unique in case a_i are distinct and all A_i are nonempty. Otherwise, different reorderings can occur, but the value of $(C)\int f d\mu$ given by the formula from above does not depend upon these different reorderings.

For Generalized Measure and Integration Theory, see [5], [8], [16], [20] and [21]. Some special computing devices for the Choquet integral appear in [4].

To calculate the values, we created a C++ program. The source code is written in C++ in the CodeBlocks development medium, 17.12 version on Windows 10 operating system, combined with GNU GCC Compiler in MinGW distribution, 6.3 version. For the matrix operations the Eigen library, version 3.3 was used.

2.2 Psychological preliminary facts

The electroencephalogram (EEG) is a recording of the electrical activity of the brain. The recorded waveforms reflect the cortical electrical activity (see [15]).

The main frequencies of the human EEG waves are: Delta (frequency of oscillation between 0.5 and 4 Hz), Theta (frequency of oscillation between 4 and 7 Hz), Alpha (frequency of oscillation between 8 and 13 Hz, it's divided into Low Alpha (8–10 Hz) and High Alpha (10-13 Hz)), Beta (frequency of oscillation between 13 and 30 Hz, it's divided into Low Beta (12.5–16 Hz), Beta

(16.5-20 Hz) and High Beta (20.5-28Hz)) and Gamma (frequency of oscillation between 30 and 140 Hz).(see [15])

In psychological trait theory, the Big Five personality traits, also known as the five-factor model (FFM), is a suggested taxonomy, or grouping, for personality traits, developed from the 1980s onwards.The Big Five theory is based therefore on semantic associations between words and not on neuropsychological experiments. This theory uses descriptors of common language and suggests five broad dimensions commonly used to describe the human personality and psyche (see [7]). The theory identifies five factors (see [7]):

-openness to experience (inventive or curious vs. consistent or cautious)

-conscientiousness (efficient or organized vs. extravagant or careless)

-extraversion (outgoing/energetic vs. solitary/reserved)

-agreeableness (friendly or compassionate vs. challenging or callous)

-neuroticism (sensitive or nervous vs. resilient or confident)

Anxiety is an emotional state characterized by a feeling of insecurity, disorder, diffuse (see [1]). It is a state of fear, agitation, insecurity and nervousness. Symptoms of anxiety include high blood pressure, helplessness, maladaptation, sadness and anxiety, manifested in the body by palpitations (fast heartbeat), trembling, sweating in the palms and insomnia (see [1]).

It is estimated that approximately about 6.5% of the world's population suffers or has suffered from medically diagnosed anxiety, but there are many more who suffer from the annoying symptoms of stress or anxiety (see [1]).

Psychologists' studies show that there is an interactive relationship between anxiety and two factors in the Big Five theory, namely, extraversion and neuroticism. More exactly, anxiety has a positive relationship with neuroticism and a negative relationship with extraversion (see [7]).

The specialists in psychology state that the anxiety is characterized by LowAlpha, HighAlpha, LowBeta and HighBeta waves (see [15]).

In this regard, we have developed a mathematical model through which EEG waves are processed in order to determine the level of anxiety.

Before exhibiting our mathematical model described in the present paper, we shall make some bibliographical comments, pertaining to some aspects of the EEG theory. A very good and documented material, dedicated to the history of the encephalography, beginning with Hans Berger, can be found in [18].

The basic material concerning EEG considered by us, when writing this paper, is contained in the monographs [1], [15] and in [12]. In [2], the role of Alpha oscillations in cognitive psychomotor, psycho-emotional and physiological aspects of human life is discussed.In [7] one lays stress upon the interpretation of the fact that brain oscillation and empirical evidence link Delta oscillation with reward motivation and Alpha oscillation with anxiety. In [9] an interesting use of the Fourier transform in the study of the encephalograms is proposed. In [10], the Delta band (1-3.5 Hz)of the EEG oscillatory activity is studied, linked to a broad variety of perceptual and cognitive operations. In [14] one lays- stress upon the fact that, up to now, reactive, as opposite to proactive behaviour, during social interaction, have not been investigated in relation to other kinds of social behaviour. A virtual interaction model is proposed. In [17] one explains that the association between neuroticism and anxiety may be additionally explained by transdiagnostic factors.

3 Using the Choquet integral to solve the inverse problem of information fusion

3.1 Linearization formula for the computation of the finite discrete Choquet integral

We shall be concerned with the case when T is finite, $T = \{x_1, x_2, \cdots, x_n\}, n \geq 1$. Hence, for the function $f : T \to \mathbb{R}_+$, there exists a permutation $\sigma : T \to T$ (we write $\sigma(x_i) = x_i^*$ for any $i = 1, 2, \cdots, n$) such that $f(x_1^*) \leq f(x_2^*) \leq \cdots \leq f(x_n^*)$.

In case the values $f(x_1), f(x_2), \cdots, f(x_n)$ are distinct, such σ is unique.

Working for the measurable space $(T, \tau) = (T, P(T))$ and for a monotone measure $\mu : P(T) \to \mathbb{R}_+$, the formula (1) from Preliminary Facts gives

$$(C)\int f d\mu = \sum_{i=1}^{n} (f(x_i^*) - f(x_{i-1}^*))\mu(\{x_i^*, x_{i+1}^*, \cdots, x_n^*\}) \qquad (2)$$

with the convention $f(x_0^*) = 0$.

Let us write $P^*(T) = \{A \subset T | A \neq \emptyset\} = P(T)\backslash\{\emptyset\}$ (hence $P^*(T)$ has $2^n - 1$ sets). For any $E \in P^*(T)$, define

$$a_E \stackrel{\text{def}}{=} \min_{x_p \in E} f(x_p) - \max_{x_q \in T\backslash E} f(x_q)$$

with the convention $\max_{x_q \in \emptyset} f(x_q) = 0$ (in the case $E = T$).

Next, define

$$b_E = \begin{cases} a_E, & if \quad a_E \geq 0, \\ 0, & if \quad a_E < 0 \end{cases}$$

It is seen that $b_E = a_E^+$.

The next theorem gives a linear formula, with respect to μ, for the computation of $(C)\int f d\mu$ (no need to order the values of f). This formula is crucial for solving the inverse problem of information fusion.

Theorem 1. *One has the formula*

$$(C)\int f d\mu = \sum_{E\in P^*(T)} b_E \mu(E) \tag{3}$$

3.2 Canonical enumeration of $P^*(T)$

In order to have complete and concrete recipes for the computation of $(C)\int f d\mu$, we shall introduce the canonical enumeration of $P^*(T) = \{E_1, E_2, \cdots, E_{2^n-1}\}$, the order of the enumeration $E_1, E_2, \cdots, E_{2^n-1}$ being explained in the sequel.

Any number $j = 1, 2, \cdots, 2^n - 1$ can be uniquely written in the binary form

$$j = \overline{j_n j_{n-1} \ldots j_1} = j_1 + 2j_2 + 2^2 j_3 + \cdots + 2^{n-1} j_n$$

where all $j_i \in \{0, 1\}$ and at least one $j_i \neq 0$.

The natural order of the numbers j coincides with the lexicographic order of the representative complexes $(j_n, j_{n-1}, \cdots, j_1)$.

For instance, if $n = 3$, hence $2^n - 1 = 7$, each $j = 1, 2, \cdots, 7$ has the form $j = \overline{j_3 j_2 j_1} = j_1 + 2j_2 + 2^2 j_3$: $j = 1 = \overline{001}$ (with $j_1 = 1; j_2 = 0; j_3 = 0$); $j = 2 = \overline{010} < j = 3 = \overline{011} < j = 4 = \overline{100} < j = 5 = \overline{101} < j = 6 = \overline{110} < j = 7 = \overline{111}$.

We shall enumerate $P^*(T)$ in lexicographic order: $E_1, E_2, \cdots, E_7$, viewed as $E_1 = E_{\overline{001}}, E_2 = E_{\overline{010}}, \cdots, E_7 = E_{\overline{111}} = E_{2^n-1}$. Practically, via this concrete exemplification, we defined the canonical enumeration of $P^*(T)$.

Consequently, the membership rule is the following:

$$x_i \in E_{\overline{j_n j_{n-1} \cdots j_1}} \Leftrightarrow j_i = 1$$

(i.e. $E_{\overline{j_n j_{n-1} \cdots j_1}} = \cup_{j_i=1}\{x_i\}$)

This membership rule generates, exactly $2^n - 1$ different sets.

For $n = 3$, one has: $E_1 = \{x_1\}, E_2 = \{x_2\}, E_3 = \{x_1, x_2\}, E_4 = \{x_3\}, E_5 = \{x_1, x_3\}, E_6 = \{x_2, x_3\}, E_7 = \{x_1, x_2, x_3\}$.

One can see that formula (3) in Theorem 1 can be written in the form

$$(C)\int f d\mu = \sum_{j=1}^{2^n-1} b_{E_j} \mu(E_j) \tag{4}$$

where any $j = 1, 2, \cdots, 2^n - 1$ is written in the form $\overline{j_n j_{n-1} \cdots j_1}$.

The results in this section are standard.

3.3 Solving the inverse problem of information fusion in this case (Identification of the monotone measure used to generate the aggregation tool)

We shall consider the elements x_i of the set $T = \{x_1, x_2, \cdots, x_n\}$ as the source of our information, e.g. any such x_i is a subject of our observation. Any observation of all subjects in T will be considered as a function $f : T \to \mathbb{R}$. We shall make l observations $f_1, f_2, \cdots f_l$, obtaining for any such observation $f_i : T \to \mathbb{R}$ the values $f_i(x_j), j = 1, 2, \cdots, n$ and the conclusion (which is numerical) $y_i \in \mathbb{R}$. Namely, the values $f_i(x_j)$ and the values y_i are the input of the system (each y_i is a value of the fusion target). The idea is to use as aggregation tool the Choquet integral with respect to the some monotone measure μ. Acting in this manner, we obtain the data set as follows:

$$\begin{array}{ccccc} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_n) & y_1 \\ f_2(x_1) & f_2(x_2) & \cdots & f_2(x_n) & y_2 \\ & & \cdots & & \\ f_l(x_1) & f_l(x_2) & \cdots & f_l(x_n) & y_l \end{array} \tag{5}$$

where we accept the existence of a monotone measure $\mu : P(T) \to \mathbb{R}_+$ such that

$$y_p = (C) \int f_p d\mu, p = 1, 2, \cdots, l \tag{6}$$

Acting in the spirit of the inverse problem of information fusion, we shall consider that the unknown object is the measure μ.

To be more precise, we shall consider that the observed values $f_i(x_j)$ and the values y_i are known, but the measure μ is unknown (i.e. the values $\mu(E_j), j = 1, 2, \cdots, 2^n - 1$ are unknown), in this case the measure being the output.

Taking into account formulae (4), (5), (6) it follows that the values $\mu(E_1)$, $\mu(E_2), \cdots, \mu(E_{2^n-1})$ must be solutions of the linear system (l equations with $2^n - 1$ unknowns).

$$\sum_{j=1}^{2^n-1} b_{p_j} \mu(E_j) = y_p, p = 1, 2, \cdots, l \tag{7}$$

Here, according to Theorem 1, we have, for any $p = 1, 2, \cdots, l$ and any $j = 1, 2, \cdots, 2^n - 1$, $b_{p_j} = b^p_{E_j}$, i.e. $b_{p_j} = a^p_{E_j}$ if $a^p_{E_j} \geq 0$ and $b_{p_j} = 0$ if $a^p_{E_j} < 0$, where $a^p_{E_j} = \min_{x_t \in E_j} f_p(x_t) - \max_{x_t \notin E_j} f_p(x_t)$.

So, to solve our problem means to solve the system (7) and, generally speaking, this is a difficult task from computational point of view. It is preferable to solve (7) approximately, using the least squares method.

To this end, we consider the function $F : \mathbb{R}^{2^n-1} \to \mathbb{R}$, given via

$$F(t_1, t_2, \cdots, t_{2^n-1}) = \sum_{p=1}^{l} (\sum_{j=1}^{2^n-1} b_{p_j} t_j - y_p)^2$$

This function is infinitely differentiable and convex, being a sum of convex functions of the form $(t_1, t_2, \cdots, t_{2^n-1}) \to (a_1 t_1 + a_2 t_2 + \cdots + a_{2^n-1} t_{2^n-1} + b)^2$.

Hence, F has a minimum, attained at the point $(t_1^0, t_2^0, \ldots, t_{2^n-1}^0)$ which is the solution of the system of equations

$$\frac{\partial F}{\partial t_k}(t_1^0, t_2^0, \ldots, t_{2^n-1}^0) = 0, k = 1, 2, \cdots, 2^n - 1$$

($2^n - 1$ equations with $2^n - 1$ unknowns t_j^0).

This system is in fact

$$\sum_{p=1}^{l} b_{p_k} (\sum_{j=1}^{2^n-1} b_{p_j} t_j^0 - y_p) = 0 \Leftrightarrow$$

$$\sum_{j=1}^{2^n-1} (\sum_{p=1}^{l} b_{p_k} b_{p_j}) t_j^0 = \sum_{p=1}^{l} b_{p_k} y_p, k = 1, 2, \cdots, 2^n - 1 \tag{8}$$

It is seen that, in case $(t_1^0, t_2^0, \cdots, t_{2^n-1}^0) = (\mu(E_1), \mu(E_2), \cdots, \mu(E_{2^n-1}))$ is an exact solution of (7), then $(t_1^0, t_2^0, \cdots, t_{2^n-1}^0)$ is a solution of last system (8) too.

Otherwise, a solution $(t_1^0, t_2^0, \ldots, t_{2^n-1}^0)$ of (8) is an approximate solution of (7). We shall consider the measure $\nu : P(T) \to \mathbb{R}_+$ given via $\nu(E_j) = t_j^0$, $j = 1, 2, \cdots, 2^n - 1$ as the (approximate) solution of our problem.

The practical matrix solution of (8) is described in the sequel.

Consider the matrices

$$X \stackrel{\text{def}}{=} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ & & \vdots & \\ b_{l1} & b_{l2} & \cdots & b_{lm} \end{pmatrix}, \text{of type } (l, m = 2^n - 1)$$

$$Y \stackrel{\text{def}}{=} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_l \end{pmatrix}, \text{of type } (l, 1)$$

and the (approximate) matrix solution

$$A = \begin{pmatrix} t_1^0 \\ t_2^0 \\ \vdots \\ t_m^0 \end{pmatrix}, \text{of type } (m = 2^n - 1, 1)$$

It is seen that (8) says that

$$X^T X A = X^T Y \tag{9}$$

In the (desirable) particular case when $X^T X$ is invertible, the solution A is given via

$$A = (X^T X)^{-1} X^T Y \tag{10}$$

Caution: It is advisable to work for $l \geq 2^n - 1$, i.e. one must perform a sufficiently large amount of observations. The reason for this restriction is given by the fact that

$$rank(X^T X) \leq min(rank(X^T), rank(X)) = rank(X) \leq l$$

So, in case $l < 2^n - 1$, it follows that $X^T X$ is not invertible and the last formula for A (i.e. formula (10)) cannot be used.

Correction of the results:

The approximate solution ν must be a monotone measure. The verification of this fact is the following: for any $E \in P^*(T)$, $E \neq T$, one must have

$$0 \leq \nu(E) \leq \nu(E \cup \{x_i\}), \text{for any } x_i \in T \setminus E$$

If this is not the case (i.e. negative values $\nu(E)$ apper, or violation of the monotonicity occurs), we modify ν as follows (it is possible for some of subsequent steps to be unnecessary):

1. First, all (possible) strictly negative values $\nu(E)$ are replaced by the value 0 (hence all $\nu(G)$, with $G \subset E$, become 0 too).

2. Next, if there exist E, F in $P^*(T)$ such that $E \subset F$ and $\nu(E) > \nu(F)$, then we modify $\nu(E)$, replacing the value $\nu(E)$ with

$$max\{\nu(G) | G \subset E \text{ and } \nu(G) \leq min\{\nu(H) | E \subset H\}\}$$

(the set under max is not empty, containing $0 = \nu(\emptyset)$).

3. The procedure at 2. continues for the modified $\nu(E)$ until "wrong" pairs (E, F) as above do not appear any more.

Finally, we obtain the monotone measure μ which will be called "the good measure".

4 Determination of the anxiety degree

Main idea of the paper is introduced at the beginning of the Introduction.

4.1 Creating the instrument (Measure). The inverse problem

We made measurements of EEG waves for $l = 70$ subjects (we think, $l = 70$ is sufficient to furnish a credible measure). For each subject, the measurement values were provided in a CSV file. Now, we construct $l = 70$ functions as follows. We consider the total set $T = \{x_1, x_2, x_3, x_4\}$, where the measured attributes are $x_1 = LowAlpha, x_2 = HighAlpha, x_3 = LowBeta, x_4 = HighBeta$. For any $p = 1, 2, \cdots, 70$, we have the function $f_p : T \to \mathbb{R}_+$, where f_p corresponds to the subject S_p and the values $f_p(x_i)$, $i = 1, 2, 3, 4$ are determined as follows: we extract from the CSV file the set of the values pertaining to i for S_p and we compute the average of these values, which is exactly $f_p(x_i)$. Also, for any $p = 1, 2, \cdots, 70$ we have the value $y_p =$ the level of anxiety of S_p, given by the psychologists. We obtained Table 1with 70 rows and 4+1 = 5 columns.

Namely, row number p contains the following input values: $f_p(x_1), f_p(x_2), f_p(x_3), f_p(x_4), y_p$. It is important to mention that the input values y_p (level of anxiety of S_p) are grades from 0 to 6 (meaning that 0 is the least anxious and 6 the most anxious).

As we have said, we decided to choose as fusion instrument, the Choquet integral of the functions f_p, $p = 1, 2, \cdots, 70$, with respect to a virtual (unknown) measure μ which will be determined. Namely, for any $p = 1, 2, \cdots, 70$ one must have (similar to (6)):

$$y_p = (C) \int f_p d\mu \qquad (11)$$

In the C++ program, we used a function to process the data from the CSV files, and to create a matrix.

In order to save typographical space, we exhibit below only one row of the Table 1

Table 1: Sample data

Low Alpha	**High Alpha**	**Low Beta**	**High Beta**	**Grade**
41482.67	27173.99	18173.78	19565.17	6

Using the devices described in section 3, we determined the "measure" ν. It has been necessary to (slightly)modify ν (see correction of the Result, end of

Section 3) and we obtained the good measure μ, whose values $\mu(E)$, arranged in the aforementioned lexicographic order, are the following:

$$\mu(E) = \begin{cases} 0 & \text{if } E = \emptyset \\ 0.0000208904, & \text{if } E = \{x_1\} \\ 0.0000104939, & \text{if } E = \{x_2\} \\ 0.000047098, & \text{if } E = \{x_1, x_2\} \\ 0.0000398943, & \text{if } E = \{x_3\} \\ 0.000230985, & \text{if } E = \{x_1, x_3\} \\ 0.000142402, & \text{if } E = \{x_2, x_3\} \\ 0.000230985, & \text{if } E = \{x_1, x_2, x_3\} \\ 0.0000143691, & \text{if } E = \{x_4\} \\ 0.0000357572, & \text{if } E = \{x_1, x_4\} \\ 0.0000143691, & \text{if } E = \{x_2, x_4\} \\ 0.000047098, & \text{if } E = \{x_1, x_2, x_4\} \\ 0.0000398943, & \text{if } E = \{x_3, x_4\} \\ 0.000230985, & \text{if } E = \{x_1, x_3, x_4\} \\ 0.000142402, & \text{if } E = \{x_2, x_3, x_4\} \\ 0.000230985, & \text{if } E = \{x_1, x_2, x_3, x_4\} \end{cases}$$

The measure was calculated using the C++ program.

4.2 Using the obtained instrument (Measure) to determine the anxiety level of other subjects. The direct problem

We used the obtained measure μ to decide over the anxiety level of 10 new subjects.We made $l = 10$ measurements. The measurement values were provided in a CSV file. For each subject p, $p = 1, 2, \cdots, 10$, we obtained the values $f_p(x_i)$, $i = 1, 2, 3, 4$ as previously. In this case the input data are $f_p(x_i)$ and the measure μ. We also have the conclusions of psychologists for these new topics with which we will compare our results. Namely, measuring these $l = 10$ new subjects, we obtained the values $f_p(x_1)$,$f_p(x_2)$,$f_p(x_3)$,$f_p(x_4)$,$p = 1, 2, \cdots, 10$. See Table 2

Table 2: New subjects

No	Low Alpha	High Alpha	Low Beta	High Beta
1.	33738.85	26911.79	15911.23	15827.22
2.	11360.2	8108.433	9596.968	11737.35
3.	35170.35	23910.76	17211.32	12380.47
4.	36224.77	27315.48	16978.18	21143.96
5.	10658.68	7869.542	8243.346	7708.505

Table 2: New subjects

No	Low Alpha	High Alpha	Low Beta	High Beta
6.	33411.1	22809.05	21446.92	14769.47
7.	36568.82	23903.5	14545.97	19463.21
8.	33738.85	26911.79	15911.23	15827.22
9.	41482.67	27173.99	18173.78	19565.17
10.	51089.42	27487.33	19538.43	20055.94

The classification given by the psychologists to the subjects (measured individuals), is represented by grades, from 0 to 6 (meaning that 0 is the least anxious and 6 the most anxious). See Table 3:

Table 3: Grades

No	Grades
1.	4
2.	2
3.	5
4.	5
5.	2
6.	5
7.	4
8.	4
9.	5
10.	5

Using the formula (11), we obtained the conclusions(the output data):

$$\begin{aligned}
y_1 &= 4.33598 \approx 4,\\
y_2 &= 2.28522 \approx 2,\\
y_3 &= 4.52631 \approx 5,\\
y_4 &= 4.59469 \approx 5,\\
y_5 &= 1.95455 \approx 2,\\
y_6 &= 5.23955 \approx 5,\\
y_7 &= 4.06521 \approx 4,\\
y_8 &= 4.33598 \approx 4,\\
y_9 &= 4.92068 \approx 5,\\
y_{10} &= 5.38052 \approx 5.
\end{aligned}$$

As can be seen, the resulting conclusions are very close to the conclusions provided by psychologists.

This procedure can be continued for any new subject, given that the results obtained in this way are very good.

5 Conclusions

1. Because the anxiety is characterized by LowAlpha, HighAlpha, LowBeta and HighBeta (as psychologists say), the measurement and variation of them will give the level of anxiety and its changes.

2. We mention that we started from the premise that the studies, the discoveries and data provided by the psychology specialists with whom we collaborated are certain. The mathematical models described in this chapter are fixed, so they can be easily adapted to the new discoveries of psychologists in the fields studied by us.

3. We have determined an aggregation model, using the data resulted from the measurements. The aggregation model consists in a monotone measure and the Choquet integration of the input data with respect to this measure and is completely new.

4. The aggregation model was used to draw conclusions regarding the level of anxiety of new subjects measured with NeuroSky.

5. The created model was validated by comparing its values with the ones obtained via classical methods (psychological tests).

6. Future studies can be considered, e.g. studies concerning epilepsy, or studies with a NeuroSky device with two active sensors.

7. The authors declare that all subjects gave their informed consent for inclusion before they participated in the study.

Acknowledgements

1. The authors are very much indebted to the editors and the reviewers for their most valuable suggestions which improved a lot the quality of the paper.

2. In writing this paper, we have worked in collaboration with the Institute of Studies, Research, Development and Innovation of Titu Maiorescu University in Bucharest, as well as with specialists of the Military Technical Academy in Bucharest. To all of them, our warm thanks.

6 Appendix: C++ functions

In the C++ programme we used one function to process the data from the csv files. Then we use another function to create a matrix of type MatrixXld.

The function *LoadEigenMatrix* loads from a CSV (comma separated values) type file the numerical values from that file, with the option to specify whether or not that file has the names of the columns on the first line. The returned result is a matrix of type MatrixXld.

The function *GetRegularData* loads from two directories on disk from the files corresponding to the measured parameters. The files corresponding to the results must have the same name on disk as the files corresponding to the measured results. The result is a matrix of type MatrixXld.

The matrix type MatrixXld from the Eigen library corresponds to a dynamically allocated matrix with long double values of 128 bytes.

```
MatrixXld LoadEigenMatrix(const char* fn, bool ignore_header)
    {
    fstream file;
    string line;
    MatrixXld mat;
    long row = 0;
    file.open(fn, ios_base::in);
    if (ignore_header)
    {
    getline(file, line);
    }
    mat = MatrixXld::Constant(1,1,0);
    string oline;
    while (getline(file,line))
    {
    row ++;
    int cnt;
    char* number;
    cnt = 0;
    number = strtok((char* )line.c_str(),",");
    while (number != NULL)
    {
    cnt ++;
    if (cnt > mat.cols())
    {
    mat.conservativeResize(NoChange, cnt);
```

```
mat(0, cnt-1) = 0;
}
long double num;
if (strcmp(number, "") == 0)
num = 0;
else
{
try
{
num = (long double)strtod((const char*)number, nullptr);
}
catch (exception e)
{
cout « "Eroare la conversia elementului (" « row « "," « cnt « ")" « endl;
}
}
mat(0, cnt - 1) = mat(0, cnt - 1) + num;
number = strtok(NULL, ",");
}
}
for (int i=0; i<mat.cols(); i++) mat(0,i) = (mat(0,i) / ((long double) row));
file.close();
return mat;
}
```

The C++ program using the function *getBinInv* which constructs a matrix of type MatrixXld with t lines and n columns, representing t subsets of a given set with n elements, each subset being represented by a line in which the elements present among the n are represented with 1, and the missing ones with 0.

```
MatrixXld getBinInv (int t, int n)
{
MatrixXld bin = MatrixXld::Constant(t, n, 0);
for (int i = 0; i < t; i++)
{
for (int j = 0; j < n; j++)
{
int val = (((i + 1)»(j)) & 1);
bin(i, j) = (long double)val;
}
```

```
}
return bin;
}
```

We used the functions: *create_measure*, *is_null_matrix*, *apply_measure*.

The function *create_measure* creates a measure in the form of a matrix of type MatrixXld having as parameters a matrix resulting from calling the*getRegularData* function, the number of measured parameters, a vector representing the columns corresponding to the desired parameters from the given matrix and the column corresponding to the result.

```
MatrixXld create_measure(MatrixXld m, int n, int v[], int result_col)
{
int t = (int) (pow(2,n) - 1);
int l = 0;
l = m.rows();
MatrixXld date = MatrixXld::Constant(l, n, 0);
MatrixXld y = MatrixXld::Constant(l, 1, 1);
for(int i =0; i<n; i++)
date.block(0,i,l,1) = m.block(0,v[i],l,1);
y.block(0,0,l,1) = m.block(0,result_col,l,1);
MatrixXld bininv = getBinInv(t, n);
MatrixXld b = MatrixXld::Constant(l, t, 0.0);
for (int p = 0; p < l; p++)
{
for (int j = 0; j < t; j++)
{
long double emin = std::numeric_limits<long double>::max();
long double emax = std::numeric_limits<long double>::min();
for (int i = 0; i < n; i++)
{
if (bininv(j, i) == 1.0)
{
if (date(p,i) < emin) emin = date(p, i);
}
if (bininv(j, i) == 0.0)
{
if (date(p,i) > emax)
emax = date(p, i);
}
}
```

```
if (j == (t - 1))
emax = 0.0;
if (emin - emax > 0.0)
b(p,j) = emin - emax;
else b(p,j) = 0.0;
}
}
fstream fout;
fout.open("b_openness.csv", ios_base::out);
fout « b « endl;
fout.close();
MatrixXld a(t,t);
a = b.transpose() * b;
fout.open("b_b_trans_openness.csv", ios_base::out);
fout « a « endl;
fout.close();
long double det = 0;
det = (long double)a.determinant();
cout« "determinant = " « det « endl;
MatrixXld miu = MatrixXld::Constant(t,1, 0.0);
if (det > 0 )
{
miu = a.inverse() * b.transpose() * y;
for (int i=0; i<t; i++)
if (miu(i,0) < 0)
miu(i,0) = 0;
}
return miu;
}
```

The function *is_null_matrix* checks if the given matrix has only 0 elements and returns the boolean value corresponding to this check.

```
bool is_null_matrix(MatrixXld x)
{
bool isnull = true;
for (int i=0; i<x.rows(); i++)
for (int j=0; j<x.cols(); j++)
{
if (x(i,j) != 0.0)
{
```

```
isnull = false;
break;
}
}
return isnull;
}
```

The function *apply_measure* applies the measure given as a parameter to a set of measured parameters presented as an array resulting from calling the *getRegularData* function with the second empty string parameter. The number of parameters used from the given matrix and a vector representing the column positions corresponding to those parameters are also specified. The result is returned as a matrix of type MatrixXld.

```
MatrixXld apply_measure(MatrixXld m3, MatrixXld measure, int n, int v[])
{
int l = m3.rows();
MatrixXld date = MatrixXld::Constant(l, n, 0);
for(int i =0; i<n; i++)
date.block(0,i,l,1) = m3.block(0,v[i],l,1);
int t = (int) (pow(2,n) - 1);
MatrixXld b = MatrixXld::Constant(l, t, 0.0);
MatrixXld bininv = getBinInv(t, n);
for (int p = 0; p < l; p++)
{
for (int j = 0; j < t; j++)
{
long double emin = std::numeric_limits<long double>::max();
long double emax = std::numeric_limits<long double>::min();
for (int i = 0; i < n; i++)
{
if (bininv(j, i) == 1.0)
{
if (date(p,i) < emin)
emin = date(p, i);
}
if (bininv(j, i) == 0)
{
if (date(p,i) > emax) emax = date(p, i);
}
}
```

```
if (j == (t - 1))
emax = 0;
if (emin - emax > 0)
b(p,j) = emin - emax;
else
b(p,j) = 0;
}
}
return (b * measure);
}
```

Using the function *isMonotoneMeasure*, we check if the calculated measurement is monotonous.

```
bool isMonotoneMeasure(MatrixXld mat, int n)
{
bool is_monotone = true;
int i,j,aux,k,lin = mat.rows(),col=mat.cols();
for(j=0;j<=lin-1;j++)
for(i=1;i<=col-1;i++)
if((j»(n-i-1))==0)
{
aux=1«(n-i-1);
k=aux+j;
if(mat(j,1)>mat(k,1))
is_monotone = false;
}
return is_monotone;
}
```

References

[1] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, DSM-5, 5th edition, Washington DC, (2013).

[2] O.M.Bazanova, D.Vernon, Interpreting EEG alpha activity, Neuroscience and Biobehavioral Reviews, 44 (2014), 94-110.

[3] L. Chang, Application of Fuzzy Comprehensive Evolution Based on Genetic Algorithm in Psychological Measurement, Scientific Programming for Smart Internet of Things (special issue), 2021, Article ID 9607006.

[4] I.Chiţescu, A.Plăviţu, Computing Choquet integrals, Fuzzy Sets and Systems, 327 (2017), 46-68.

[5] G. Choquet, Theory of capacities, Annales de L'Institut Fourier, 5 (1953-1954), 131-295.

[6] D. Denneberg, M. Grabisch, Measure and integral with purely ordinal scales, J. Math. Psychology, 48 (2004), 15-27.

[7] G.Gennady, G.Knyazev, EEG correlates of personality types, Netherlands Journal of Psychology, 62(2) (2006), 78-87.

[8] M.Grabisch, Set Functions, Games and Capacities in Decision Making, Springer, (2016).

[9] A.M.Grass, F.A.Gibbs, A Fourier transform of the electroencephalogram, J.Neurophysiol, 1 (1938), 521-526.

[10] T.Harmony, The functional significance of delta oscillations in cognitive processing, Frontiers in Integrative Neuroscience, https://doi.org/10.3389/fnint.2013.00083

[11] D.E.Hinkle, W.Wiersma, S.G.Jurs, Applied Statistics for the Behavioral Sciences, 5th edition, Houghton Mifflin Company, Boston, New York, (2003).

[12] N. Kane, J. Acharya, S. Benickzy, L. Caboclo, S. Finnigan, P. W. Kaplan, H. Shibasaki, R. Pressler, M. J A M van Putten, A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings, Revision 2017, Corrected 2019, Clin.Neurophy.Pract, 4 (2019), 133.

[13] J.M. Keller, J. Osborn, Training the Fuzzy Integral, International Journal of Approximate Reasoning, 15 (1996), 1-24.

[14] G.Knyazev, E.Merkulova, A.Savostyanov, A.Bocharov, A.Saprigyn, Personality and EEG correlates of reactive social behavior, Neuropsychologia, 124 (2019), 98-107.

[15] E. Niedermeyer and F. Lopes da Silva, eds, Electroencephalography, Basic Principles, Clinical Applications and Related Fields, 4th edition. Lippincott, Williams and Wilkins, Philadelphia, Pennsylvania, (1999).

[16] E.Pap (editor), Handbook of Measure Theory, two volumes, Elsevier, North Holland, (2002).

[17] D.J.Paulus, S.Vanwoerden, P.J.Norton, C.Sharp, From neuroticism to anxiety. Examining unique contributions of three transdiagnostic vulnerability factors, Personality and Individual Differences, 94 (2016), 38-43.

[18] H.W.Shipton, EEG analysis: a history and a prospectus, Annual Review of Biophysics and Bioengineering, 4 (1975), 1-15.

[19] H.-H. Tsai, I.-Y. Lu, The evaluation of service quality using generalized Choquet integral, Information Sciences, 176 (2006), 640-663.

[20] Z.Wang, G. J. Klir, Generalized Measure Theory, Springer, (2009).

[21] Z.Wang, R. Yang, K-S. Leung, Nonlinear Integrals and Their Applications in Data Mining, World Scientific, (2010).

Ion Chiţescu,
Department of Mathematics,
National University of Science and Technology Politehnica Bucharest,
Splaiul Independenței 313, Bucharest, 060042, Romania
Email: ionchitescu@yahoo.com

Mădălina Giurgescu (Manea) - Corresponding Author,
Department of Mathematics,
National University of Science and Technology Politehnica Bucharest, Piteşti University Center,
Târgul din Vale 1, Piteşti, 110040, Romania
Email: madalina_giurgescu@yahoo.com

Titi Paraschiv,
Department of Computer Science,
Military Technical Academy of Bucharest,
George Coșbuc 39-49, Bucharest, 050141, Romania
Email: titi.paraschiv@prof.utm.ro