

Multi-Dimensional Fuzzy Transforms for Attribute Dependencies

Ferdinando Di Martino^{1,2} Vincenzo Loia² Salvatore Sessa¹

¹Dipartimento di Costruzioni e Metodi Matematici in Architettura, Università degli Studi di Napoli Federico II,
80134 Napoli, Italy

²Dipartimento di Matematica e Informatica, Università degli Studi di Salerno,
84084 Salerno, Italy
Email: fdimarti@unina.it, loia@unisa.it, sessa@unina.it

Abstract — We explore attribute dependencies in the datasets by using direct and inverse fuzzy transforms. Our algorithm optimizes the fuzzy partitions of the universe of the attributes and moreover establishes if the set of the data points is sufficiently dense with respect to the chosen partitions: two specific regression indexes measure the reliability of our model. The known “El Nino” dataset is the basis of our experiments, whose results are consistent with the regression analysis made with the same data.

Keywords — attribute dependence, fuzzy transform, index of determinacy, regression.

1 Introduction

Regression is a well known statistical supervised technique used in data mining (cfr. [5, 6, 7, 8, 9]). The dependency between numerical attributes is studied via an equation of regression: the dependent attribute is modeled as a function of the independent attributes or predictors. We present a fast regression algorithm in which we use the fuzzy transforms (for short, F-transforms) for exploring numerical attribute dependencies in datasets. A *direct fuzzy transform* [1, 11, 12, 13, 14] gives a correspondence between a set of continuous functions defined on the interval $[a, b]$ and a set of n -dimensional vectors defined suitably. An *inverse fuzzy transform* gives the converse correspondence, that is an n -dimensional vector is transformed into a continuous function which approximates the original function up to a small quantity ϵ . The method based on the F-transforms has been used in other topics like image processing [2, 3, 4, 12], geology [11].

In [10] the attribute dependencies in data analysis are established via F-transforms: the dependency of an attribute X_z from the attributes X_1, \dots, X_k is made by setting $X_z = H(X_1, \dots, X_k)$, where $H: [a_1, b_1] \times \dots \times [a_k, b_k] \rightarrow [a_z, b_z]$ is a continuous function and $[a_i, b_i]$ is the domain of X_i , $i = 1, \dots, k$. The approximating function $H_{F,n}$ is obtained via the *discrete F-transforms* and the difference between H and $H_{F,n}$ at the their common points is evaluated by means of the index of determinacy. We devote particular attention to the solution of the following problems:

a) to find the best fuzzy partition of the attribute domains in the construction of the direct F-transform. We observe that the index of determinacy calculated with coarse grained partitions is less than that one calculated with finer partitions;

b) our method analyzes if the set of the assigned data points is sufficiently dense w. r. t. the chosen partitions.

Our algorithm makes a fuzzy partition of $[a_i, b_i]$, $i=1, \dots, k$, in n fuzzy sets and establishes a threshold value for the indexes of determinacy after some training tests. The attribute dependency is found if the index overcomes the threshold value, otherwise we pass to a partition of $[a_i, b_i]$ in $n+1$ fuzzy sets by checking additionally if the set of the assigned data points is sufficiently dense w.r.t. the chosen partition. Of course if this condition is violated, then the process is stopped.

In Section 2 we give the definitions of F-transforms in k (≥ 2) variables. In Section 3 we present our process, in Section 4 a sample simulation shows how to find the threshold value for the indexes of determinacy. The known dataset “El Nino” is the basis of our tests contained in Section 5. Section 6 concludes this paper.

2 Multi-dimensional F-transforms

In accordance to [12], let $n \geq 2$ and x_1, x_2, \dots, x_n be points of $[a, b]$, called nodes, such that $x_1 = a < x_2 < \dots < x_n = b$. The fuzzy sets $A_1, \dots, A_n : [a, b] \rightarrow [0, 1]$, called basic functions, form a *fuzzy partition* of $[a, b]$ if the following hold:

- (1) $A_i(x_i) = 1$ for every $i = 1, 2, \dots, n$;
- (2) $A_i(x) = 0$ if $x \notin (x_{i-1}, x_{i+1})$ for every $i = 2, \dots, n$;
- (3) $A_i(x)$ is a continuous function on $[a, b]$ for every $i = 1, 2, \dots, n$;
- (4) $A_i(x)$ strictly increases on $[x_{i-1}, x_i]$ for every $i = 2, \dots, n$ and strictly decreases on $[x_i, x_{i+1}]$ for every $i = 1, \dots, n-1$;

$$(5) \sum_{i=1}^n A_i(x) = 1 \text{ for every } x \in [a, b].$$

The partition $\{A_1(x), \dots, A_n(x)\}$ is said *uniform* if moreover the following hold:

- (6) $n \geq 3$ and the nodes are equidistant, that is $x_i = a + h \cdot (i-1)$ for every $i = 1, 2, \dots, n$ where $h = (b-a)/(n-1)$;
- (7) $A_i(x_i - x) = A_i(x_i + x)$ for every $x \in [0, h]$ and $i = 2, \dots, n-1$;

(8) $A_{i+1}(x) = A_i(x-h)$ for every $x \in [x_i, x_{i+1}]$ and $i = 1, \dots, n-1$.

By limiting ourselves to the discrete case (cfr. [12] in the case of a continuous function), let f be a real function defined on assigned points p_1, \dots, p_m of $[a, b]$. Let P be the set of these points and suppose that it is *sufficiently dense* w. r. t. the fixed partition $\{A_1, A_2, \dots, A_n\}$, that is for every $i = 1, \dots, n$ there exists an index $j \in \{1, \dots, m\}$ such that $A_i(p_j) > 0$. Thus we define the numerical vector $[F_1, F_2, \dots, F_n]$ as the *direct F-transform* of f w.r.t. the basic functions $\{A_1, A_2, \dots, A_n\}$, where each F_i is given by

$$F_i = \frac{\sum_{j=1}^m f(p_j) A_i(p_j)}{\sum_{j=1}^m A_i(p_j)} \quad (1)$$

for every $i=1, \dots, n$. Then we define the *inverse F-transform* of f w. r. t. $\{A_1, A_2, \dots, A_n\}$ by setting for every $j = 1, \dots, m$:

$$f_{F,n}(p_j) = \sum_{i=1}^n F_i A_i(p_j) \quad (2)$$

Then the following theorem [12] holds:

Theorem 1. Let $f(x)$ be a real function assigned on a set $P = \{p_1, \dots, p_m\} \subseteq [a, b]$. Then for every $\varepsilon > 0$, there exist an integer $n(\varepsilon)$ and a related fuzzy partition $\{A_1, A_2, \dots, A_{n(\varepsilon)}\}$ of $[a, b]$ such that P is sufficiently dense w. r. t. $\{A_1, A_2, \dots, A_{n(\varepsilon)}\}$ and for every $p_j \in [a, b]$, $j = 1, \dots, m$, the following inequality holds:

$$|f(p_j) - f_{F,n(\varepsilon)}(p_j)| < \varepsilon \quad (3)$$

As in [15], we extend these definitions to the multi-dimensional case. Let k be given intervals $[a_i, b_i]$ ($i=1, \dots, k$) and $[a_1, b_1] \times [a_2, b_2] \times \dots \times [a_k, b_k]$ be our universe of discourse. Let $x_{11}, x_{12}, \dots, x_{1n_1} \in [a_1, b_1], \dots, x_{k1}, x_{k2}, \dots, x_{kn_k} \in [a_k, b_k]$ be $n_1 + \dots + n_k$ assigned points, called nodes, such that $x_{i1} = a_i < x_{i2} < \dots < x_{in_i} = b_i$ for every $i = 1, \dots, k$. Let $\{A_{i1}, A_{i2}, \dots, A_{in_i}\}$ be a fuzzy partition of $[a_i, b_i]$ for every $i = 1, \dots, k$ and $f(x_1, x_2, \dots, x_k)$ be a real function by assuming determined values in m points $p_j = (p_{j1}, p_{j2}, \dots, p_{jk}) \in [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_k, b_k]$ for every $j=1, \dots, m$. We say that the set $P = \{(p_{11}, p_{12}, \dots, p_{1k}), (p_{21}, p_{22}, \dots, p_{2k}), \dots, (p_{m1}, p_{m2}, \dots, p_{mk})\}$ is *sufficiently dense* w. r. t. the chosen partitions $\{A_{11}, A_{12}, \dots, A_{1n_1}\}, \dots, \{A_{k1}, A_{k2}, \dots, A_{kn_k}\}$ if for each k -tuple $\{h_1, \dots, h_k\} \in \{1, \dots, n_1\} \times \dots \times \{1, \dots, n_k\}$ there exists a point $p_j = (p_{j1}, p_{j2}, \dots, p_{jk}) \in P$, $j \in \{1, \dots, m\}$, such that $A_{1h_1}(p_{j1}) \cdot A_{2h_2}(p_{j2}) \cdot \dots \cdot A_{kh_k}(p_{jk}) > 0$. Then we can

define the (h_1, h_2, \dots, h_k) -th component $F_{h_1 h_2 \dots h_k}$ of the *direct F-transform* of f w.r.t. the basic functions $\{A_{11}, A_{12}, \dots, A_{1n_1}\}, \dots, \{A_{k1}, A_{k2}, \dots, A_{kn_k}\}$ as

$$F_{h_1 h_2 \dots h_k} = \frac{\sum_{j=1}^m f(p_{j1}, p_{j2}, \dots, p_{jk}) \cdot A_{1h_1}(p_{j1}) \cdot A_{2h_2}(p_{j2}) \cdot \dots \cdot A_{kh_k}(p_{jk})}{\sum_{j=1}^m A_{1h_1}(p_{j1}) \cdot A_{2h_2}(p_{j2}) \cdot \dots \cdot A_{kh_k}(p_{jk})} \quad (4)$$

The *inverse F-transform* of f w.r.t. the basic functions $\{A_{11}, A_{12}, \dots, A_{1n_1}\}, \dots, \{A_{k1}, A_{k2}, \dots, A_{kn_k}\}$ is the following function defined for each point $p_j = (p_{j1}, p_{j2}, \dots, p_{jk}) \in [a_1, b_1] \times \dots \times [a_k, b_k]$ as

$$f_{F, n_1 n_2 \dots n_k}^F(p_{j1}, p_{j2}, \dots, p_{jk}) = \sum_{h_1=1}^{n_1} \sum_{h_2=1}^{n_2} \dots \sum_{h_k=1}^{n_k} F_{h_1 h_2 \dots h_k} \cdot A_{1h_1}(p_{j1}) \cdot \dots \cdot A_{kh_k}(p_{jk}) \quad (5)$$

for every $j=1, \dots, m$. The following extension of Theorem 1 holds:

Theorem 2. Let $f(x_1, x_2, \dots, x_k)$ be a function assigned on the set of points $P = \{(p_{11}, p_{12}, \dots, p_{1k}), (p_{21}, p_{22}, \dots, p_{2k}), \dots, (p_{m1}, p_{m2}, \dots, p_{mk})\} \subseteq [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_k, b_k]$. Then for every $\varepsilon > 0$, there exist k integers $n_1(\varepsilon), \dots, n_k(\varepsilon)$ and fuzzy partitions $\{A_{11}, A_{12}, \dots, A_{1n_1(\varepsilon)}\}, \dots, \{A_{k1}, A_{k2}, \dots, A_{kn_k(\varepsilon)}\}$

such that the set P is sufficiently dense w.r.t. the above fuzzy partitions and for every $p_j = (p_{j1}, p_{j2}, \dots, p_{jk}) \in P$, $j=1, \dots, m$, the following inequality holds:

$$\left| f(p_{j1}, p_{j2}, \dots, p_{jk}) - f_{n_1(\varepsilon) n_2(\varepsilon) \dots n_k(\varepsilon)}^F(p_{j1}, p_{j2}, \dots, p_{jk}) \right| < \varepsilon \quad (6)$$

The proof follows the same lines of the similar Theorem 1.

3 Attribute dependencies

We modify slightly the algorithm of [10], where the data are represented in the following matrix:

	X_1	\dots	X_i	\dots	X_r
O_1	p_{11}	\cdot	p_{1i}	\cdot	p_{1r}
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
O_j	p_{j1}	\cdot	p_{ji}	\cdot	p_{jr}
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
O_m	p_{m1}	\cdot	p_{mi}	\cdot	p_{mr}

being $X_1, \dots, X_i, \dots, X_r$ the attributes, $O_1, \dots, O_j, \dots, O_m$ the objects and p_{ji} the value of the attribute X_i . By setting $a_i = \min\{p_{1i}, \dots, p_{mi}\}$ and $b_i = \max\{p_{1i}, \dots, p_{mi}\}$, we define the interval $[a_i, b_i]$ of the given values of the attribute X_i . We can consider the attribute X_z , $z \in \{1, \dots, r\}$, dependent from k independent attributes X_1, \dots, X_k , (with $k \leq r < m$ and $z \notin \{1, \dots, k\}$) via a function H defined as

$$X_z = H(X_1, \dots, X_k) \quad (7)$$

We use the F-transforms method in the following process:

1) We assume $n_1 = n_2 = \dots = n_k = n$ without loss of generality and we refer to the model (7). After normalization in $[0,1]$ of the values of the independent and dependent attributes, we start with $n = 3$.

2) A uniform fuzzy partition $\{A_{i1}, A_{i2}, \dots, A_{in_i}\}$ is assigned in each interval $[a_i, b_i]$ by setting for every $i = 1, \dots, k$ and $j = 2, \dots, k-1$:

$$A_{i1}(x) = \begin{cases} 0.5 \cdot (1 + \cos \frac{\pi}{h_i}(x - x_{i1})) & \text{if } x \in [x_{i1}, x_{i2}] \\ 0 & \text{otherwise} \end{cases}$$

$$A_{ij}(x) = \begin{cases} 0.5 \cdot (1 + \cos \frac{\pi}{h_i}(x - x_{ij})) & \text{if } x \in [x_{i(j-1)}, x_{i(j+1)}] \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$A_{in}(x) = \begin{cases} 0.5 \cdot (1 + \cos \frac{\pi}{h_i}(x - x_{in})) & \text{if } x \in [x_{i(n-1)}, x_{in}] \\ 0 & \text{otherwise} \end{cases}$$

where $h_i = (b_i - a_i)/(n - 1)$ and $x_{ij} = a_i + h_i \cdot (j-1)$.

3) For each k -tuple $\{h_1, \dots, h_k\} \in \{1, \dots, n\}^k$ we see that exists at least an object O_j with values $(p_{j1}, p_{j2}, \dots, p_{jk})$ such that $A_{h_1}(p_{j1}) \cdot A_{h_2}(p_{j2}) \cdot \dots \cdot A_{h_k}(p_{jk}) > 0$. If this condition is not verified, the set $P = \{(p_{j1}, p_{j2}, \dots, p_{jk}), j=1, \dots, m$ is not sufficiently dense w. r. t. the basic functions (8) and hence the process is stopped.

4) For simplicity, we put $H(p_{j1}, p_{j2}, \dots, p_{jk}) = p_{jz}$ for every $j=1, 2, \dots, m$. In accordance to (4), the (h_1, h_2, \dots, h_k) -th component $F_{h_1 h_2 \dots h_k}$ of the direct F-transform of H is defined as

$$F_{h_1 h_2 \dots h_k} = \frac{\sum_{j=1}^m p_{jz} \cdot A_{h_1}(p_{j1}) \cdot \dots \cdot A_{h_k}(p_{jk})}{\sum_{j=1}^m A_{h_1}(p_{j1}) \cdot \dots \cdot A_{h_k}(p_{jk})} \quad (9)$$

Then, in accordance to (5), the inverse F-transform $H_{n_1 n_2 \dots n_k}^F$ of H is defined as

$$H_{n_1 n_2 \dots n_k}^F(p_{j1}, p_{j2}, \dots, p_{jk}) = \sum_{h_1=1}^{n_1} \sum_{h_2=1}^{n_2} \dots \sum_{h_k=1}^{n_k} F_{h_1 h_2 \dots h_k} \cdot A_{h_1}(p_{j1}) \cdot \dots \cdot A_{h_k}(p_{jk}) \quad (10)$$

5) The difference between $H_{n_1 n_2 \dots n_k}^F$ and H (or X_z) is estimated in the points $(p_{j1}, p_{j2}, \dots, p_{jm}), j=1, \dots, m$, via the following statistical indexes of determinacy:

$$r_c^2 = \frac{\sigma_{H^F}^2}{\sigma_{X_z}^2} = \frac{\sum_{j=1}^m (H_{n_1 n_2 \dots n_k}^F(p_{j1}, p_{j2}, \dots, p_{jk}) - \hat{p}_z)^2}{\sum_{j=1}^m (p_{jz} - \hat{p}_z)^2} \quad (11)$$

$$r_c'^2 = 1 - \left[(1 - r_c^2) \cdot \frac{m-1}{m-k-1} \right] \quad (12)$$

where \hat{p}_z is the mean of the values of X_z . We also look at a threshold value α determined experimentally: if r_c^2 and $r_c'^2$ are less or equal to α , then we consider correct our model (7) and stop the process, otherwise we set $n \leftarrow n+1$ and go back to 2). We fix the threshold α at a confidence value.

Strictly speaking, Theorem 2 is unable to evaluate the correctness of our model (7) because it does not give a method of calculation of the k indexes $n_1(\epsilon), \dots, n_k(\epsilon)$ and of the basic functions w. r. t. which the set $\{p_{j1}, p_{j2}, \dots, p_{jk}, j=1, \dots, m\}$ is sufficiently dense in such a way

$$\left| H(p_{j1}, p_{j2}, \dots, p_{jk}) - H_{n_1(\epsilon) \dots n_k(\epsilon)}^F(p_{j1}, p_{j2}, \dots, p_{jk}) \right| < \epsilon \quad (13)$$

In other words, the substitution of H obtained with $H_{n_1(\epsilon) \dots n_k(\epsilon)}^F$ (up to an arbitrary quantity ϵ) is possible by Theorem 2, but we use the statistical index (12) from a practical point of view. Indeed it allows additionally to compare various regression models that intend to explain the same dependent variable from a different number k of explanatory variables.

4 A sample simulation

The calculation of the threshold value α for the indexes of determinacy (11) and (12) is made with training tests. As sample example, we present the case of a dataset with 1000 records and two fields, that is $m=1000, r=2, k=1, X_1 =$ "Area of a circle", $X_2 =$ "Radius of a circle", $a_1 = 0, b_1 = \pi, a_2 = 0$ and $b_2 = 1$. Hence the values of the attributes are $p_{j1} \in [0, \pi]$ and $p_{j2} \in [0, 1]$ for every $j = 1, \dots, 1000$. If $z = 2$, we have $X_2 = H(X_1) = (X_1/\pi)^{1/2}$. In the interval $[0, \pi]$ we consider a maximum number of 500 basic functions of the type (8) since we have seen that the set $P = \{p_{j1} : j = 1, \dots, 1000\}$ is not sufficiently dense w. r. t. the chosen partitions for $n \geq 500$. Table 1 contains the values of the indexes r_c^2 and $r_c'^2$ for several values of n and it is plain that a good choice is $\alpha = 0.9$ for $n \leq 7$.

Table 1. Indexes of determinacy

n	r_c^2	r'^2_c
3	0.587000	0.588000
5	0.826898	0.827725
7	0.900207	0.901108
10	0.943000	0.944000
15	0.969749	0.970719
25	0.986085	0.987072
50	0.995084	0.996080
300	0.999000	1.000000
400	0.999000	1.000000
500	0.999000	1.000000

For sake of completeness, we show the plots in Figure 1 and Figure 2 of the values of X_2 and H_n^F for $n = 3$ and $n = 300$, respectively.

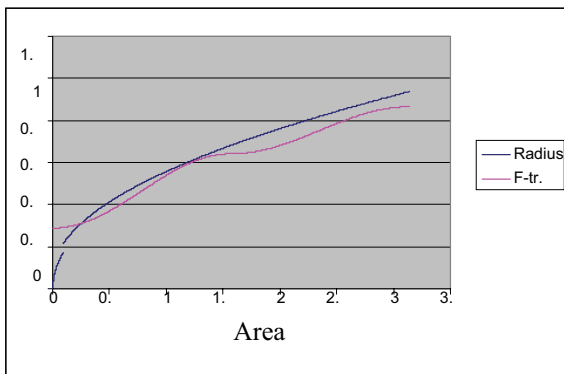


Figure 1. Graph of X_2 and H_3^F .

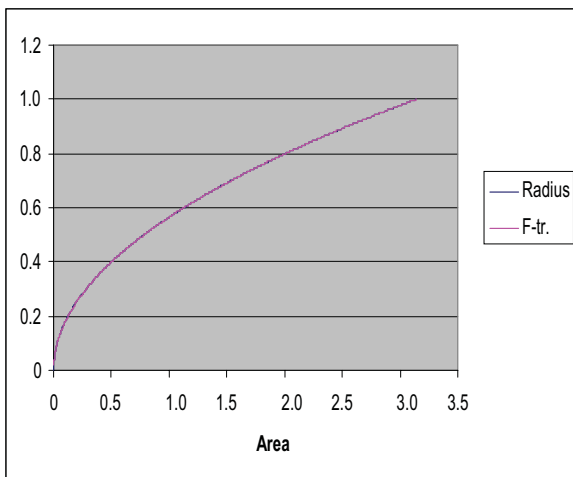


Figure 2. Graph of X_2 and H_{300}^F .

5 A complete experiment

We have downloaded from the known dataset “El Nino” (http://kdd.ics.uci.edu/databases/el_nino/el_nino.data.html) the oceanographic and surface meteorological data measured from a series of buoys positioned throughout the equatorial ocean Pacific. The data are formed from the following attributes:

- X_1 = date,
- X_2 = latitude,
- X_3 = longitude,
- X_4 = zonal winds (west < 0, east > 0),
- X_5 = meridional winds (south < 0, north > 0),
- X_6 = relative humidity,
- X_7 = air temperature in degrees Celsius,
- X_8 = sea surface temperature in degrees Celsius.

The data were measured from the buoys in various locations since 1980. The regression analysis establishes that significant relationships between the variables were not observed except for $X_8=H(X_7)$ and $X_8=H(X_1,X_7)$ which are of linear type like we will show in the sequel.

The training tests give $\alpha = 0.8$ as a threshold value. For single attribute dependencies, Table 2 shows the possible maximal number n_{max} of basic functions which can be used for deducing H_n^F . If n_{max} is overcome, the set of the attribute values is not sufficiently dense w. r. t. the chosen partitions. Moreover the absence of n_{max} in some entry of Table 2 means that the dependence model is not analyzable with the F-transforms method.

Table 2. n_{max} usable to explore single dependencies

	Date	Latit	Lon.	Zon. wind	Mer. wind	Rel. hum.	Air temp	Sea sur. tem..
Date	13	3	14	48	5	14	3	
Latitude	25	3	14	48	5	21	3	
Longit.	25	13	14	48	5	21	3	
Zonal winds	25	13	3	48	4	21	3	
Merid. winds	25	13	3	14		21	3	
Relative humid.	25	13		10	48		20	3
Air temper.	25	13	3	7	33			3
Sea sur. temper.	25	13	3	7	33	3	21	

The most meaningful model is $X_8=H(X_7)$ as Figure 3 shows: the plot is obtained for $n = 10$ and moreover we have $r_c^2 = r'^2_c = 0.83$. In the analysis of multiple attribute dependencies, Table 3 shows that only the model $X_8 = H(X_1,X_7)$ has the indexes of determinacy greater than $\alpha = 0.8$. Figure 4 shows the related 3D graph.

6 Conclusions

The F-transforms analyze all the possible dependencies between attributes in the datasets via a function H whose inverse F-transform is inside the value of the indexes of determinacy compared with a threshold parameter α . Our approach has the advantage to control always that the set of the attribute values is sufficiently dense w.r.t. the basic functions forming a uniform partition of the interval of context related to each attribute, otherwise the algorithm is stopped.

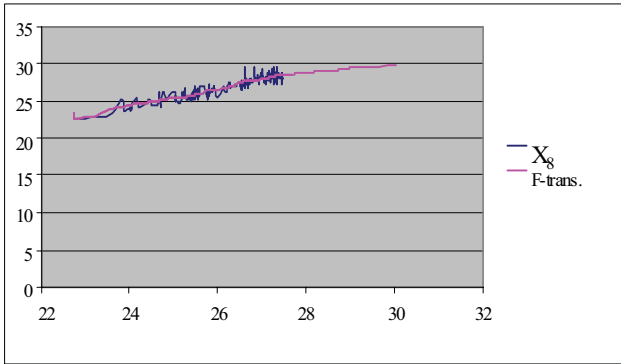


Figure 3. Graph of $X_8 = H(X_7)$ and H_{10}^F .

Table 3. Best values of indexes of determinacy for X_8

X_z	X_1, \dots, X_k	r_c^2	$r_c'^2$
X_8	X_1, X_7	0.827	0.828
X_8	X_6, X_7	0.341	0.338
X_8	X_4, X_5	0.266	0.266
X_8	X_1, X_6, X_8	0.110	0.111

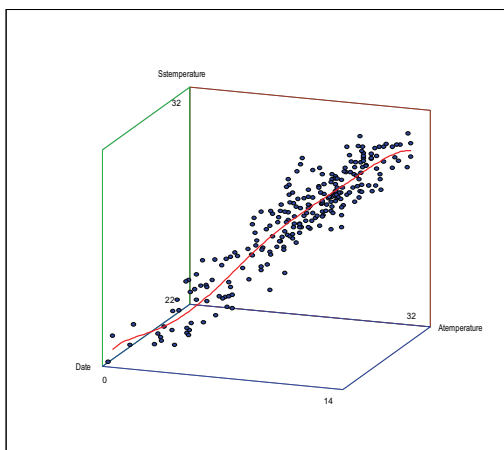


Figure 4. Plot 3D of $X_8 = H(X_1, X_7)$.

This method shall be integrated in future works with data mining classification and with results of regression data, such as decision tree learning and association mining algorithms.

References

[1] M. Daňková and M. Štěpnička, Fuzzy transform as an additive normal form, *Fuzzy Sets and Systems* 157 (8) (2006) 1024–1035.

[2] F. Di Martino, V. Loia, I. Perfilieva and S. Sessa, An image coding/decoding method based on direct and inverse fuzzy transforms, *Internat. Journal of Approximate Reasoning* 48 (2008) 110–131.

[3] F. Di Martino, V. Loia and S. Sessa, Fuzzy transforms for compression and decompression of color videos, *Information Sciences*, submitted.

[4] F. Di Martino and S. Sessa, Compression and decompression of images with discrete fuzzy transforms, *Information Sciences* 177 (2007) 2349–2362.

[5] N.R. Draper and H. Smith, *Applied Regression Analysis*, John Wiley & Sons Inc., New York, 1998.

[6] J. Han and M. Kamber, *Data Mining: Concepts and Techniques* (2nd Edition), Morgan-Kaufman, New York, 2006.

[7] R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis*, Prentice-Hall International, London, 1992.

[8] Y. S. Lee and S. J. Yen, Classification based on attribute dependency, Proceedings of 6th Internat. Conference DaWaK' 04 (Zaragoza, 1-3/9/04) *Lecture Notes in Computer Sciences*, Vol. 5192, 2004, 259 – 268.

[9] S. Mitra, S.K. Pal and P. Mitra, Data mining in soft computing framework: a survey, *IEEE Transactions on Neural Networks* 13 (1) (2002) 3–14.

[10] I. Perfilieva, V. Novák and A. Dvořák, Fuzzy transforms in the analysis of data, *Internat. Journal of Approximate Reasoning* 48 (2008) 36–46.

[11] I. Perfilieva, Fuzzy transforms: application to reef growth problem, in: R. B. Demicco, G. J. Klir (eds.), *Fuzzy Logic in Geology*, Academic Press, Amsterdam, 2003, pp. 275–300.

[12] I. Perfilieva, Fuzzy transforms: theory and applications, *Fuzzy Sets and Systems* 157 (8) (2006) 993–1023.

[13] I. Perfilieva, Fuzzy transforms: a challenge to conventional transforms, In: P.W. Hawkes (Ed.), *Advances in Images and Electron Physics*, vol. 147, Elsevier Academic Press, San Diego, 2007, pp. 137–196.

[14] M. Štěpnička and R. Valášek, Fuzzy transforms for functions with two variables, in: J. Ramik, V. Novák (eds.), *Methods for Decision Support in Environment with Uncertainty-Applications in Economics, Business and Engineering*, IRAFM, Ostrava, Czech Republic, 2003, pp. 96–102.

[15] M. Štěpnička and R. Valášek, Numerical solution of partial differential equations with help of fuzzy transform, in: *Proc. Int. Conf. on FUZZ – IEEE*, Reno, 2005, pp. 1104–1109.