Dariusz Jacek Jakóbczak
Katedra Podstaw Informatyki i Zarządzania
Wydział Elektroniki i Informatyki
Politechnika Koszalińska

Data Forecasting and Extrapolation via Probability Distribution and Nodes Combination

Keywords: information retrieval, data extrapolation, curve interpolation, PNC method, probabilistic modeling, forecasting

1. Introduction

Information retrieval and data forecasting are still the opened questions not only in mathematics and computer science. For example the process of planning can meet such a problem: what is the next value that is out of our knowledge, for example any wanted value by tomorrow. This planning may deal with buying or selling, with anticipating costs, expenses or with foreseeing any important value. The key questions in planning and scheduling, also in decision making and knowledge representation [1] are dealing with appropriate information modeling and forecasting. Two-dimensional data can be regarded as points on the curve. Classical polynomial interpolations and extrapolations (Lagrange, Newton, Hermite) are useless for data forecasting, because values that are extrapolated (for example the stock quotations or the market prices) represent continuous or discrete data and they do not preserve a shape of the polynomial. This paper is dealing with data forecasting by using the method of Probabilistic Nodes Combination (PNC) and extrapolation as the extension of interpolation. The values which are retrieved, represented by curve points, consist of information which allows us to extrapolate and to forecast some data for example before making a decision [2].

If the probabilities of possible actions are known, then some criteria are to be applied: Laplace, Bayes, Wald, Hurwicz, Savage, Hodge-Lehmann [3] and others [4]. But this paper considers information retrieval and data forecasting based only on 2D nodes. Proposed method of Probabilistic Nodes Combination (PNC) is used in data reconstruction and forecasting. PNC method uses two-dimensional data for knowledge representation [5] and computational foundations [6]. Also medicine [7], industry and manufacturing are looking for the methods connected with geometry of the curves [8]. So suitable data representation and precise reconstruction or extrapolation [9] of the curve is a key factor in many applications of artificial intelligence: forecasting, planning, scheduling and decision making.

The author wants to approach a problem of curve interpolation [10-12] and data forecasting by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points and outside of this set. The functions that are used in computations represent whole family of elementary functions together with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. Nowadays methods apply mainly polynomial functions, for example Bernstein polynomials in Bezier curves, splines and NURBS [13]. But Bezier curves do not represent the interpolation method and cannot be used for extrapolation. Numerical methods for data reconstruction are based on polynomial or trigonometric functions, for example Lagrange, Newton, Aitken and Hermite methods. These methods have some weak sides [14] and are not sufficient for curve interpolation and extrapolation in the situations when the curve cannot be build by polynomials or trigonometric functions. Proposed 2D point retrieval and forecasting method is the functional modeling via any elementary functions and it helps us to fit the curve.

Author presents novel Probabilistic Nodes Combination (PNC) method of curve interpolation-extrapolation and takes up PNC method of two-dimensional curve modeling via the examples using the family of Hurwitz-Radon matrices (MHR method) [15], but not only (other nodes combinations). The method of PNC requires minimal assumptions: the only information about a curve is the set of at least two nodes $p_i = (x_i,y_i) \in \boldsymbol{R}^2$, $i = 1,2,\ldots n$. Proposed PNC method is applied in data forecasting and information retrieval via different coefficients: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power. Function for PNC calculations is chosen individually at each modeling and it represents probability distribution function of parameter $\alpha \in [0;1]$ for every point situated between two successive interpolation knots. For more precise modeling knots ought to be settled at key points of the curve, for example local minimum or maximum, highest point of the curve in a particular orientation, convexity changing or curvature extrema.

The goal of this paper is to answer the question: how to build the data model by a set of knots [16] and how to extrapolate the points?

2. Data Simulation and Extrapolation

The method of PNC is computing points between two successive nodes of the curve: calculated points are interpolated and parameterized for real number $\alpha \in [0;1]$ in the range of two successive nodes. PNC method uses the combinations of nodes $p_1=(x_1,y_1)$, $p_2=(x_2,y_2),\ldots$, $p_n=(x_n,y_n)$ as $h(p_1,p_2,\ldots,p_m)$ and $m = 1,2,\ldots n$ to interpolate second coordinate y as (2) for first coordinate c in (1):

$$c = \alpha\cdot xi + (1-\alpha)\cdot xi+1, \quad i = 1,2,\ldots n-1, \tag{1}$$

$$y(c) = \gamma \cdot y_i + (1-\gamma) y_{i+1} + \gamma(1-\gamma)\cdot h(p_1, p_2,\ldots, p_m), \tag{2}$$

$\alpha \in [0;1]$, $\gamma = F(\alpha) \in [0;1]$, $F:[0;1]\to[0;1]$, $F(0)=0$, $F(1)=1$ and F is strictly monotonic.

PNC extrapolation requires α outside of [0;1]: $\alpha < 0$ (anticipating points right of last node for $c > x_n$) or $\alpha > 1$ (extrapolating values left of first node for $c < x_1$), $\gamma=F(\alpha)$, $F:P\to\mathbf{R}$, $P \supset [0;1]$, $F(0)=0$, $F(1)=1$. Here are the examples of h computed for MHR method [17]:

$$h(p_1, p_2) = \frac{y_1}{x_1} x_2 + \frac{y_2}{x_2} x_1 \tag{3}$$

or

$$h(p_1, p_2, p_3, p_4) = \frac{1}{x_1^2 + x_3^2}(x_1x_2y_1 + x_2x_3y_3 + x_3x_4y_1 - x_1x_4y_3) + $$

$$+ \frac{1}{x_2^2 + x_4^2}(x_1x_2y_2 + x_1x_4y_4 + x_3x_4y_2 - x_2x_3y_4) .$$

Three other examples of nodes combinations:

$h(p_1, p_2) = \frac{y_1x_2}{x_1y_2} + \frac{y_2x_1}{x_2y_1}$ or $h(p_1, p_2) = x_1x_2 + y_1y_2$ or the simplest $h(p_1, p_2,\ldots, p_m) = 0$.

Nodes combination is chosen individually for each data and it depends on the type of information modeling. Formulas (1)-(2) represent curve parameterization as $\alpha \in P$:

$$x(\alpha) = \alpha\cdot x_i + (1-\alpha)\cdot x_{i+1}$$

and

$$y(\alpha) = F(\alpha)\cdot y_i + (1-F(\alpha)) y_{i+1} + F(\alpha)(1-F(\alpha))\cdot h(p_1, p_2,\ldots, p_m),$$

$$y(\alpha) = F(\alpha)\cdot (y_i - y_{i+1} + (1-F(\alpha))\cdot h(p_1, p_2,\ldots, p_m)) + y_{i+1} .$$

Proposed parameterization gives us the infinite number of possibilities for calculations (determined by choice of F and h) as there is the infinite number of data for reconstruction and forecasting. Nodes combination is the individual feature of each modeled data. Coefficient $\gamma = F(\alpha)$ and nodes combination h are key factors in PNC interpolation and forecasting.

2.1. Extended distribution functions in PNC forecasting

Points settled between the nodes are computed using PNC method. Each real number $c \in [a;b]$ is calculated by a convex combination $c = \alpha \cdot a + (1 - \alpha) \cdot b$ for $\alpha = \frac{b-c}{b-a} \in [0;1]$. Key question is dealing with coefficient γ in (2). The simplest way of PNC calculation means $h = 0$ and $\gamma = \alpha$ (basic probability distribution). Then PNC represents a linear interpolation and extrapolation. MHR method [18] is not a linear interpolation. MHR [19] is the example of PNC modeling. Each interpolation requires specific distribution of parameter α and γ (1)-(2) depends on parameter $\alpha \in [0;1]$:

$$\gamma = F(\alpha),\ F:[0;1]\rightarrow[0;1],\ F(0) = 0,\ F(1) = 1$$

and F is strictly monotonic. Coefficient γ is calculated using different functions (polynomials, power functions, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan or arc cot, also inverse functions) and choice of function is connected with initial requirements and data specifications. Different values of coefficient γ are connected with applied functions $F(\alpha)$. These functions $\gamma = F(\alpha)$ represent the examples of probability distribution functions for random variable $\alpha \in [0;1]$ and real number $s > 0$: $\gamma=\alpha^s$, $\gamma=sin(\alpha^s\cdot\pi/2)$, $\gamma=sin^s(\alpha\cdot\pi/2)$, $\gamma=1-cos(\alpha^s\cdot\pi/2)$, $\gamma=1-cos^s(\alpha\cdot\pi/2)$, $\gamma=tan(\alpha^s\cdot\pi/4)$, $\gamma=tan^s(\alpha\cdot\pi/4)$, $\gamma=log_2(\alpha^s+1)$, $\gamma=log_2^s(\alpha+1)$, $\gamma=(2^\alpha-1)^s$, $\gamma=2/\pi\cdot arcsin(\alpha^s)$, $\gamma=(2/\pi\cdot arcsin\alpha)^s$, $\gamma=1-2/\pi\cdot arccos(\alpha^s)$, $\gamma=1-(2/\pi\cdot arccos\alpha)^s$, $\gamma=4/\pi\cdot arctan(\alpha^s)$, $\gamma=(4/\pi\cdot arctan\alpha)^s$, $\gamma=ctg(\pi/2-\alpha^s\cdot\pi/4)$, $\gamma=ctg^s(\pi/2-\alpha\cdot\pi/4)$, $\gamma=2-4/\pi\cdot arcctg(\alpha^s)$, $\gamma=(2-4/\pi\cdot arcctg\alpha)^s$, $\gamma=\beta\cdot\alpha^2+(1-\beta)\cdot\alpha$, $\gamma=\beta\cdot\alpha^4+(1-\beta)\cdot\alpha,\ldots$, $\gamma=\beta\cdot\alpha^{2k}+(1-\beta)\cdot\alpha$ for $\beta\in[0;1]$ and $k\in \mathbf{N}$ or $\gamma = 1-(1-\alpha)\cdot s^\alpha$.

Functions above, used in γ calculations, are strictly monotonic for random variable $\alpha\in[0;1]$ as $\gamma = F(\alpha)$ is probability distribution function. There is one important probability distribution in mathematics: beta distribution where for example $\gamma=3\alpha^2-2\alpha^3$, $\gamma=4\alpha^3-3\alpha^4$ or $\gamma=2\alpha-\alpha^2$. Also inverse functions F^{-1} are appropriate for γ calculations. Choice of function and value s depends on data specifications and individual requirements during data interpolation.

Extrapolation demands that α is out of range [0;1], for example $\alpha\in(1;2]$ or $\alpha\in[-1;0)$, with $\gamma = F(\alpha)$ as probability distribution function and then F is called extended distribution function in the case of extrapolation. Some of these functions γ are useless for data forecasting because they do not exist ($\gamma=\alpha^{½}$, $\gamma=\alpha^{¼}$) if $\alpha < 0$ in (1). Then it is possible to change parameter $\alpha < 0$ into corresponding $\alpha > 1$ and formulas (1)-(2) turn to equivalent equations:

$$c = \alpha\cdot xi+1 + (1-\alpha)\cdot xi,\quad i = 1,2,\ldots n-1, \tag{4}$$

$$y(c) = \gamma\cdot y_{i+1} + (1-\gamma)y_i + \gamma(1-\gamma)\cdot h(p_1, p_2,\ldots, p_m)\,. \tag{5}$$

PNC forecasting for $\alpha < 0$ or $\alpha > 1$ uses function F as extended distribution function for the arguments from $P \supset [0;1]$, $\gamma = F(\alpha)$, $F{:}P \to \mathbf{R}$, F(0)=0, F(1)=1 and F has to be strictly monotonic only for $\alpha \in [0;1]$. Data simulation and modeling for $\alpha < 0$ or $\alpha > 1$ is done using the same function $\gamma = F(\alpha)$ that is earlier defined for $\alpha \in [0;1]$.

3. PNC Extrapolation and Data Trends

Unknown data are modeled (interpolated or extrapolated) by the choice of nodes, determining specific nodes combination and probabilistic distribution function to show trend of values: increasing, decreasing or stable. Less complicated models take $h(p_1,p_2,\ldots,p_m) = 0$ and then the formula of interpolation (2) looks as follows:

$$y(c) = \gamma \cdot y_i + (1-\gamma)y_{i+1}.$$

It is linear interpolation for basic probability distribution ($\gamma = \alpha$).

<u>Example 1</u>

Nodes are (1;3), (3;1), (5;3) and (7;3), $h = 0$, extended distribution $\gamma = \alpha^2$, extrapolation is computed with (4)-(5) for $\alpha > 1$:

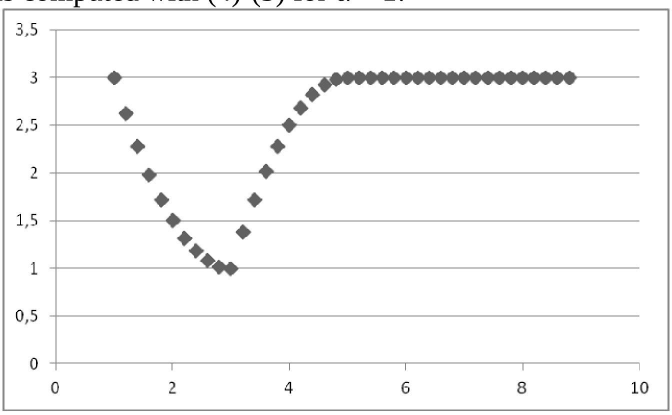

Fig. 1. PNC for 9 interpolated points between nodes and 9 extrapolated points.

Anticipated points (stable trend): (7.2;3), (7.4;3), (7.6;3), (7.8;3), (8;3), (8.2;3), (8.4;3), (8.6;3), (8.8;3) for α = 1.1, 1.2, …, 1.9.

Example 2

Nodes (1;3), (3;1), (5;3) and (7;2), h = 0, extended distribution $\gamma=F(\alpha)=\alpha^2$. Forecasting is computed as (4)-(5) with $\alpha > 1$:

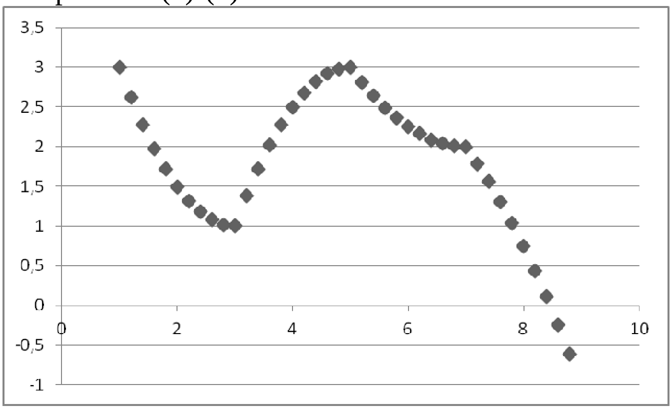

Fig. 2. PNC with 9 interpolated points between nodes and 9 extrapolated points.

Extrapolated points (decreasing trend): (7.2;1.79), (7.4;1.56), (7.6;1.31), (7.8;1.04), (8;0.75), (8.2;0.44), (8.4;0.11), (8.6;-0.24), (8.8;-0.61) for α = 1.1, 1.2, …, 1.9.

Example 3

Nodes (1;3), (3;1), (5;3) and (7;4), h = 0, extended distribution $\gamma=F(\alpha)=\alpha^3$:

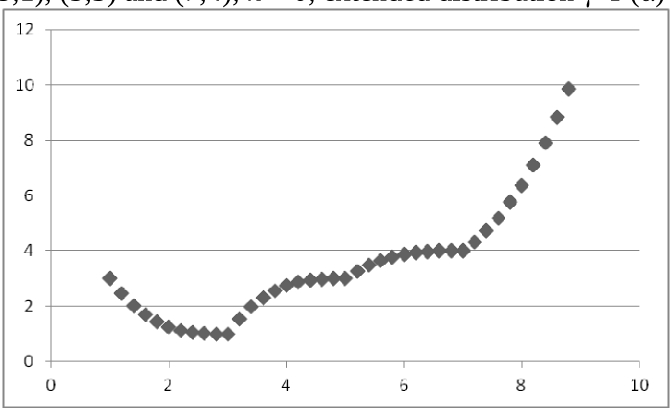

Fig. 3. PNC for 9 interpolated points between nodes and 9 extrapolated points.

Forecast (increasing trend): (7.2;4.331), (7.4;4.728), (7.6;5.197), (7.8;5.744), (8;6.375), (8.2;7.096), (8.4;7.913), (8.6;8.832), (8.8;9.859) for α = 1.1, 1.2, …, 1.9.

These three examples 1-3 (Fig.1-3) with nodes combination h = 0 differ at fourth node and extended probability distribution functions $\gamma=F(\alpha)$. Much more possibilities of modeling are connected with a choice of nodes combination

$h(p_1,p_2,\ldots,p_m)$. MHR method [20] uses the combination (3) with good features connected with orthogonal rows and columns at Hurwitz-Radon family of matrices [21-22]:

$$h(p_i, p_{i+1}) = \frac{y_i}{x_i} x_{i+1} + \frac{y_{i+1}}{x_{i+1}} x_i$$

and then (2): $y(c) = \gamma \cdot y_i + (1-\gamma) y_{i+1} + \gamma(1-\gamma) \cdot h(p_i, p_{i+1})$.

Here are two examples 4 and 5 of PNC method with MHR combination (3).

<u>Example 4</u>

Nodes are (1;3), (3;1) and (5;3), extended distribution $\gamma = F(\alpha) = \alpha^2$. Forecasting is computed with (4)-(5) for $\alpha > 1$:

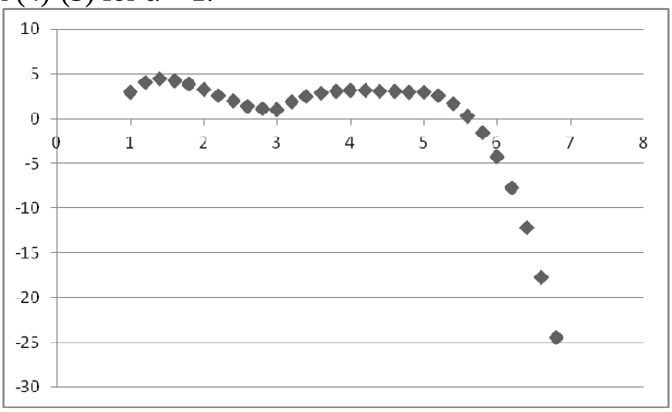

Fig. 4. PNC modeling with 9 interpolated points between nodes and 9 extrapolated points.

Extrapolation (decreasing trend): (5.2;2.539), (5.4;1.684), (5.6;0.338), (5.8;-1.603), (6;-4.25), (6.2;-7.724), (6.4;-12.155), (6.6;-17.68), (6.8;-24.443) for α = 1.1, 1.2, …, 1.9.

Example 5

Nodes (1;3), (3;1) and (5;3), extended distribution $\gamma = F(\alpha) = \alpha^{1.5}$. This forecasting is computed with (4)-(5) for $\alpha > 1$:

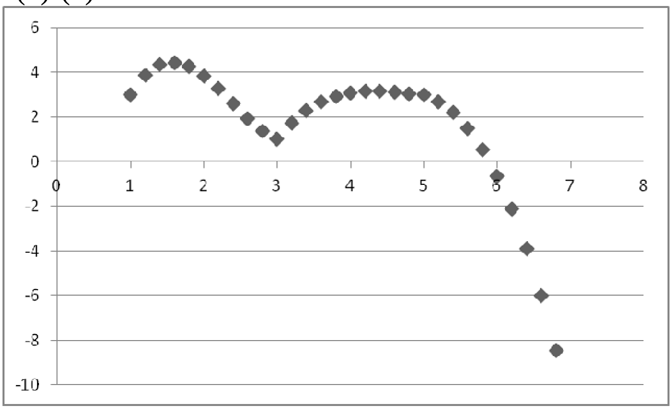

Fig. 5. PNC modeling with 9 interpolated points between nodes and 9 extrapolated points.

Value forecasting (decreasing trend): (5.2;2.693), (5.4;2.196), (5.6;1.487), (5.8;0.543), (6;-0.657), (6.2;-2.136), (6.4;-3.915), (6.6;-6.016), (6.8;-8.461) for α =1.1, 1.2, …, 1.9.

Now let us consider PNC method with other functions F than power functions, $\alpha < 0$ for extrapolation (1)-(2) and nodes combination h=0.

Example 6

Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and extended distribution $F(\alpha)=sin(\alpha\cdot\pi/2)$, h=0:

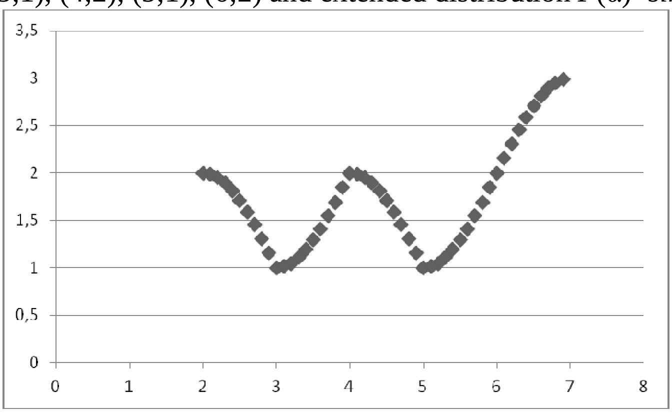

Fig. 6. PNC modeling with 9 interpolated points between nodes and 9 extrapolated points.

Extrapolation points (increasing trend): (6.1;2.156), (6.2;2.309), (6.3;2.454), (6.4;2.588), (6.5;2.707), (6.6;2.809), (6.7;2.891), (6.8;2.951), (6.9;2.988) for α = -0.1, -0.2, …, -0.9.

Example 7

Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and extended distribution $\gamma=F(\alpha)=sin^3(\alpha\cdot\pi/2)$, $h = 0$:

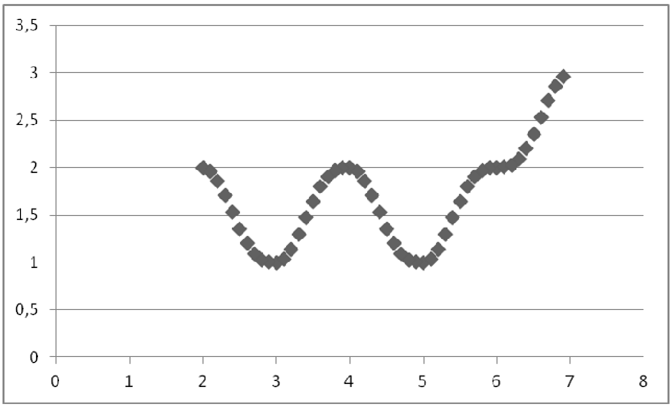

Fig. 7. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node.

Forecast points (increasing trend): (6.1;2.004), (6.2;2.03), (6.3;2.094), (6.4;2.203), (6.5;2.354), (6.6;2.53), (6.7;2.707), (6.8;2.86), (6.9;2.964) for α = -0.1, -0.2, …, -0.9.

These two examples 6 and 7 (Fig.6-7) with nodes combination h=0 and the same set of nodes differ only at extended probability distribution functions $\gamma = F(\alpha)$. Fig.8 is the example of nodes combination h as (3) in MHR method.

Example 8

Nodes (2;2), (3;1), (4;1), (5;1), (6;2) and extended distribution function $\gamma=F(\alpha)=2^{\alpha}-1$:

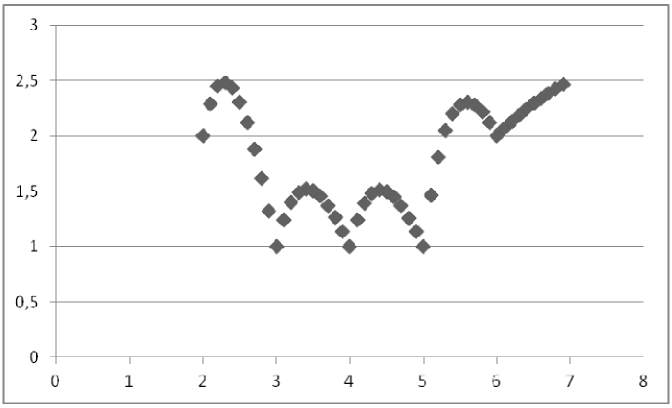

Fig. 8. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node.

Forecast points (increasing trend): (6.1;2.067), (6.2;2.129), (6.3;2.188), (6.4;2.242), (6.5;2.293), (6.6;2.34), (6.7;2.384), (6.8;2.426), (6.9;2.464) for α = -0.1, -0.2, …, -0.9.

Examples that are calculated above have one function $\gamma = F(\alpha)$ and one combination h for all ranges between nodes. But it is possible to create a model with functions $\gamma_i = F_i(\alpha)$ and combinations h_i individually for every range of nodes $(p_i;p_{i+1})$. Then it enables very precise modeling of data between each successive pair of nodes. Each data point is interpolated or extrapolated by PNC via three factors: the set of nodes, probability distribution function $\gamma = F(\alpha)$ and nodes combination h. These three factors are chosen individually for each data, therefore this information about modeled points seems to be enough for specific PNC data retrieval and forecasting. Function γ is selected via the analysis of known points before extrapolation, we may assume $h = 0$ at the beginning and after some time exchange h by more adequate.

These eight examples illustrate the forecasting of some values in planning process, for example anticipation of some costs or expenses and foreseeing the prices or other significant data in the process of planning.

4. Conclusions

The paper is dealing with information retrieval and data forecasting. The method of Probabilistic Nodes Combination (PNC) enables interpolation and extrapolation of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. Function for γ calculations is chosen individually at each case and it is treated as probability distribution function: γ depends on initial requirements and data specifications. PNC method leads to point extrapolation and interpolation via discrete set of fixed knots. Main features of PNC method are: PNC method develops a linear interpolation and extrapolation into other functions as probability distribution functions; PNC is a generalization of MHR method via different nodes combinations; nodes combination and coefficient γ are crucial in the process of data probabilistic retrieval and forecasting. Future works are going to precise the choice and features of nodes combinations and coefficient γ, also to implementation of PNC in handwriting and signature recognition.

References

1. Brachman, R.J., Levesque, H.J.: Knowledge Representation and Reasoning. Morgan Kaufman, San Francisco (2004)
2. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: Reasoning About Knowledge. MIT Press (1995)
3. Straffin, P.D.: Game Theory and Strategy. Mathematical Association of America, Washington, D.C. (1993)
4. Watson, J.: Strategy – An Introduction to Game Theory. University of California, San Diego (2002)
5. Markman, A.B.: Knowledge Representation. Lawrence Erlbaum Associates (1998)
6. Sowa, J.F.: Knowledge Representation: Logical, Philosophical and Computational Foundations. Brooks/Cole, New York (2000)
7. Soussen, C., Mohammad-Djafari, A.: Polygonal and Polyhedral Contour Reconstruction in Computed Tomography. IEEE Transactions on Image Processing 11(13), 1507-1523 (2004)
8. Tang, K.: Geometric Optimization Algorithms in Manufacturing. Computer – Aided Design & Applications 2(6), 747-757 (2005)
9. Kozera, R.: Curve Modeling via Interpolation Based on Multidimensional Reduced Data. Silesian University of Technology Press, Gliwice (2004)
10. Collins II, G.W.: Fundamental Numerical Methods and Data Analysis. Case Western Reserve University (2003)
11. Chapra, S.C.: Applied Numerical Methods. McGraw-Hill (2012)
12. Ralston, A., Rabinowitz, P.: A First Course in Numerical Analysis – Second Edition. Dover Publications, New York (2001)
13. Schumaker, L.L.: Spline Functions: Basic Theory. Cambridge Mathematical Library (2007)
14. Dahlquist, G., Bjoerck, A.: Numerical Methods. Prentice Hall, New York (1974)

15. Jakóbczak, D.: 2D and 3D Image Modeling Using Hurwitz-Radon Matrices. Polish Journal of Environmental Studies 4A(16), 104-107 (2007)
16. Jakóbczak, D.: Shape Representation and Shape Coefficients via Method of Hurwitz-Radon Matrices. Lecture Notes in Computer Science 6374 (Computer Vision and Graphics: Proc. ICCVG 2010, Part I), Springer-Verlag Berlin Heidelberg, 411-419 (2010)
17. Jakóbczak, D.: Curve Interpolation Using Hurwitz-Radon Matrices. Polish Journal of Environmental Studies 3B(18), 126-130 (2009)
18. Jakóbczak, D.: Application of Hurwitz-Radon Matrices in Shape Representation. In: Banaszak, Z., Świć, A. (eds.) Applied Computer Science: Modelling of Production Processes 1(6), pp. 63-74. Lublin University of Technology Press, Lublin (2010)
19. Jakóbczak, D.: Object Modeling Using Method of Hurwitz-Radon Matrices of Rank k. In: Wolski, W., Borawski, M. (eds.) Computer Graphics: Selected Issues, pp. 79-90. University of Szczecin Press, Szczecin (2010)
20. Jakóbczak, D.: Implementation of Hurwitz-Radon Matrices in Shape Representation. In: Choraś, R.S. (ed.) Advances in Intelligent and Soft Computing 84, Image Processing and Communications: Challenges 2, pp. 39-50. Springer-Verlag, Berlin Heidelberg (2010)
21. Jakóbczak, D.: Object Recognition via Contour Points Reconstruction Using Hurwitz-Radon matrices. IGI Global books „Image Processing: Concepts, Methodologies, Tools, and Applications", Hershey PA, USA, 998-1018 (2013)
22. Jakóbczak, D.: Curve Parameterization and Curvature via Method of Hurwitz-Radon Matrices. Image Processing & Communications- An International Journal 1-2(16), 49-56 (2011)

Abstract

Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D data interpolation and extrapolation. Nodes are treated as characteristic points of information retrieval and data forecasting. PNC modeling via nodes combination and parameter γ as probability distribution function enables 2D point extrapolation and interpolation. Two-dimensional information is modeled via nodes combination and some functions as continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Extrapolated values are used as the support in data forecasting.

Streszczenie

Autorska metoda Probabilistycznej Kombinacji Węzłów- Probabilistic Nodes Combination (PNC) jest wykorzystywana do interpolacji i ekstrapolacji dwuwymiarowych danych. Węzły traktowane są jako punkty charakterystyczne informacji, która ma być odtwarzana lub przewidywana. Dwuwymiarowe dane są interpolowane lub ekstrapolowane z wykorzystaniem różnych funkcji rozkładu prawdopodobieństwa: potęgowych, wielomianowych, wykładniczych, logarytmicznych, trygonometrycznych, cyklometrycznych. W pracy pokazano propozycję metody ekstrapolowania danych jako pomoc w przewidywaniu trendu dla nieznanych wartości.