

# On the equivalence of the forecast value construction in the “pyramidal” extrapolation method and cubic forecast

<https://doi.org/10.31713/MCIT.2020.15>

Yuriy Turbal

department of computer science and applied mathematic  
National university of water and environmental engineering  
Rivne, Ukraine  
e-mail: y.v.turbal@nuwm.edu.ua

Mariana Turbal

department of computer science and applied mathematic  
National university of water and environmental engineering  
Rivne, Ukraine  
e-mail: y.v.turbal@nuwm.edu.ua

Driwi Abd Alkaleg

department of computer science and applied mathematic  
National university of water and environmental engineering  
Rivne, Ukraine  
e-mail: a.a.drivi@nuwm.edu.ua

Al Shukri Sumayya Ali

department of computer science and applied mathematic  
National university of water and environmental engineering  
Rivne, Ukraine  
e-mail: s.a.alshukri@asmarya.edu.ly

**Abstract**—The paper is devoted to the new method of short time series extrapolation, which was developed in recent years and was called “pyramidal”. This method can be used for predicting the deterministic component of an arbitrary economic, environmental and geophysical parameters. Numerical results show significant advantages of the proposed method in comparison with approaches to extrapolate, based on the use of polynomials, including Newton’s extrapolation . In this paper we study the procedure for constructing the forecast value in accordance with the pyramidal method and show that it is equivalent to the cubic forecast.

**Keywords**— extrapolation; forecast; divided differences; interpolation, time series recognition problems

## I. INTRODUCTION

The most common methods of short time-series forecasting are the extrapolation methods [6]. Among the known methods of extrapolation we can distinguish methods based on interpolation polynomials – the most commonly used Newton - Gregory backward difference formula [3], generalized polynomial interpolation based on the systems of Chebyshev functions – polynomials, exponential, trigonometric functions [2], methods based on the special functions and trend analyze [4]-[9], methods based on spline – cubic, B-spline, Bezier curves [10] and methods, based on the statistical approaches [6],[12]-[13].

There are a number of problems dealing with the study of small time series [8]. For example, in the absence of additional information is often impossible to conclude that process is determinate or indeterminate, which significantly effects on the model construction. In determinate case is obvious that for any set of points there are many curves that pass through them or something to bring, and it is difficult to argue that one curve (model) is precisely the law that comprehensively describes the

phenomenon and will effectively predict its behavior in the future. The hypothesis, which corresponds to the observation point, is effective if it is appropriate to predict with reasonable accuracy the process starting from any step of observations. In practice, the finding an effective hypothesis is impossible. The main idea of this article is to develop a technique that allows you to build the optimal predict point without direct use of any specific classes of extrapolation function. In the paper [1] was proposed a new method of short time series extrapolation, which can be used for predicting of economic, environmental and geophysical parameters. This method is based on the property of the rows of modified finite differences that the best cubic approximation is in the range of convexity. Numerical results show significant advantages of the proposed method in comparison with approaches to extrapolate, based on the use of polynomials, including Newton’s extrapolation. This method is comparatively new and requires deep in-depth research and data series validation. In this paper we study the procedure for constructing the forecast value in accordance with the pyramidal method and show that it is equivalent to the cubic forecast.

## II. PIRAMIDAL METHOD OF DATA EXTRAPOLATION

Pyramidal method of data extrapolation was proposed in [1] . The main feature of this method is construction a special divided differences and finding their order, for which a better predicted value in a certain sense can be found. Then the value of the original function at the point located outside the interpolation interval is based on the predictive value for the divided differences using a special computational procedure.

Let  $f(x)$  be a continuous function of one independent variable defined over a closed interval  $[a,b]$  having continuous derivatives and let values  $f_1, f_2, \dots, f_n$  of this function are

known at the points  $x_1, x_2, \dots, x_n$  respectively. It is needed to assess the value of this function at the point  $x > x_n$ . Consider the finite differences modified as follows:

$$\Delta^j f_i = \frac{\Delta^{j-1} f_{i+1} - \Delta^{j-1} f_i}{r_i^j - l_i^j}, \quad (1)$$

$$\text{where } r_i^j = \begin{cases} x^c_{i+\frac{j}{2}}, j=2k, \\ x_{i+\left[\frac{j}{2}\right]+1}, j=2k+1, \end{cases} \quad l_i^j = \begin{cases} x^c_{i+\frac{j}{2}-1}, j=2k, \\ x_{i+\left[\frac{j}{2}\right]}, j=2k+1, \end{cases}$$

$$j = \overline{1, n-1}, i = \overline{1, n-j}.$$

Define the middle points  $x_i^c = (x_i + x_{i+1})/2$ . Let  $f_i^c, i = \overline{1, n-1}$  be the functions values in intermediate points  $x_i^c, i = \overline{1, n-1}$ . Then we can define the finite differences in the intermediate points as follows:

$$\Delta^j f_i^c = \frac{\Delta^{j-1} f_{i+1}^c - \Delta^{j-1} f_i^c}{\hat{r}_i^j - \hat{l}_i^j}, \quad (2)$$

where

$$\hat{r}_i^j = \begin{cases} x_{i+\frac{j}{2}+1}, j=2k, \\ x^c_{i+\left[\frac{j}{2}\right]+1}, j=2k+1, \end{cases} \quad \hat{l}_i^j = \begin{cases} x_{i+\frac{j}{2}}, j=2k, \\ x^c_{i+\left[\frac{j}{2}\right]}, j=2k+1, \end{cases}$$

$$j = \overline{1, n-2}, i = \overline{1, n-j}.$$

Consider the values of the form:

$$\tilde{\Delta}^i f_{n-i}^c = 2 \frac{(\Delta^{i-2} f_{n-i+2} - \Delta^{i-2} f_{n-i+1}^c)}{(r_i - c_i)(r_i - l_i)} - \frac{(\Delta^{i-2} f_{n-i+1}^c - \Delta^{i-2} f_{n-i+1})}{(c_i - l_i)(r_i - l_i)}, \quad (3)$$

where

$$r_i = \begin{cases} x_{n-\frac{i}{2}+1}, i=2k, \\ x^c_{n-\left[\frac{i}{2}\right]}, i=2k+1, \end{cases}$$

$$c_i = \begin{cases} x^c_{n-\frac{i}{2}}, i=2k, \\ x_{n-\left[\frac{i}{2}\right]}, i=2k+1, \end{cases} \quad l_i = \begin{cases} x_{n-\frac{i}{2}}, i=2k, \\ x^c_{n-\left[\frac{i}{2}\right]-1}, i=2k+1, \end{cases}$$

$$ll_i = \begin{cases} x^c_{n-\frac{i}{2}-1}, i=2k, \\ x_{n-\left[\frac{i}{2}\right]-1}, i=2k+1. \end{cases}$$

$$i = \overline{2, n-1}.$$

$$\text{According to [1] if for any index } i \text{ the condition} \quad \frac{(\Delta^{i-2} f_{n-i+1}^c - \Delta^{i-2} f_{n-i+1})}{c_i - l_i} = \frac{(\Delta^{i-2} f_{n-i+1}^c - \Delta^{i-2} f_{n-i}^c)}{c_i - ll_i} \quad (4)$$

are satisfied and the condition of linearity of slope change velocity for the broken line which nodes at the points  $(ll_i, \Delta^{i-2} f_{n-i}^c)$ ,  $(l_i, \Delta^{i-2} f_{n-i+1}^c)$ ,  $(c_i, \Delta^{i-2} f_{n-i+1}^c)$ ,  $(r_i, \Delta^{i-2} f_{n-i+2}^c)$ ,  $(rr_i, \Delta^{i-2} f_{n-i+2}^c)$  it is satisfied equality:

$$\tilde{\Delta}^i f_{n-i}^c = \Delta^i f_{n-i}^c. \quad (5)$$

Then a procedure of finding of the function unknown value at the point  $x_n^c$  can be built as follows:

$$\Delta^{j-1} f_{n-j+1}^c = \Delta^{j-1} f_{n-j}^c + \Delta^j f_{n-j}^c (\hat{r}_{n-j}^j - \hat{l}_{n-j}^j), \quad j = \overline{i, 1}. \quad (6)$$

Consequently, the following algorithm of extrapolation was proposed:

1. Construct the table of finite difference according to (1).
2. Index  $i^*$  is finding according to the formula

$$i^* = \operatorname{argmin}_i \left| \frac{(\Delta^{i-2} f_{n-i+1}^c - \Delta^{i-2} f_{n-i+1})}{c_i - l_i} - \frac{(\Delta^{i-2} f_{n-i+1}^c - \Delta^{i-2} f_{n-i}^c)}{c_i - ll_i} \right|$$

Calculate the value  $\tilde{\Delta}^{i^*} f_{n-i^*}^c$ .

3. The predictive function value formula:

$$\Delta^{j-1} f_{n-j+1}^c = \Delta^{j-1} f_{n-j}^c + \Delta^j f_{n-j}^c (\hat{r}_{n-j}^j - \hat{l}_{n-j}^j), \quad \text{where } \tilde{\Delta}^{i^*} f_{n-i^*}^c = \Delta^{i^*} f_{n-i^*}^c, j = \overline{i^*, 1}.$$

Without loss of generality we can consider uniform grid,  $x_k - x_{k-1} = 1$ . In this case finite differences (3) can be easy to calculate. The illustration is given on the Table 1. In this transposed table, the values  $\Delta^k f_l, \Delta^k f_{l+1}, \Delta^k f_{l+2}, \Delta^k f_{l+3}$  are known,  $\Delta^k f_{l+4}$  is unknown. Other values recorded in selected cells are also unknown. According to (3) we can find  $4\Delta^k f_{l+3} - 8\Delta^k f_{l+2} + 4\Delta^k f_{l+1}$  and it is easy to find another



$$A_2 = \begin{pmatrix} x^3 & z1 & x & 1 \\ (x+d)^3 & z2 & x+d & 1 \\ (x+2d)^3 & z3 & x+2d & 1 \\ (x+3d)^3 & z4 & x+3d & 1 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} x^3 & x^2 & z1 & 1 \\ (x+\Delta)^3 & (x+\Delta)^2 & z2 & 1 \\ (x+2\Delta)^3 & (x+2\Delta)^2 & z3 & 1 \\ (x+3\Delta)^3 & (x+3\Delta)^2 & z4 & 1 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} x^3 & x^2 & x & z1 \\ (x+\Delta)^3 & (x+\Delta)^2 & x+\Delta & z2 \\ (x+2\Delta)^3 & (x+2\Delta)^2 & x+2\Delta & z3 \\ (x+3\Delta)^3 & (x+3\Delta)^2 & x+3\Delta & z4 \end{pmatrix}.$$

Obviously, finding the appropriate determinants is cumbersome task. We can use symbolic operation in Mathcad environment. Thus

$$\det A = 12d^6,$$

$$c1 = \det A_1 / \det A =$$

$$-(2*d^3*z1 - 6*d^3*z2 + 6*d^3*z3 - 2*d^3*z4) / (12*d^6),$$

$$d1 = \det A_2 / \det A =$$

$$(30*d^4*z2 - 12*d^4*z1 - 24*d^4*z3 + 6*d^4*z4 -$$

$$6*d^3*x*z1 + 18*d^3*x*z2 - 18*d^3*x*z3 + 6*d^3*x*z4) / (12*d^6),$$

$$e1 = \det A_3 / \det A =$$

$$(36*d^5*z2 - 22*d^5*z1 - 18*d^5*z3 + 4*d^5*z4 -$$

$$6*d^3*x^2*z1 + 18*d^3*x^2*z2 -$$

$$18*d^3*x^2*z3 + 6*d^3*x^2*z4 - 24*d^4*x*z1 + 60*d^4*x*z2 -$$

$$48*d^4*x*z3 + 12*d^4*x*z4) / (12*d^6),$$

$$f1 = \det A_4 / \det A =$$

$$(12*d^6*z1 + 2*d^3*x^3*z1 + 12*d^4*x^2*z1 - 6*d^3*x^3*z2 -$$

$$30*d^4*x^2*z2 + 6*d^3*x^3*z3 + 24*d^4*x^2*z3 -$$

$$2*d^3*x^3*z4 - 6*d^4*x^2*z4 + 22*d^5*x*z1 -$$

$$36*d^5*x*z2 + 18*d^5*x*z3 - 4*d^5*x*z4) / (12*d^6).$$

After substitution all expression into  $c_1(x + 4d)^3 + d_1(x + 4d)^2 + e_1(x + 4d) + f_1$  we get :

$$c_1(x + 4d)^3 + d_1(x + 4d)^2 + e_1(x + 4d) + f_1 = 4z2 - z1 - 6z3 + 4z4.$$

We've get expression (3). Thus, the task of determining the forecast value in the corresponding row of the pyramidal method is equivalent to the cubic forecast based on the last 4 values of the data series.

#### IV CONCLUSIONS

Thus, in this paper we show that the procedure of the predicted value constructing in the pyramidal method is equivalent to the cubic forecast. This result makes it possible to assert that the pyramidal method allows one to construct an optimal cubic forecast in a certain sense in the corresponding row of the pyramid of finite differences. In paper [1] authors

also discuss the cubic forecast, but the corresponding statements are not strict enough. Ultimately, pyramidal method allows to build the forecast based on the optimal degree polynom. The degree of the polinom is chosen in accordance with minimizing the error.

The second interesting result that follows from our article is the method of constructing a cubic forecast of an arbitrary function. Indeed, to construct a cubic forecast by arbitrary 4 points, it is needed to build a fragment of the finite difference table (see Table 1) and carry out simple calculations in accordance with the procedure described in Section 2. In this case, there is no need to solve a system of linear algebraic equations to find the corresponding coefficients.

#### REFERENCES

- [1] Yuriy Turbal, Andriy Bomba, Anastasiia Sokh, Olena Radoveniuk, Mariana Turbal, "Pyramidal method of small time series extrapolation," International journal of computing science and mathematic, vol.10, N4, 2019, pp.122-130
- [2] Bomba, A.Ya., Turbal, Yu.V., 'Data analysis method and problems of identification of trajectories of solitary waves', *Journal of Automation and Information Sciences*, Vol.5, 2015, pp. 34-43.
- [3] Amstrong, J. S. "Extrapolation for Time-Series and Cross-Sectional Data", Retrieved from [http://repository.upenn.edu/marketing\\_papers/148](http://repository.upenn.edu/marketing_papers/148), 2001.
- [4] Adondo D.O., "A local extrapolation methods for hyperbolic conservation laws: the eno and goodman-leveque underlying shemes and sufficient conditions for tvd property", Kansas state university, Manhattan, Kansas, 2008, 85 p.,
- [5] Brezinski C., M. Zaglia M.R., "Extrapolation methods", Applied numerical Mathematics, Vol. 15, Issue 2, 1995, pp. 123-131.
- [6] Chatfield, C., "Time-series Forecasting", Chapman & Hall , 2001, 265 p
- [7] Goura G. S. L., Badcock K. J., Woodgate M. A, and Richards B. E , "Extrapolation Effects on Coupled Computational Fluid Dynamics/Computational Structural Dynamics Simulations", AIAA Journal, Vol. 41, No. 2, 2003, pp. 312-314.
- [8] Hyndman Rob J. and Kostenko A. V. , "Minimum sample size requirements for seasonal forecasting models", *FORESIGHT*, Issue 6, 2007.
- [9] Johanneson P. , "Extrapolation of load histories and spectra", *Fatigue & Fracture of Engineering Materials & Structures* , pp. 209-217.
- [10] Kostinsky, A.S. , "On the principles of a spline extrapolation concerning geophysical data", Reports of the National Academy of Sciences of Ukraine , Vol.2, 2014, pp.111-117.
- [11] Kim W-G, Park J-Y, Jang J.at al ., "Methodologies for long-term creep life extrapolation of Gr. 91 Steel", *The 8-th Pacific Rim International Congress on Advanced Materials and Processing Proceedings* .
- [12] Levi J. C. And Barody I. I., "A Comparative Study of Extrapolation Methods for Creep Data at Small Strains", London : Her Majesty's Stationery Office, 1069, 29 p.
- [13] Zhan, Z., Fu, Y., Yang, R., Xi, Z. et al. (2012), "A Bayesian Inference based Model Interpolation and Extrapolation," *SAE Int. J. Mater. Manf.* v.5(2), pp.357-364, doi:10.4271/2012-01-0223
- [14] Yuriy Turbal, Andriy Bomba, Anastasiia Sokh, Olena Radoveniuk, Mariana Turbal, "Spatial generalization of the pyramidal data etrapolation," Bulletin of Taras Shevchenko National University of Kyiv , Series Physics & Mathematics , vol.1,2017, pp. 94-99.
- [15] Sidi A. , "Practacal extrapolation methods. Theory and applications.", *University of Camdridge*, 2003, 543 p.
- [16] Pace C.N., Shaw K.L., "Linear extrapolation method of analyzing solvent denaturation curves", *Texas A&M University*, 2007, 7 p.