The Recursive Future And Past Equation Based On The Ananda-Damayanthi Normalized Similarity Measure Considered To Exhaustion {File Closing Version+2}

ISSN 1751-3030

Author:
Ramesh Chandra Bagadi

Affiliation 1:

Data Scientist
International School Of Engineering (INSOFE)
2nd Floor, Jyothi Imperial, Vamsiram Builders, Janardana Hills, Above South India Shopping Mall, Old Mumbai Highway, Gachibowli, Hyderabad, TelanganaState, 500032, India.

Email: ramesh.bagadi@insofe.edu.in
Tel:+91 9440032711

Abstract

In this research investigation, the author has presented a Recursive Past Equation and a Recursive Future Equation based on the Ananda-Damayanthi Normalized Similarity Measure considered to Exhaustion [1] (please see the addendum of [1] as well).

The Ananda-Damayanthi Normalized Similarity Measure

Considering any two Real Numbers, their Ananda-Damayanthi Similarity Measure is given by the Smaller of the Two Numbers. The Ananda-Damayanthi Normalized Similarity Measure is given by the Ratio of the Smaller to The Larger [1].

The Recursive Future Equation

Given a Time Series $Y = \{y_1, y_2, y_3, \ldots\ldots, y_{n-1}, y_n\}$

we can find y_{n+1} using the following Recursive Future Equation

$$y_{n+1} = \frac{\left\{\sum_{k=1}^{n} y_k \left\{\left(\frac{S_{k0}}{L_{k0}}\right)\right\}\right\} + \sum_{k=1}^{n}\sum_{j=1}^{m}\left\{L_{k(j-1)} - S_{k(j-1)}\right\}\left\{\frac{S_{kj}}{L_{kj}}\right\}}{\sqrt{\left\{\sum_{k=1}^{n}\left\{\left(\frac{S_{k0}}{L_{k0}}\right)^2\right\}\right\} + \left\{\sum_{k=1}^{n}\sum_{j=1}^{m}\left\{\frac{S_{kj}}{L_{kj}}\right\}^2\right\}}}$$

where

$S_{kj} = Smaller\ \ of\ \left(L_{k(j-1)} - S_{k(j-1)}\right)\ and\ \ y_{n+1}$ and $L_{kj} = Larger\ \ of\ \left(L_{k(j-1)} - S_{k(j-1)}\right)\ and\ \ y_{n+1}$

(This will be detailed in the next section)

where $j = 1\ to\ m$ is a Number which makes the Difference Residual $\left(L_{k(j-1)} - S_{k(j-1)}\right)$ tend to Zero.

From the above Recursive Equation, we can solve for y_{n+1}.

Proof:

We consider y_k and find the Ananda-Damayanthi Similarity [1] between y_k and y_{n+1} which we refer as $\left\{\frac{S_k}{L_k}\right\} = \left\{\frac{Smaller\ \ of\ \ y_k\ \ and\ \ y_{n+1}}{Larger\ \ of\ \ y_k\ \ and\ \ y_{n+1}}\right\}$. We now consider the lack of similarity part, i.e., $(L_k - S_k)$ and again find the Similarity between y_k and $(L_k - S_k)$ (this is the Difference Residual of First Order) which (the aforementioned Similarity) we refer to as $\left\{\frac{S_{k1}}{L_{k1}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_k - S_k\right)\ and\ y_k}{Larger\ \ of\ \left(L_k - S_k\right)\ and\ y_k}\right\}$ wherein the Difference Residual of Second Order is $(L_{k1} - S_{k1})$. And similarly, we find $\left\{\frac{S_{k2}}{L_{k2}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_{k1} - S_{k1}\right)\ and\ y_k}{Larger\ \ of\ \left(L_{k1} - S_{k1}\right)\ and\ y_k}\right\}$, $\left\{\frac{S_{k3}}{L_{k3}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_{k2} - S_{k2}\right)\ and\ y_k}{Larger\ \ of\ \left(L_{k2} - S_{k2}\right)\ and\ y_k}\right\}$, ………., $\left\{\frac{S_{km}}{L_{km}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_{k(m-1)} - S_{k(m-1)}\right)\ and\ y_k}{Larger\ \ of\ \left(L_{k(m-1)} - S_{k(m-1)}\right)\ and\ y_k}\right\}$. Note that we represent the second index by j which goes from 0 to m. We now add them all. Similarly, we consider such terms for $k = 1$ to n and compute such aforementioned quantities and add them all. We now Normalize (L^2 Norm), i.e., divide each of this value by the quantity $\sqrt{\left\{\sum_{k=1}^{n}\left\{\left(\frac{S_{k0}}{L_{k0}}\right)^2\right\}\right\} + \left\{\sum_{k=1}^{n}\sum_{j=1}^{m}\left\{\frac{S_{kj}}{L_{kj}}\right\}^2\right\}}$. We equate this value to

y_{n+1} as the RHS is the Total Normalized Similarity contribution from each element of the Time Series Set $Y=\{y_1,y_2,y_3,\ldots\ldots,y_{n-1},y_n\}$ with respect to y_{n+1}. Note that the Similarity term corresponding to the Difference Residual of Zeroth Order can be represented as $\left\{\frac{S_{k0}}{L_{k0}}\right\}$ which is actually $\left\{\frac{S_k}{L_k}\right\}$ itself.

Defining Error

We define Error in the following fashion:

For the Recursive Future Equation:

Method 1

Given a Time Series $Y=\{y_1,y_2,y_3,\ldots\ldots,y_{n-1},y_n\}$ we consider only $Y=\{y_1,y_2,y_3,\ldots\ldots,y_{n-1}\}$ and use the aforementioned Recursive Future Equation to find the n^{th} term. Say this is ${}^p y_n$ where the p stands for the 'predicted' or 'forecasted' value. Then, the Error is defined by

$$\varepsilon_F=\left(\frac{y_n-{}^p y_n}{y_n}\right)$$

Method 2

Given a Time Series $Y=\{y_1,y_2,y_3,\ldots\ldots,y_{n-1},y_n\}$ we consider it and use the aforementioned Recursive Future Equation to find the $(n+1)^{th}$ term. Say this is ${}^p y_{n+1}$ where the p stands for the 'predicted' or 'forecasted' value. We now consider the Time Series Set $Y=\{y_2,y_3,\ldots\ldots,y_{n-1},y_n,{}^p y_{n+1}\}$ and use the aforementioned Recursive Past Equation to generate the term previous to y_2, i.e., ${}^p y_1$. Then, the Error is defined by

$$\varepsilon_F=\left(\frac{y_1-{}^p y_1}{y_1}\right)$$

The Recursive Past Equation

Given a Time Series $Y=\{y_1,y_2,y_3,\ldots\ldots,y_{n-1},y_n\}$

we can find y_0 using the following Recursive Past Equation

$$y_n = \frac{\left\{\sum_{k=0}^{n-1} y_k \left\{\left(\frac{S_{k0}}{L_{k0}}\right)\right\}\right\} + \sum_{k=0}^{n-1}\sum_{j=1}^{m}\left\{L_{k(j-1)} - S_{k(j-1)}\right\}\left\{\frac{S_{kj}}{L_{kj}}\right\}}{\sqrt{\left\{\sum_{k=0}^{n-1}\left\{\left(\frac{S_{k0}}{L_{k0}}\right)^2\right\}\right\} + \left\{\sum_{k=0}^{n-1}\sum_{j=1}^{m}\left\{\frac{S_{kj}}{L_{kj}}\right\}^2\right\}}}$$

where

$S_{kj} = Smaller\ \ of\ \left(L_{k(j-1)} - S_{k(j-1)}\right)\ and\ \ y_n$ and $L_{kj} = Larger\ \ of\ \left(L_{k(j-1)} - S_{k(j-1)}\right)\ and\ \ y_n$

where $j = 1\ to\ m$ is a Number which makes the Difference Residual $\left(L_{k(j-1)} - S_{k(j-1)}\right)$ tend to Zero.

From the above Recursive Equation, we can solve for y_0.

Proof:

We consider y_k and find the Ananda-Damayanthi Similarity [1] between y_k and y_n which turns out to be $\left\{\frac{S_k}{L_k}\right\} = \left\{\frac{Smaller\ \ of\ \ y_k\ \ and\ y_n}{Larger\ \ of\ \ y_k\ \ and\ y_n}\right\}$. We now consider the lack of similarity part, i.e., $(L_k - S_k)$ and again find the Similarity between y_k and $(L_k - S_k)$ (this is the Difference Residual of First Order) which (the aforementioned Similarity) we refer to as $\left\{\frac{S_{k1}}{L_{k1}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_k - S_k\right)\ and\ y_k}{Larger\ \ of\ \left(L_k - S_k\right)\ and\ y_k}\right\}$ wherein the Difference Residual of Second Order is $(L_{k1} - S_{k1})$. And similarly, we find $\left\{\frac{S_{k2}}{L_{k2}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_{k1} - S_{k1}\right)\ and\ y_k}{Larger\ \ of\ \left(L_{k1} - S_{k1}\right)\ and\ y_k}\right\}$, $\left\{\frac{S_{k3}}{L_{k3}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_{k2} - S_{k2}\right)\ and\ y_k}{Larger\ \ of\ \left(L_{k2} - S_{k2}\right)\ and\ y_k}\right\}$, ..., $\left\{\frac{S_{km}}{L_{km}}\right\} = \left\{\frac{Smaller\ \ of\ \left(L_{k(m-1)} - S_{k(m-1)}\right)\ and\ y_k}{Larger\ \ of\ \left(L_{k(m-1)} - S_{k(m-1)}\right)\ and\ y_k}\right\}$. Note that we represent the second index by j which goes from 0 to m. We now add them all. Similarly, we consider such terms for $k = 0$ to $n-1$ and compute such aforementioned quantities and add them all. We now Normalize (L^2 Norm), i.e., divide each of this value by the quantity $\sqrt{\left\{\sum_{k=0}^{n-1}\left\{\left(\frac{S_{k0}}{L_{k0}}\right)^2\right\}\right\} + \left\{\sum_{k=0}^{n-1}\sum_{j=1}^{m}\left\{\frac{S_{kj}}{L_{kj}}\right\}^2\right\}}$. We equate this value to y_n as the RHS is the Total Normalized Similarity contribution from each element of the Time Series Set $Y = \left\{y_0, y_1, y_2, y_3, \ldots\ldots, y_{n-1}\right\}$ with respect to y_n. Note that the Similarity term corresponding to the Difference Residual of Zeroth Order can be represented as $\left\{\frac{S_{k0}}{L_{k0}}\right\}$ which is actually $\left\{\frac{S_k}{L_k}\right\}$ itself.

Defining Error

We define Error in the following fashion:

For the Recursive Past Equation:

Method 1

Given a Time Series $Y = \{y_1, y_2, y_3, \ldots\ldots, y_{n-1}, y_n\}$ we consider only $Y = \{y_2, y_3, \ldots\ldots, y_{n-1}, y_n\}$ and use the aforementioned Recursive Future Past to find the 1^{st} term. Say this is ${}^{p}y_1$ where the p stands for the 'predicted' or 'forecasted' value. Then, the Error is defined by

$$\varepsilon_P = \left(\frac{y_1 - {}^{p}y_1}{y_1} \right)$$

Method 2

Given a Time Series $Y = \{y_1, y_2, y_3, \ldots\ldots, y_{n-1}, y_n\}$ we consider it and use the aforementioned Recursive Future Equation to find the term previous to y_1. Say this is ${}^{p}y_0$ where the p stands for the 'predicted' or 'forecasted' value. We now consider the Time Series Set $Y = \{{}^{p}y_0, y_1, y_2, y_3, \ldots\ldots, y_{n-1}\}$ and use the aforementioned Recursive Future Equation to generate the term next to y_{n-1}, i.e., ${}^{p}y_n$. Then, the Error is defined by

$$\varepsilon_F = \left(\frac{y_n - {}^{p}y_n}{y_n} \right)$$

Computation Complexity

For the World's fastest Japaneese Super-Computer which can compute 1 Quadrillion Computations per second

we can use the equation $2^{(m+n)} = 10^{15}$ to calculate the Maximum Number of Terms of the Time Series n for which we wish to predict the $(n+1)^{th}$ term and m is the Number Of Difference Residual Terms we wish to consider for each term, to find the n for a given m so that the $(n+1)^{th}$ term is computed in one second.

Furthermore, if we take $m = 8 or 10$ (beyond which the value of the Difference Residuals is near vanishing) and for different amounts of times we can spare for getting the computed answer, the Number of Terms of the Time Series n that we can consider is given as follows:

Serial Number	*Duration Of Computation*	*Number of Terms n To Consider*
1	1 Second	$21.64043 - m$
2	1 Minute	$25.66808 - m$
3	1 Hour	$29.69574 - m$
4	1 Day	$34.2807 - m$
5	1 Week	$37.0886 - m$
6	1 Month (31 Days)	$39.2349 - m$
7	1 Year	$42.79246 - m$

That is, if the Time Series Set were to contain n number of terms (as shown in the table for varying values of m, namely 8 and 10, then the Duration of Computation is tabulated above.

For Forecasting Future Element

We have $2^{(n+mn)}$ number of 6th Order Polynomial Equations of the kind as shown in equation A to solve as these account for all the cases of the Time Series Set Elements being greater or lesser than the future $(n+1)^{th}$ element to be computed, as these equations are being represented by the aforementioned Recursive Future Equation. Only one among them is the correct equation and this can be found by using this thusly computed $(n+1)^{th}$ value and omitting the first element y_1, using the Time Series Set $Y = \{y_2, y_3, \ldots\ldots, y_{n-1}, y_n, y_{n+1}\}$ we predict the element y_1 using the aforementioned Recursive Past Equation. And one of the $2^{(m+mn)}$ number of 6th Order Polynomial Equations of the kind as shown in equation A which gives the best true value of y_1 can be considered as the correct equation and its future element forecast of y_{n+1} as the correct forecast.

For Forecasting Past (to the First) Element

We have $2^{(n+mn)}$ number of 6th Order Polynomial Equations of the kind as shown in equation A to solve as these account for all the cases of the Time Series Set Elements being greater or lesser than the past element y_0 to be computed, as these equations are being represented by the aforementioned Recursive Past Equation. Only one among them is the correct equation and this can be found by using this thusly computed y_0 value and omitting the latest element y_n, using the Time Series Set $Y = \{y_0, y_1, y_2, y_3, \ldots\ldots, y_{n-1}\}$ we predict the element y_n using the aforementioned Recursive Future Equation. And one of the $2^{(n+mn)}$ number of 6th Order Polynomial Equations of the kind as shown in equation A which gives the best true value of y_n can be considered as the correct equation and its past element forecast of y_0 as the correct forecast.

References

1.Bagadi, R. (2016). Proof Of As To Why The Euclidean Inner Product Is A Good Measure Of Similarity Of Two Vectors. *PHILICA.COM Article number 626*. See the Addendum as well.
http://philica.com/display_article.php?article_id=626
2.http://www.vixra.org/author/ramesh_chandra_bagadi
3.http://philica.com/advancedsearch.php?author=12897