Forecast Framework

Time Series Forecasting

Theoretical base

Igor Uspeniev

Document version: 1.0.9 (1 November 2016)

CONTENTS

CONVENTIONAL SIGNS

N – set of natural numbers.
Z – set of integer numbers.
R – set of real numbers.
◀ – begin of proof.
▶ – end of proof.
$[a,b]$ – interval between a and b that contains a and b.
$a \in A$ – element a belongs to set A.
$A=\{a,b,c\}$ – set A contains elements a, b, c.
$A \subset B$ – subset A is contained in set B.
$A \subseteq B$ – subset A is contained in set B or coincides with it.
$A = B \cup C$ – set A is a union of sets B and C.
$\{x_i\}$ – infinite set of elements x_i.
$\{(a_i,b_i):\ldots\}$ – set of elements (a_i,b_i) with conditions ... .
(a_0,b_0) – point with coordinates a_0, b_0.
$U(x_0)$ – neighbourhood of point x_0.
$U(x_0,\varepsilon)$ – ε-neighbourhood of point x_0.
$\exists v:\ldots$ – existential quantification: there is at least one v, that... .
$\forall v$ – universal quantification: for any v.
$v = f(t)$ – variable v is a function of variable t.
$f(t,v,g,u)$ – function of several variables: t,v,g,u.
$u(t_0)$ – value of function $u(t)$ in point t_0.
$g(u(t))$ – function composition: $y=u(t)$, $g(y)$; application of one function to the result of another function.
$\sum_{i=m}^{n} d_i$ – sum of $(n-m+1)$ elements $d_m, d_{m+1},\ldots,d_i,\ldots,d_{n-1},d_n$.
$i=1,\ldots,n$ – number i consequently takes on a natural values from 1 to n inclusively.
$t \to b$ – variable t approaches to point b.
$f(x)=\begin{cases} a:\ldots \\ b:\ldots \end{cases}$ – function $f(x)$ takes a value a with conditions ..., and takes a value b with conditions ... .
dt – differential of argument t.
$v'(t), v'_t, \frac{dv}{dt}, v'$ – derivative of the function $v = f(t)$.
$v^{(n)}$ – derivative with n order of function v.
a^b – exponentiation with the base a and the exponent b.
$\ln a$ – natural logarithm of the number a (on the base of e).
$\log_a b$ – logarithm of the number b on the base of a.
$\sin a$, $\cos a$, $\tan a$, $\cot a$ – sine, cosine, tangent, cotangent of the number a.
$\arcsin a$, $\arccos a$, $\arctan a$, $\mathrm{arccot}\, a$ – arcsine, arccosine, arctangent, arccotangent of the number a.
$\sinh a$, $\cosh a$, $\tanh a$, $\coth a$ – hyperbolic sine, cosine, tangent, cotangent of the number a.

1. PREFACE

The time series forecasting problem is usually associated with the exploration of process behavior with not well studied nature. In this case there is a set of external observable variables that change values over time, forming a set of time series. The document is devoted to an exposition of the theory of time series forecasting, the argumentation of its fundamental principles and the research of forecasting task solving. The goal – to promote the development of universal methods of time series analysis and forecasting without expert analysis.

At the present time the conventional approach of time series forecasting is a heuristic. Each time series is forecasted with direct or implicit expert influence with a predefined set of approximating functions and subjectively evaluating of the result. Even systems known as self-learning use predefined templates or principles. And most importantly – none of the existing forecasting methods can be analytically substantiated by forecast optimality criteria, because these criterias currently are not exist. Statistical methods only gives the answer to the admissibility question of the selected prediction method until the first unsatisfactory mismatch detection of the expected value with the observed.

The more complex the behavior of the observed values, the harder to choose the heuristic model and lower its accuracy. Regardless of the adopted heuristics only in special cases it is possible to choose the model of the observed value behavior and forecast time series within the acceptable margin of error. If there is a sufficient information about the interdependency of the internal process objects, it is possible to decompose the problem for forecasting its individual parts. However, this also does not guarantee the success of heuristic matching the behaviors of each part. Possibilities of experts including a fully automated are always limited, and starting from a certain level of entropy time series data source appear as is unpredictable random values for any heuristic algorithm.

With this approach, only a well-studied and decomposed process can be successfully predicted by its time series values.

To build a more effective approach theoretical basis of forecasting is needed: its functional definition, the existence and uniqueness conditions, the fundamental possibility of finding a solution, and determining method.

P.S.: The author apologizes for possible stylistic or other sort of errors in document, English is not my native language. I will be grateful to you for any information about errors in document. Thanks a lot!

2. GENERAL INFORMATION AND DEFINITIONS

Time Series

Time Series is a finite observed values set of observed value v in different time moments t. $\Omega = \{(t_i, v_i) : t \in T; v \in V; i \in N\}$, where T is a range of definition of time, V is a range of values of observed value. Cardinality of a set can be infinite, but practically analysed number of elements is always finite. As observed value can be considered:

- Value of unknown initial function $v(t)$,
- The value of the numerical property of conventionally closed physical system. The change in time of the numerical property can also be represented as an unknown initial function $v(t)$.

Time series forecasting is determining of probability distribution of expected value v in desired time moments t.

For convenience of treatment, we will assume T as united range of definition which includes time moments of time series and time moments of desired forecast points.

Depending of initial function nature the values of this function in time series can be represented as a single values or as confidential interval specifically defined for each individual element.

In case of the physical system properties forecasting it is essential to know about the encountered data measurement inaccuracies and incomplete closure of the physical system, introducing the random deviations in the source time series data. In this case there is a confidential interval distinct for each measure, so the time series represented as a set of elements $\{(t_i, v\min_i, v\max_i)\}$. This can be equivalently represented as notation $\{(t_i, v_i, \varepsilon_i)\}$, where $v_i = \frac{v\max_i + v\min_i}{2}$ is a middle of confidential interval, and $\varepsilon_i = \frac{v\max_i - v\min_i}{2}$ is an acceptable deviation in this point. In this case the values of initial function in time moment t_i contained in ε_i-neighbourhood of point v_i: $v(t_i) \in U(v_i, \varepsilon_i)$.

If case of accurate algebraic function forecasting represented as time series without measurement inaccuracies the acceptable deviation at any time point approaches to zero: $\varepsilon \to 0 \Leftrightarrow v(t_i) \to v_i$.

Instrument of time series forecasting $\Omega = \{(t_i, v_i) : t \in T; v \in V; i = 1, \ldots, n; n \in N\}$ with finite element number is a set $\{m_j(t)\}$ of function-models of initial function $v(t)$. From all conceivable variety of functions belong to set $\{m_j(t)\}$ only those which satisfy the acceptable deviation for all elements of time series:

$$m_j(t_i) \in U(v_i, \varepsilon_i) : i = 1, \ldots, n \quad (2.1)$$

It should be noted that in this definition even when $\varepsilon_i \to 0 : \forall i$ the set $\{m_j(t)\}$ will consist of infinite number of function.

Function-models in set $\{m_j(t)\}$ are not equisignificant by usage priority in forecasting. The method of numerical determination and it's argumentation will be discussed further. So far, we will define and use the function $\rho(m_j)$ usage priority of function-model without particularization. The probability distribution of the expected value v in desired time moment t_0 equals to the total probability distribution of sheaf of function-models taking into account their usage priority:

$$F_v(t_0, v_0) = P(v \le v_0) \equiv \frac{\sum_{i=0}^{n} P(m_i(t_0) \le v_0) \cdot \rho(m_i)}{\sum_{i=0}^{n} \rho(m_i)} : \quad F_v(t_0, v_0) \in [0,1]; n \in N; P(x) = \begin{cases} 0 : x = false \\ 1 : x = true \end{cases} \quad (2.2)$$

Axiom of function-model supportability

Function-models are determined exclusively by the time series data source created by the initial function. Neither form of the functions models, or their range of values and properties are not defined a priori and are not randomly assigned.

In the absence of this axiom unambiguous association between the condition and the decision can not be built, and the optimal solution criteria can not be the not only proven, but even considered.

This seemingly simple axiom defines all logical reasoning in the theory of time series forecasting. In the following, we prove a series of conclusions which are consequences of the axioms of function-model supportability.

Entropy of function

Now we will introduce the definition of entropy of function:

Entropy of function or process is a measure of the indefinitiveness and complexity of the formal description of the function or process.

Methods of numerical analysis of the entropy of function will be considered further.

Пока же примем условное определение: чем выше неопределенность функции, тем выше ее энтропия. Следовательно, наименьшей энтропией обладает константа, неизменная во времени, а наибольшей энтропией обладает истинно случайная функция, мера неопределенности которой не поддается описанию.

As far, we will use a conditional definition: the higher the function complexity, the higher the entropy. Therefore, the lower entropy has a constant unchanging in time, and the greatest entropy has a truly random function with a measure of indefinitiveness that can not be described.

Max entropy of function-model

The maximum entropy of function-models determined by the number of time series elements.

◄ Since the function-models defined exclusively by the time series data source, the entropy of these function-models is defined exclusively by the time series data source. The amount of information contained in the set of elements of the time series is limited to the capacity of a given set. Therefore, the maximum entropy of function-models determined by the number of time series elements. ►

Min entropy of function-model

The minimum entropy of function-model is determined by the necessary complexity of function-model appertained to time series elements.

◄ By the definition of the function-models set, only those functions belongs to set that satisfy the acceptable deviation. Therefore the entropy of function-model should be sufficient to cover the variance of the time series. ►

Entropy of function models increase

Increasing the number of time series elements leads to an increase or invariance of entropy of function-models formed by this time series.

◄ Suppose that for finite time series $\Omega_n = \left\{ \left(t_i, v_i, \varepsilon_i \right) : t \in T; v \in V; i = 1,\ldots,n; n \in N \right\}$ there is a set M_n of function-models defined. Entropy of each function-model belongs to interval $E_j \in \left[a,b \right]$, where the lower limit defined by necessary complexity and the upper limit defined by the capacity of the given set Ω_n.

If the time series will be extended by element $\left(t_{n+1}, v_{n+1}, \varepsilon_{n+1} \right)$ then the resulting time series $\Omega_{n+1} = \left\{ \left(t_i, v_i, \varepsilon_{n+1} \right) : t \in T; v \in V; i = 1,\ldots,n+1; n \in N \right\}$ will corresponds to the set of function-models M_{n+1} that contains union of sets:

- Subset $\overline{M_n} \subseteq M_n$ where function-models satisfy the acceptable deviation of added element: $m_j \left(t_{n+1} \right) \in U \left(v_{n+1}, \varepsilon_{n+1} \right)$,
- Set $\overline{M_{n+1}}$ where all function-models have maximum possible entropy for Ω_{n+1}.

$M_{n+1} = \overline{M_n} \cup \overline{M_{n+1}}$.

Therefore, increasing the number of time series elements leads to an increase or invariance of entropy of function-models formed by this time series. ►

Finitude of necessary argument

If initial function v is forecastable non-random function then there exists at least one function-model m_0 that satisfies the acceptable deviation (2.1) in the time interval $t \in T$ and m_0 obtaining requires finite number of time series elements.

◄ We will prove this statement by contradiction: let us assume that function v is forecastable non-random function, but there is not any function-model m that satisfies the acceptable deviation (2.1) in the time interval $t \in T$ and can be obtained by finite number of time series elements.

Suppose that there is a time series $\Omega_n = \{(t_i, v_i, \varepsilon_i) : t \in T; i = 1, \ldots, n\}$ and the set of function-models M_n directly obtained from data of this time series. There is no any function-model in M_n that satisfies the acceptable deviation (2.1) in the time interval $t \in T$. By adding elements $(t_{n+j}, v_{n+j}, \varepsilon_{n+j}) : \forall i, \forall t_j$ to the time series the entropy of function-models in set M_{n+j} will increase compared to M_n, in the other case there is at least one function-model in M_n that satisfies the acceptable deviation (2.1) in the time interval $t \in T$. Thus, the minimal entropy of function-models can not be limited, it means that function-models are random functions, so the initial function is a rundom function and can not be forecasted. This points to the wrong source of the statement by contradiction.

Therefore, if initial function v is forecastable non-random function then there exists at least one function-model m_0 that satisfies the acceptable deviation (2.1) in the time interval $t \in T$ and m_0 obtaining requires finite number of time series elements. ►

Usage priority of function-model

For general matching of function-model m_0 to initial function it is necessary and sufficient that $m_0(t)$ that corresponds to time series $\Omega_n = \{(t_i, v_i, \varepsilon_i) : t \in T; v \in V; i = 1, \ldots, n; n \in N\}$ by acceptable deviation condition (2.1) retain this property for any $\Omega_{n+k} = \{(t_i, v_i, \varepsilon_i) : t \in T; v \in V; i = 1, \ldots, n+k; n \in N; k \in N\}$ element extension. Every additional element $(t_{n+j}, v_{n+j}, \varepsilon_{n+j})$ that satisfies the acceptable deviation (2.1) will increase probability of function-model $m_0(t)$ matching to the initial function v.

Let us adopt the symbols:

$E_v(k)$ – number of time series elements that defines the maximum entropy of function-model to forecast function v for k elements in time series. As we have defined in conclusion about max entropy of function-model, $E_v(k) = k$.

$E_m(k)$ – number of time series elements that defined in fact the entropy of function-model m for k elements in time series.

Thus, $E_v(k) - E_m(k)$ is a number of elements that does not result to entropy increase of function-model m. This difference specifies the excess quantity of elements with entropy invariance of function-model.

Suppose that there is a function-model m with maximum possible entropy for k elements of time series Ω_k, so $E_m(k) = E_v(k) = k$. Now we extend the time series by element $(t_{n+1}, v_{n+1}, \varepsilon_{n+1})$ and express the usage priority function of function-model m in forecasting initial function v as a division between excess quantity of elements to excess of function divergence:

$$\alpha(m_{n+1}) = \frac{E_v(k+1) - E_m(k+1)}{\left| v(t_{n+j}) - m(t_{n+j}) \right|} \quad (2.3)$$

If $(t_{n+1}, v_{n+1}, \varepsilon_{n+1})$ leads to function-model m entropy increase then $E_v(k+1) - E_m(k+1) = 0$.

If $\left(t_{n+1},v_{n+1},\varepsilon_{n+1}\right)$ doesn't lead to function-model m entropy increase then $E_v\left(k+1\right)-E_m\left(k+1\right)=1$.

$\left|v\left(t_{n+j}\right)-m\left(t_{n+j}\right)\right|$ defines excess of function divergence. If case of values matching and entropy invariance $\alpha\left(m_{n+1}\right)\to\infty$, so usage priority of function-model is maximum.

Now we will express usage priority function of function-model m in forecasting initial function v for a set of elements:

$$\rho\left(m\right)=\frac{E_v\left(n\right)-E_m\left(n\right)}{\sum_{i=1}^{n}\left(\left|v\left(t_i\right)-m\left(t_i\right)\right|\right)} \quad (2.4)$$

Accounting for $E_v\left(n\right)=n$ we get:

$$\rho\left(m\right)=\frac{n-E_m\left(n\right)}{\sum_{i=1}^{n}\left(\left|v\left(t_i\right)-m\left(t_i\right)\right|\right)} \quad (2.5)$$

In case of analitical formulation of $v\left(t\right)$ and $m\left(t\right)$ then having labeled entropy of function f increasing on interval T as $\int_T U\left(f\right)\cdot dt$ formula will become:

$$\rho\left(m\right)=\frac{\int_T\left(U\left(v\right)-U\left(m\right)\right)\cdot dt}{\int_T\left|v\left(t\right)-m\left(t\right)\right|\cdot dt} \quad (2.6)$$

It should be noted that $\rho\left(m\right)$ is relatively value and can be used only in function-model compare.

Thus, formula $\left(2.2\right)$ becomes:

$$\begin{cases} F_v\left(t_0,v_0\right)=P\left(v\le v_0\right)\equiv\dfrac{\sum_{i=0}^{n}P\left(m_i\left(t_0\right)\le v_0\right)\cdot\rho\left(m_i\right)}{\sum_{i=0}^{n}\rho\left(m_i\right)} \\ \rho\left(m_i\right)=\dfrac{n-E_{m_i}\left(n\right)}{\sum_{j=1}^{n}\left(\left|v\left(t_j\right)-m_i\left(t_j\right)\right|\right)} \end{cases} : F_v\left(t_0,v_0\right)\in\left[0,1\right];\, n\in N;\, P\left(x\right)=\begin{cases}0: x=\mathit{false}\\ 1: x=\mathit{true}\end{cases} \quad (2.7)$$

If $\varepsilon\to 0$ then $\sum_{j=1}^{n}\left(\left|v\left(t_j\right)-m_i\left(t_j\right)\right|\right)\to 0$:

$$F_v\left(t_0,v_0\right)=\lim_{\varepsilon\to 0}\frac{\sum_{i=0}^{n}P\left(m_i\left(t_0\right)\le v_0\right)\cdot\dfrac{n-E_{m_i}\left(n\right)}{\sum_{j=1}^{n}\left(\left|v\left(t_j\right)-m_i\left(t_j\right)\right|\right)}}{\sum_{i=0}^{n}\dfrac{n-E_{m_i}\left(n\right)}{\sum_{j=1}^{n}\left(\left|v\left(t_j\right)-m_i\left(t_j\right)\right|\right)}}\equiv\frac{\sum_{i=0}^{n}P\left(m_i\left(t_0\right)\le v_0\right)\cdot\left(n-E_{m_i}\left(n\right)\right)}{\sum_{i=0}^{n}\left(n-E_{m_i}\left(n\right)\right)} \quad (2.8)$$

Thus, in case of $\varepsilon\to 0$ formula $\left(2.7\right)$ becomes:

$$\begin{cases} F_v\left(t_0,v_0\right)=P\left(v\le v_0\right)\equiv\dfrac{\sum_{i=0}^{n}P\left(m_i\left(t_0\right)\le v_0\right)\cdot\rho\left(m_i\right)}{\sum_{i=0}^{n}\rho\left(m_i\right)} \\ \rho\left(m_i\right)=n-E_{m_i}\left(n\right) \end{cases} : F_v\left(t_0,v_0\right)\in\left[0,1\right];\, n\in N;\, P\left(x\right)=\begin{cases}0: x=\mathit{false}\\ 1: x=\mathit{true}\end{cases} \quad (2.9)$$

At this stage the method of full function-models $\{m_i\}$ set generation by time series data source is sill undefined. Also is unexpressed the number of elements $E_{m_i}(n)$ that defines entropy of function-model. This will be analyzed in the next section.

3. THEOREM OF UNIVERSAL FORM

Usage restrictions

To date, all the variety of functions can not be represented in the form of arithmetic operations with elementary functions and their compositions. Consequently, function formulating can not be limited to this definition, since setting the problem to make forecasting of unknown function possible, we have to maximize the analyzed set of functions.

On the other hand while extending the set of function definition, we are inevitably challenged with the question of the functions defining method. And first of all with question – consider whether the unknown function is differentiable.

The answer to this question is irreversibly separates the way of solving the problem.

Consider the positive response to the question of differentiability.

In this case, the set of the functions greatly expanded, as a significant, and perhaps immeasurably greater number of conceivable functions have only a differential representation.

For example, differential equation $\frac{du}{d\varphi} = e^{-\varphi^2}$ has a solution in quadratures $u = \int e^{-\varphi^2} \cdot d\varphi + a : a = const$, the integral of which can not be found. But this is not to preclude to function $u(\varphi)$ existance and to be one of possible initial functions or function-models that create or forecast the time series. Functions expressed in quadrature like $u(\varphi)$ is infinite set, and their forms and compositions may be much more complex.

Moreover, the functions given in the form of differential equations are solvable in quadratures only in exceptional cases. Significantly most of the functions exists only in the form of differential equations without the possibility of transformation them to form $v = f(t)$ even in quadratures.

As is known, to quadratures can be reduced first-order differential equation with an integrating factor, homogeneous and quasi-homogeneous equation, Bernoulli equation and Riccati equation. Higher order differential equations have a solution in quadratures only in the case of linear representation, and even a slight lack of homogeneity leads only to the evaluation criterion of solution verifying as in Liouville's formula. Known common methods for solving inhomogeneous differential equation does not exist, but the functions that satisfy this equation, objectively exist and can act as time series functional data source and need to be used as function-models for forecasting.

Thus, in the case of an affirmative answer to the question of differentiability, especially with infinitely differentiable, we are able to greatly expand the variety of functions due to their representation in the form of differential equations.

However, non-differentiable functions are lost from view. Consider the negative or uncertain answer to the question of differentiability.

This way of reasoning is much more difficult, because this brings up the question of the describing method of non-differentiable functions, each of which can be very specific, ranging from the simplest options contained in the definition of the module, to special definitions for interrupted and descrete functions or special functions such as the Dirichlet function.

Nondifferentiable functions do not belong to one group for which special methods of functional analysis are developed. They just ***do not*** belong to the group of differentiable functions. At the same time, denying the possibility of differentiation does not give additional theoretical tools of functional analysis, but only eliminates many known tools common to differentiable functions.

Within this theory presented, made a choice in favor of differentiable functions. In the future it is possible expansion of functions set toward some non-differentiable functions, but in the framework of these theoretical expressions all the functions in will be considered infinitely differentiable on the interval T of the time series and it's forecasting.

Goal

To apply the functions given in the form of differential equations, as time-series forecasting function-models necessary to determine the method of equivalent transformation of any differential equation to the universal form. This form should be varied only on the basis of the numerical value of entropy, while retaining the possibility of a universal representation of all conceivable functions defined by differential equations of arbitrary form.

To achieve this we will consider a few statements will prove them in a general way and then consider the specific examples.

Differential form without function of argument

Suppose that there is an infinitely differentiable function defined arbitrary differential equation, which can include the argument value, and elementary functions (we will call them radicals $\{R_i\}$):

$$F\left(t,\left\{v^{(i)}\right\},\left\{R_j\right\}\right)=0: i=0,\ldots,z;\ j=1,\ldots,\mu;\ \{z,\mu\}\in N \quad (3.1)$$

We will prove that $F\left(t,\left\{v^{(i)}\right\},\left\{R_j\right\}\right)=0$ can be equivalently expressed by a differential equation of higher order with the set of initial conditions, consisting only of arithmetic operations, the operations of differentiation and not containing the argument value and radicals:

$$F\left(t,\left\{v^{(i)}\right\},\left\{R_j\right\}\right)=0: i=0,\ldots,z;\ j=1,\ldots,\mu;\ \{z,\mu\}\in N \Leftrightarrow \begin{cases}\sum_{i=0}^{n} a_i\cdot\prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}}=0:\left\{n,m,p_{i,j}\right\}\in N; a_i\in R\\ v^{(i)}\left(t_0\right)=v_0^{(i)}:\ i=z,\ldots,m-1; m>z\end{cases} \quad (3.2)$$

Equivalence inversion provided by initial conditions of differential values as soving of Cauchy problem.

◄ Proof consists of several stages.

1. Initial function.

Every conceivable function consists of the components: the argument values, arithmetic, elementary functions and their compositions, the integration and differentiation operations. In this case, all elementary functions (for the sake of brevity, let's call them radicals) have either simple arithmetic, or differential interdependencies. In particular:

$$f_1=\ln x \quad f_1'=(x)^{-1}\cdot x'$$

$$f_2=\log_a x \quad f_2'=\frac{1}{\ln a}\cdot(x)^{-1}\cdot x'$$

$$f_3=x^a \quad f_3'=a\cdot x^{a-1}\cdot x' \quad \Rightarrow \quad f_3'=a\cdot(x)^{-1}\cdot x'\cdot f_3$$

$$f_4=a^x \quad f_4'=\ln a\cdot x'\cdot a^x \quad \Rightarrow \quad f_4'=\ln a\cdot x'\cdot f_4$$

$$f_5=\sin x \quad f_5'=\cos x\cdot x' \quad f_5''=-\sin x\cdot(x')^2+\cos x\cdot x'' \quad \Rightarrow \quad f_5''=-(x')^2\cdot f_5+(x')^{-1}\cdot x''\cdot f_5'$$

$$f_6=\cos x \quad f_6'=-\sin x\cdot x' \quad f_6''=-\cos x\cdot(x')^2-\sin x\cdot x'' \quad \Rightarrow \quad f_6''=-(x')^2\cdot f_6+(x')^{-1}\cdot x''\cdot f_6'$$

$$f_7=\tan x \quad f_7'=\frac{x'}{\cos^2 x}=x'\cdot\left(f_6\right)^{-2}$$

$$f_8=\cot x \quad f_8'=-\frac{x'}{\sin^2 x}=-x'\cdot\left(f_5\right)^{-2}$$

$$f_9=\arcsin x \quad f_9'=\frac{x'}{\sqrt{1-x^2}} \quad f_9''=\frac{x''-x^2\cdot x''+x\cdot x'\cdot x'}{\left(1-x^2\right)^{\frac{3}{2}}}=\left((x')^{-3}\cdot x''-(x)^2\cdot(x')^{-3}\cdot x''+x\cdot(x')^{-1}\right)\cdot\left(f_9'\right)^3$$

$$f_{10}=\arccos x \quad f_{10}'=-\frac{x'}{\sqrt{1-x^2}} \quad f_{10}''=\left((x')^{-3}\cdot x''-(x)^2\cdot(x')^{-3}\cdot x''+x\cdot(x')^{-1}\right)\cdot\left(f_{10}'\right)^3$$

$$f_{11}=\arctan x \quad f_{11}'=\frac{x'}{1+x^2} \quad f_{11}''=\frac{x''\cdot\left(1+x^2\right)-2\cdot x\cdot(x')^2}{\left(1+x^2\right)^2}=\left((x')^{-2}\cdot x''+x^2\cdot(x')^{-2}\cdot x''-2\cdot x\right)\cdot\left(f_{11}'\right)^2$$

$f_{12} = \operatorname{arccot} x \quad f'_{12} = -\dfrac{x'}{1+x^2} \quad f''_{12} = -\left((x')^{-2} \cdot x'' + x^2 \cdot (x')^{-2} \cdot x'' - 2 \cdot x \right) \cdot (f'_{12})^2$

$f_{13} = \operatorname{sh} x \quad f'_{13} = x' \cdot \operatorname{ch} x = x' \cdot f_{14}$

$f_{14} = \operatorname{ch} x \quad f'_{14} = x' \cdot \operatorname{sh} x = x' \cdot f_{13}$

$f_{15} = \operatorname{th} x \quad f'_{15} = \dfrac{x'}{\operatorname{ch}^2 x} = x' \cdot (f_{14})^{-2}$

$f_{16} = \operatorname{cth} x \quad f'_{16} = -\dfrac{x'}{\operatorname{sh}^2 x} = -x' \cdot (f_{13})^{-2}$

Thus, multiple differentiation of radicals leads to one of the three cases:

- Simple arithmetic operations, such as in the case of the log, arcsine, arccosine, arctangent, arccotangent etc.,
- Cyclic dependencies between a finite number of radical differentials $f\left(R, R', \ldots, R^{(q)}\right) = 0$, for example, in the case of sine, cosine, etc.,
- Cyclic dependencies between a finite number of several radical differentials $f\left(R_1, R'_1, \ldots, R_1^{(q_1)}, R_2, R'_2, \ldots, R_2^{(q_2)}, \ldots, R_\mu, R'_\mu, \ldots, R_\mu^{(q_\mu)}\right) = 0$, for example, between the hyperbolic sine and hyperbolic cosine, logarithm and a natural logarithm, cosine and tangent, and so forth..

The first two cases are special and the third case is the general which includes first and second.

2. Complementation of finite variability in radical differentiation.

Thus, the radicals in its differentiation have finite variability, so equation $F\left(t, \left\{v^{(i)}\right\}, \left\{R_j\right\}\right) = 0 : i = 0, \ldots, z;\ j = 1, \ldots, \mu;\ \{z, \mu\} \in N$ that contains finite set of radicals can be differentiated finite times $F_i = F^{(i)}$ to build this system of equations:

$$\begin{cases} F_0\left(t, \left\{v^{(i)}\right\}, R_1, R_2, \ldots, R_\mu\right) = 0 : i = 0, \ldots, z \\ F_1\left(t, \left\{v^{(i)}\right\}, R_1, R_1^{(1)}, R_2, R_2^{(1)} \ldots, R_\mu, R_\mu^{(1)}\right) = 0 : i = 0, \ldots, z+1 \\ F_2\left(t, \left\{v^{(i)}\right\}, R_1, R_1^{(1)}, R_1^{(2)}, R_2, R_2^{(1)}, R_2^{(2)} \ldots, R_\mu, R_\mu^{(1)}, R_\mu^{(2)}\right) = 0 : i = 0, \ldots, z+2 \\ \ldots \\ F_k\left(t, \left\{v^{(i)}\right\}, R_1, R_1^{(1)}, \ldots, R_1^{(q_1)}, R_2, R_2^{(1)}, \ldots, R_2^{(q_2)} \ldots, R_\mu, R_\mu^{(1)}, \ldots, R_\mu^{(q_\mu)}\right) = 0 : i = 0, \ldots, z+k \end{cases}$$

where $k = 1 + \sum_{i=0}^{m} q_i$ due to the finiteness of the differential relationships between radicals. According to the fundamental theorem of algebra, this system of equations in the set of variables $\left\{R_j^{(h)}\right\}$ corresponds to equation without radicals: $\overline{F}\left(t, \left\{v^{(i)}\right\}\right) = 0 : i = 0, \ldots, z+k$ that contains only values of argument, function differentials and arithmetical operations: addition, subtraction, multiplication and division, and also raising the values t и $v^{(i)}$ to integer power:

$$\sum_{i=0}^{x} \left(\omega_i \cdot \prod_{j=0}^{z+k} \left(v^{(j)}\right)^{b_{i,j}} \cdot t^{d_i} \right) = 0 : b_{i,j} \in Z;\ d_i \in Z; x \in N; \omega_i \in R \quad (3.3)$$

Conversion from (3.1) to (3.3) is inversible in case of defining starting conditions on each additional level of differentiating: $v^{(j)}(t_0) = v_0^{(j)} : i = z, \ldots, z+k-1;\ \{z, k\} \in N$

3. Factorization of argument value

Yaking into account the linear dependency between summands $\omega_i \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{i,j}} \cdot t^{d_i}$, we will differentiate them x times and put them into Wronskian determinant:

$$\begin{vmatrix} \omega_0 \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{0,j}} \cdot t^{d_0} & \omega_1 \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{1,j}} \cdot t^{d_1} & \cdots & \omega_x \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{x,j}} \cdot t^{d_x} \\ \left(\omega_0 \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{0,j}} \cdot t^{d_0}\right)' & \left(\omega_1 \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{1,j}} \cdot t^{d_1}\right)' & \cdots & \left(\omega_x \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{x,j}} \cdot t^{d_x}\right)' \\ \vdots & \vdots & \ddots & \vdots \\ \left(\omega_0 \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{0,j}} \cdot t^{d_0}\right)^{(x)} & \left(\omega_1 \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{1,j}} \cdot t^{d_1}\right)^{(x)} & \cdots & \left(\omega_x \cdot \prod_{j=0}^{z+k}\left(v^{(j)}\right)^{b_{x,j}} \cdot t^{d_x}\right)^{(x)} \end{vmatrix} = 0 \quad (3.4)$$

Disclosure determinant leads to the removal of the brackets and the reduction of $t^{\sum_{i=0}^{x} d_i - \frac{x\cdot(x+1)}{2}}$, so our differential equation got rid of argument t:

$$\sum_{i=0}^{n} a_i \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}} = 0: \quad \left\{n, m, p_{i,j}\right\} \in N; a_i \in R \quad (3.5)$$

Conversion from (3.3) to (3.5) is inversible in case of defining starting conditions on each additional level of differentiating $v^{(i)}(t_0) = v_0^{(i)}: \ i = \mu, \ldots, m-1; m > \mu$ that defines Cauchy problem.

Function satisfying (3.1) will satisfy (3.5). And back to front with starting conditions of Cauchy problem function satisfying (3.5) will satisfy (3.1).►

The reduction to a predetermined view

Now we will prove that the equation (3.5) by finite number of reexpressions can be reduced to equation with indefinitiveness only in differentials.

◄ Let the function given by the equation: $\sum_{i=0}^{n} a_i \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}} = 0$, where $\{a_i\}, \{p_{i,j}\}$ are unknown constants.

We introduce an additional function $g_{0,i} = 1: \forall i$, at the first stage, it will play the role of a neutral factor:

$$\sum_{i=0}^{n} a_i \cdot g_{0,i} \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}} = 0 \quad (3.6)$$

Divide (3.6) on $g_{0,n} \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{n,j}}$:

$$\sum_{i=0}^{n-1}\left\{a_i \cdot \frac{g_{0,i}}{g_{0,n}} \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}-p_{n,j}}\right\} + a_n = 0 \qquad (3.7)$$

Differentiate (3.7) by dt:

$$\left(\sum_{i=0}^{n-1} a_i \cdot \frac{g_{0,i}}{g_{0,n}} \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}-p_{n,j}} + a_n\right)' = 0$$

$$\sum_{i=0}^{n-1} a_i \cdot \left(\left(\frac{g_{0,i}}{g_{0,n}}\right)' \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}-p_{n,j}} + \frac{g_{0,i}}{g_{0,n}} \cdot \prod_{j=0}^{m}\left(v^{(j)}\right)^{p_{i,j}-p_{n,j}} \cdot \sum_{j=0}^{m}\left(p_{i,j}-p_{n,j}\right) \cdot \frac{v^{(j+1)}}{v^{(j)}}\right) + \left(a_n\right)' = 0$$

$$\sum_{i=0}^{n-1} a_i \cdot \left(\left(\frac{g_{0,i}}{g_{0,n}} \right)' + \frac{g_{0,i}}{g_{0,n}} \cdot \sum_{j=0}^{m} \left(p_{i,j} - p_{n,j} \right) \cdot \frac{v^{(j+1)}}{v^{(j)}} \right) \cdot \prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{i,j} - p_{n,j}} = 0 \quad (3.8)$$

Multiply (3.8) on $\prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{n,j}}$:

$$\sum_{i=0}^{n-1} a_i \cdot \left(\left(\frac{g_{0,i}}{g_{0,n}} \right)' + \frac{g_{0,i}}{g_{0,n}} \cdot \sum_{j=0}^{m} \left(p_{i,j} - p_{n,j} \right) \cdot \frac{v^{(j+1)}}{v^{(j)}} \right) \cdot \prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{i,j}} = 0$$

$$\sum_{i=0}^{n-1} a_i \cdot \left(\frac{g_{0,i}' \cdot g_{0,n} - g_{0,i} \cdot g_{0,n}'}{\left(g_{0,n} \right)^2} + \frac{g_{0,i}}{g_{0,n}} \cdot \sum_{j=0}^{m} \left(p_{i,j} - p_{n,j} \right) \cdot \frac{v^{(j+1)}}{v^{(j)}} \right) \cdot \prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{i,j}} = 0$$

$$\sum_{i=0}^{n-1} a_i \cdot \left(g_{0,i}' \cdot g_{0,n} - g_{0,i} \cdot g_{0,n}' + g_{0,i} \cdot g_{0,n} \cdot \sum_{j=0}^{m} \left(p_{i,j} - p_{n,j} \right) \cdot \frac{v^{(j+1)}}{v^{(j)}} \right) \cdot \prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{i,j}} = 0 \qquad (3.9)$$

Introducing a new function $g_{1,i} = g_{0,i}' \cdot g_{0,n} - g_{0,i} \cdot g_{0,n}' + g_{0,i} \cdot g_{0,n} \cdot \sum_{j=0}^{m} \left(p_{i,j} - p_{n,j} \right) \cdot \frac{v^{(j+1)}}{v^{(j)}}$ we will get:

$$\sum_{i=0}^{n-1} a_i \cdot g_{1,i} \cdot \prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{i,j}} = 0 \qquad (3.10)$$

Thus, by using $g_{1,i}$ instead of $g_{0,i}$ we have reexpressed (3.6) to (3.10) which contains one less summand. Iteratively repeating steps from (3.6) to (3.10) and introducing $g_{2,i}$, $g_{3,i}$,..., $g_{n,i}$ we will get equation:

$$\sum_{i=0}^{n-n} a_i \cdot g_{n,i} \cdot \prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{i,j}} = a_0 \cdot g_{n,0} \cdot \prod_{j=0}^{m} \left(v^{(j)} \right)^{p_{0,j}} = 0 \quad (3.11)$$

Equation (3.11) is equivalent to equation $g_{n,0} = 0$ that defined by recurrent formula:

$$\begin{cases} g_{k+1,i} = g_{k,i}' \cdot g_{k,n-k} - g_{k,i} \cdot g_{k,n-k}' + g_{k,i} \cdot g_{k,n-k} \cdot \sum_{j=0}^{m} \left(P_{i,j} - P_{n-k,j} \right) \cdot \frac{v^{(j+1)}}{v^{(j)}} \\ g_{0,i} = 1 \end{cases} \qquad (3.12)$$

Compare the method of specifying the function by equation (3.12) and by equation (3.5).

Equation (3.12) has higher differential order but less unknown constants – we have got rid of $\{a_i\}$. In the meantime equation (3.12) also reduced to sum of differential multiplication, but contrary to equation (3.5) powers in exponentiating in equation (3.12) depend only on n and m.

The values of the factors in the equation (3.12) depends on the exponents of the equation (3.5).

Such reexpression of reducing coefficient indefinitiveness from (3.5) to (3.12) can be performed twice – the second stage will operate with the result equation of the first stage. In this case we will get differential equation with fully known constants, summands and multipliers. Indefinitiveness will be only in differentials.►

Consider the example.

Suppose that function $y(t)$ defined by differential equation:

$$a_0 \cdot (y)^{p_{0,0}} \cdot (y')^{p_{0,1}} + a_1 \cdot (y)^{p_{1,0}} \cdot (y')^{p_{1,1}} + a_2 \cdot (y)^{p_{2,0}} \cdot (y')^{p_{2,1}} = 0 \quad (3.13)$$

Here $\{a_i\}$ и $\{p_{i,j}\}$ – unknown scalar constants. We introduce the additional function $g_{0,i} = 1 : \forall i$, at the first stage, it will play the role of a neutral factor:

$$a_0 \cdot g_{0,0} \cdot (y)^{p_{0,0}} \cdot (y')^{p_{0,1}} + a_1 \cdot g_{0,1} \cdot (y)^{p_{1,0}} \cdot (y')^{p_{1,1}} + a_2 \cdot g_{0,2} \cdot (y)^{p_{2,0}} \cdot (y')^{p_{2,1}} = 0 \quad (3.14)$$

Divide (3.14) on $g_{0,2} \cdot (y)^{p_{2,0}} \cdot (y')^{p_{2,1}}$:

$$a_0 \cdot \frac{g_{0,0}}{g_{0,2}} \cdot (y)^{p_{0,0}-p_{2,0}} \cdot (y')^{p_{0,1}-p_{2,1}} + a_1 \cdot \frac{g_{0,1}}{g_{0,2}} \cdot (y)^{p_{1,0}-p_{2,0}} \cdot (y')^{p_{1,1}-p_{2,1}} + a_2 = 0 \quad (3.15)$$

Differentiate (3.15) on dt:

$$\left(a_0 \cdot \frac{g_{0,0}}{g_{0,2}} \cdot (y)^{p_{0,0}-p_{2,0}} \cdot (y')^{p_{0,1}-p_{2,1}} + a_1 \cdot \frac{g_{0,1}}{g_{0,2}} \cdot (y)^{p_{1,0}-p_{2,0}} \cdot (y')^{p_{1,1}-p_{2,1}} + a_2 \right)' = 0$$

$$a_0 \cdot \left(\left(\frac{g_{0,0}}{g_{0,2}} \right)' \cdot (y)^{p_{0,0}-p_{2,0}} \cdot (y')^{p_{0,1}-p_{2,1}} + \frac{g_{0,0}}{g_{0,2}} \cdot (y)^{p_{0,0}-p_{2,0}} \cdot (y')^{p_{0,1}-p_{2,1}} \cdot \left((p_{0,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{0,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) +$$

$$+ a_1 \cdot \left(\left(\frac{g_{0,1}}{g_{0,2}} \right)' \cdot (y)^{p_{1,0}-p_{2,0}} \cdot (y')^{p_{1,1}-p_{2,1}} + \frac{g_{0,1}}{g_{0,2}} \cdot (y)^{p_{1,0}-p_{2,0}} \cdot (y')^{p_{1,1}-p_{2,1}} \cdot \left((p_{1,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{1,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) = 0$$

$$a_0 \cdot \left(\left(\frac{g_{0,0}}{g_{0,2}} \right)' + \frac{g_{0,0}}{g_{0,2}} \cdot \left((p_{0,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{0,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) \cdot (y)^{p_{0,0}-p_{2,0}} \cdot (y')^{p_{0,1}-p_{2,1}} +$$

$$+ a_1 \cdot \left(\left(\frac{g_{0,1}}{g_{0,2}} \right)' + \frac{g_{0,1}}{g_{0,2}} \cdot \left((p_{1,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{1,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) \cdot (y)^{p_{1,0}-p_{2,0}} \cdot (y')^{p_{1,1}-p_{2,1}} = 0 \quad (3.16)$$

Multiply (3.16) on $(y)^{p_{2,0}} \cdot (y')^{p_{2,1}}$ and expand the differential of the division:

$$a_0 \cdot \left(\frac{g_{0,0}' \cdot g_{0,2} - g_{0,0} \cdot g_{0,2}'}{(g_{0,2})^2} + \frac{g_{0,0}}{g_{0,2}} \cdot \left((p_{0,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{0,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) \cdot (y)^{p_{0,0}} \cdot (y')^{p_{0,1}} +$$

$$+ a_1 \cdot \left(\frac{g_{0,1}' \cdot g_{0,2} - g_{0,1} \cdot g_{0,2}'}{(g_{0,2})^2} + \frac{g_{0,1}}{g_{0,2}} \cdot \left((p_{1,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{1,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) \cdot (y)^{p_{1,0}} \cdot (y')^{p_{1,1}} = 0$$

Multiply both part of equation on $(g_{0,2})^2$:

$$a_0 \cdot \left(g_{0,0}' \cdot g_{0,2} - g_{0,0} \cdot g_{0,2}' + g_{0,0} \cdot g_{0,2} \cdot \left((p_{0,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{0,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) \cdot (y)^{p_{0,0}} \cdot (y')^{p_{0,1}} +$$

$$+ a_1 \cdot \left(g_{0,1}' \cdot g_{0,2} - g_{0,1} \cdot g_{0,2}' + g_{0,1} \cdot g_{0,2} \cdot \left((p_{1,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{1,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \right) \cdot (y)^{p_{1,0}} \cdot (y')^{p_{1,1}} = 0 \quad (3.17)$$

Now we have got two summands instead of the in the starting equation (3.13). Indroduce function:

$$\begin{cases} g_{1,0} = g_{0,0}' \cdot g_{0,2} - g_{0,0} \cdot g_{0,2}' + g_{0,0} \cdot g_{0,2} \cdot \left((p_{0,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{0,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \\ g_{1,1} = g_{0,1}' \cdot g_{0,2} - g_{0,1} \cdot g_{0,2}' + g_{0,1} \cdot g_{0,2} \cdot \left((p_{1,0}-p_{2,0}) \cdot \frac{y'}{y} + (p_{1,1}-p_{2,1}) \cdot \frac{y''}{y'} \right) \end{cases} \quad (3.18)$$

Using (3.18) in (3.17) will lead to:

$$a_0 \cdot g_{1,0} \cdot (y)^{p_{0,0}} \cdot (y')^{p_{0,1}} + a_1 \cdot g_{1,1} \cdot (y)^{p_{1,0}} \cdot (y')^{p_{1,1}} = 0 \quad (3.19)$$

Repeat these reexpression with (3.19) to get rid of one more summand:

$$a_0 \cdot \frac{g_{1,0}}{g_{1,1}} \cdot (y)^{p_{0,0}-p_{1,0}} \cdot (y')^{p_{0,1}-p_{1,1}} + a_1 = 0$$

$$\left(a_0 \cdot \frac{g_{1,0}}{g_{1,1}} \cdot (y)^{p_{0,0}-p_{1,0}} \cdot (y')^{p_{0,1}-p_{1,1}} + a_1 \right)' = 0$$

$$a_0 \cdot \left(\left(\frac{g_{1,0}}{g_{1,1}} \right)' \cdot (y)^{p_{0,0}-p_{1,0}} \cdot (y')^{p_{0,1}-p_{1,1}} + \frac{g_{1,0}}{g_{1,1}} \cdot (y)^{p_{0,0}-p_{1,0}} \cdot (y')^{p_{0,1}-p_{1,1}} \cdot \left(\left(p_{0,0}-p_{1,0}\right) \cdot \frac{y'}{y} + \left(p_{0,1}-p_{1,1}\right) \cdot \frac{y''}{y'} \right) \right) = 0$$

$$a_0 \cdot \left(g_{1,0}' \cdot g_{1,1} - g_{1,0} \cdot g_{1,1}' + g_{1,0} \cdot g_{1,1} \cdot \left(\left(p_{0,0}-p_{1,0}\right) \cdot \frac{y'}{y} + \left(p_{0,1}-p_{1,1}\right) \cdot \frac{y''}{y'} \right) \right) \cdot (y)^{p_{0,0}-p_{1,0}} \cdot (y')^{p_{0,1}-p_{1,1}} = 0$$

Using change $g_{2,0} = g_{1,0}' \cdot g_{1,1} - g_{1,0} \cdot g_{1,1}' + g_{1,0} \cdot g_{1,1} \cdot \left(\left(p_{0,0}-p_{1,0}\right) \cdot \frac{y'}{y} + \left(p_{0,1}-p_{1,1}\right) \cdot \frac{y''}{y'} \right)$ we have got:

$$a_0 \cdot g_{2,0} \cdot (y)^{p_{0,0}-p_{1,0}} \cdot (y')^{p_{0,1}-p_{1,1}} = 0 \quad \Rightarrow \quad g_{2,0} = 0$$

Taking in account that $g_{0,i} = 1 : \forall i$ we have got:

$$\begin{cases} g_{1,0} = \left(p_{0,0}-p_{2,0}\right) \cdot \frac{y'}{y} + \left(p_{0,1}-p_{2,1}\right) \cdot \frac{y''}{y'} \\ g_{1,1} = \left(p_{1,0}-p_{2,0}\right) \cdot \frac{y'}{y} + \left(p_{1,1}-p_{2,1}\right) \cdot \frac{y''}{y'} \\ g_{1,0}' \cdot g_{1,1} - g_{1,0} \cdot g_{1,1}' + g_{1,0} \cdot g_{1,1} \cdot \left(\left(p_{0,0}-p_{1,0}\right) \cdot \frac{y'}{y} + \left(p_{0,1}-p_{1,1}\right) \cdot \frac{y''}{y'} \right) = 0 \end{cases} \qquad (3.20)$$

Thus, we obtain a differential equation with known exponents and multiplying coefficients dependent on the exponents of the equation (3.13):

$$\begin{cases} \frac{y'''}{y} + b_0 \cdot \frac{(y'')^2}{y \cdot y'} + b_1 \cdot \frac{(y'')^3}{(y')^3} + b_2 \cdot \frac{y' \cdot y''}{(y)^2} + b_3 \cdot \frac{(y')^3}{(y)^3} = 0 \\ b_k = f_k\left(\left\{p_{i,j}\right\}\right) \end{cases} \qquad (3.21)$$

With this result we will perform operation of reducing coefficient indefinitiveness again:

$$\frac{y'''}{y} + b_0 \cdot \frac{(y'')^2}{y \cdot y'} + b_1 \cdot \frac{(y'')^3}{(y')^3} + b_2 \cdot \frac{y' \cdot y''}{(y)^2} + b_3 \cdot \frac{(y')^3}{(y)^3} = 0$$

$$(y)^2 \cdot (y')^{-3} \cdot y''' + b_0 \cdot (y)^2 \cdot (y')^{-4} \cdot (y'')^2 + b_1 \cdot (y)^3 \cdot (y')^{-6} \cdot (y'')^3 + b_2 \cdot y \cdot (y')^{-2} \cdot y'' + b_3 = 0 \quad (3.21)$$

Introducing change:

$$\begin{cases} g_{1,0} = 2 \cdot \frac{y'}{y} - 3 \cdot \frac{y''}{y'} + \frac{y^{(4)}}{y'''} \\ g_{1,1} = 2 \cdot \frac{y'}{y} - 4 \cdot \frac{y''}{y'} + 2 \cdot \frac{y'''}{y''} \\ g_{1,2} = 3 \cdot \frac{y'}{y} - 6 \cdot \frac{y''}{y'} + 3 \cdot \frac{y'''}{y''} \\ g_{1,3} = \frac{y'}{y} - 2 \cdot \frac{y''}{y'} + \frac{y'''}{y''} \end{cases} \qquad (3.22)$$

Using change (3.22) in (3.21) we have got:

$$g_{1,0}\cdot(y)^2\cdot(y')^{-3}\cdot y'''+b_0\cdot g_{1,1}\cdot(y)^2\cdot(y')^{-4}\cdot(y'')^2+b_1\cdot g_{1,2}\cdot(y)^3\cdot(y')^{-6}\cdot(y'')^3+b_2\cdot g_{1,3}\cdot y\cdot(y')^{-2}\cdot y''=0$$

$$\frac{g_{1,0}}{g_{1,3}}\cdot y\cdot(y')^{-1}\cdot(y'')^{-1}\cdot y'''+b_0\cdot\frac{g_{1,1}}{g_{1,3}}\cdot y\cdot(y')^{-2}\cdot y''+b_1\cdot\frac{g_{1,2}}{g_{1,3}}\cdot(y)^2\cdot(y')^{-4}\cdot(y'')^2+b_2=0 \quad (3.23)$$

Introducing change:

$$\begin{cases} g_{2,0}=g_{1,0}{}'\cdot g_{1,3}-g_{1,0}\cdot g_{1,3}{}'+g_{1,0}{}'\cdot g_{1,3}\cdot\left(\frac{y'}{y}-\frac{y''}{y'}-\frac{y'''}{y''}+\frac{y^{(4)}}{y'''}\right)\\ g_{2,1}=g_{1,1}{}'\cdot g_{1,3}-g_{1,1}\cdot g_{1,3}{}'+g_{1,1}\cdot g_{1,3}\cdot\left(\frac{y'}{y}-2\cdot\frac{y''}{y'}+\frac{y'''}{y''}\right)\\ g_{2,2}=g_{1,2}{}'\cdot g_{1,3}-g_{1,2}\cdot g_{1,3}{}'+g_{1,2}\cdot g_{1,3}\cdot\left(2\cdot\frac{y'}{y}-4\cdot\frac{y''}{y'}+2\cdot\frac{y'''}{y''}\right)\end{cases} \quad (3.24)$$

Using change (3.24) in (3.23) we have got:

$$g_{2,0}\cdot y\cdot(y')^{-1}\cdot(y'')^{-1}\cdot y'''+b_0\cdot g_{2,1}\cdot y\cdot(y')^{-2}\cdot y''+b_1\cdot g_{2,2}\cdot(y)^2\cdot(y')^{-4}\cdot(y'')^2=0$$

$$\frac{g_{2,0}}{g_{2,2}}\cdot(y)^{-1}\cdot(y')^3\cdot(y'')^{-3}\cdot y'''+b_0\cdot\frac{g_{2,1}}{g_{2,2}}\cdot(y)^{-1}\cdot(y')^2\cdot(y'')^{-1}+b_1=0 \quad (3.25)$$

Introducing change:

$$\begin{cases} g_{3,0}=g_{2,0}{}'\cdot g_{2,2}-g_{2,0}\cdot g_{2,2}{}'+g_{2,0}\cdot g_{2,2}\cdot\left(-\frac{y'}{y}+3\cdot\frac{y''}{y'}-3\cdot\frac{y'''}{y''}+\frac{y^{(4)}}{y'''}\right)\\ g_{3,1}=g_{2,1}{}'\cdot g_{2,2}-g_{2,1}\cdot g_{2,2}{}'+g_{2,1}\cdot g_{2,2}\cdot\left(-\frac{y'}{y}+2\cdot\frac{y''}{y'}-\frac{y'''}{y''}\right)\end{cases} \quad (3.26)$$

Using change (3.26) in (3.25) we have got:

$$g_{3,0}\cdot(y)^{-1}\cdot(y')^3\cdot(y'')^{-3}\cdot y'''+b_0\cdot g_{3,1}\cdot(y)^{-1}\cdot(y')^2\cdot(y'')^{-1}=0$$

$$\frac{g_{3,0}}{g_{3,1}}\cdot y'\cdot(y'')^{-2}\cdot y'''+b_0=0$$

$$g_{3,0}{}'\cdot g_{3,1}-g_{3,0}\cdot g_{3,1}{}'+g_{3,0}\cdot g_{3,1}\cdot\left(\frac{y''}{y'}-2\cdot\frac{y'''}{y''}+\frac{y^{(4)}}{y'''}\right)=0 \quad (3.27)$$

Thus, from (3.13) with unknown constants we have got differential equation (3.27) and recursively defined functions without inknown constants.

It should be noted that this transformation preserves the equivalence and is not a trick or erroneous information loss. Here uncertainty factors switches to the initial values uncertainty in the formulation of the Cauchy problem.

Thus, any differential equation with an arbitrary set of elementary functions, and their conpositions can be reduced to a differential equation of higher order, which is the sum of the differentials in the works of well-known integer powers and known integer scalar factors.

The complexity (entropy) of initial function is converted to the maximum order of differentiation on the stage (3.5), which in turn uniquely determine a differential equation on stage (3.12).

Forecasting

Thus, we have obtained a universal form of the differential equation for any function that has only uncertainty in the order of differentiation. This option is a numerical measure of the uncertainty (complexity) of the differential equation, that is, the minimum number of time series elements needed to define the function. It should be noted that the number of elements is not guaranteed to be sufficient, as the Cauchy problem clearly defining the uniqueness of solutions of differential equations where initial conditions defined separately for each individual derivative. This does not contradict the assertion of the finitude of necessary arguments submitted in section 2, as indeed some of the points may not lead to an increasing of entropy of function, satisfying earlier models found.

Time series forecasting is reduced to an iterative sequence of actions:

1. Take a premise about the number n of elements of the time series, forming the entropy function model on this iteration. For the first iteration it takes a value of 1, which is the minimum number of elements which can built time series forecasting model.
2. For a given value take a function-model $\sum_{i=1}^{n} a_i \cdot \prod_{j=1}^{n} \left(m^{(j)}\right)^{p_{i,j}} = 0$ where n defined on previous srtage, and transform this model by the reduction to a predetermined view to get the universal form $g\left(\left\{v^{(i)}\right\}\right) = 0$ with recursively defined functions. This equation defined bundle of function-models depended by starting conditions.
3. Using numerical methods check satisfying of defined function-models the acceptable deviation for all elements of time (2.1). If so, we will use this function-models set on the next step.
4. Use acceptable function-models to forecast time series by formuls defined in (2.7) or (2.9) depending on existence of intervals of acceptable deviation.
5. Go to the next iteration of the loop. In practice, the number of iterations will be determined by the necessary number of elements for the growth of entropy function models. Starting with a certain amount of elements numerical methods for solving differential equations will lead to functions similar to that found in the previous iteration.

This algorithm solves the problem in a general way. In practice due to the existence of possible special requirements of calculation limitations algorithm may be modified and simplified. However, the closer will be simplified algorithm to the above, the forecasting problem is solved more accurately, and the result is more objective.