SM Journal of Biometrics & Biostatistics ISSN: 2573-5470

Research Article © Nichelatti M. 2020

Forming the Differential Equations from Simple Polynomial Regression Models

Michele Nichelatti*

Service of Biostatistics, Niguarda Ca' Granda Hospital, Italy

Abstract

It is presented a method useful to form ordinary differential equations from nonlinear regression models of the type $y = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, a power series allowing to directly retrieving the differential model from raw data after fitting, to be compared with the differential model expected for the biological system which is studied. In particular for any possible value of n, the highest power at which the independent variable is raised, the paper gives the method to get the differential equation having the polynomial as solution. The use of power series allows some practical advantages when dealing with differential equations, and one of these - in some cases - is the capability of retrieving a function as solution of differential equation without having to know the specific rule to solve the differential equation itself.

Keywords: Polynomial regression models; Differential equations

Introduction

In some cases, to obtain an estimate of the behavior of dependent variable with respect to an independent one can be just the first task, and we would be much more interested in retrieving from the regression the differential equation governing the system [1]. For example, when dealing with a linear regression model given by the simple equation

$$y = a_1 x + a_0 \tag{1}$$

Where x is the independent variable, and y is the dependent one, $a_0, a_1 \neq 0$ being two arbitrary constants, the differential equation governing the system may be easily obtained deriving with respect to x both sides of equation (1)

$$\partial_x y = a_1 \tag{2}$$

so that, inserting equation (2) into equation (1) gives the differential equation

$$x\partial_x y - y + a_0 = 0 \tag{3}$$

having equation (1) as general solution. Equation (3) is a linear first order ordinary differential equation, also known as Clairaut's equation.

Equation (1) is the simpler form of a power series of the type $y = \sum_{k=0}^{+\infty} a_k x^k$, which is widely used in mathematics: for example, it can be used to solve the Malthus differential equation without utilizing any solution technique proper of differential equations; indeed, it is enough to know its first derivative $\partial_x y = \sum_{k=0}^{+\infty} k a_k x^{k-1}$ and the power series of the exponential function $e^x = \sum_{k=0}^{+\infty} \frac{x^k}{k!}$ [2]; in particular, the power series with a maximum finite power m (now a polynomial) shows both real and complex roots, depending on its possible factorization.

Polynomials are used in mathematics to approximate various kinds of function [3], but they are also used to model a wide range of phenomena in physics, social science and economics [4], as well as in biological ones, like in the systems governed by a mass action law or in the Lotka-Volterra predator-prey systems [5].

Differential Equations from Polynomial Models

The approach to polynomial univariate models is quite more complicated than those used for equation (1). Assuming a result of the type

$$y = a_2 x^2 + a_1 x \tag{4}$$

We will face with two arbitrary constants a_1, a_2 such that the first and second derivatives respectively are

$$\partial_x y = 2a_2 x + a_1 \tag{5}$$

$$\partial_{xx}^2 y = 2a_2 \,. \tag{6}$$

In this case, and in all next cases, we tacitly assume that y is continuous and derivable at least n times over its domain and that none of the coefficients are zero, so that we can eliminate the a_2 constant, since we have

$$a_2 = \frac{1}{2}\partial_{xx}^2 y \tag{7}$$

and therefore, first of equations (5) reads

Submitted: 30 December 2019 | **Accepted:** 16 January, 2020 | **Published:** 22 January, 2020

***Corresponding author:** Michele Nichelatti, Service of Biostatistics, Niguarda Ca' Granda Hospital, Milan, Italy, Email: michele.nichelatti@ospedaleniguarda.it

Citation: Nichelatti M (2020) Forming the Differential Equations from Simple Polynomial Regression Models. SM J Biometrics Biostat 5: 4. doi: https://dx.doi.org/10.36876/smjbb611439

$$\partial_x y = x\partial^2_{xx} y + a_1 \tag{8}$$

with

$$a_1 = -x\partial^2_{xx} y + \partial_x y \tag{9}$$

thus, inserting the values found for a_1, a_2 in equation (4), we obtain

$$\frac{1}{2}x^2\partial^2_{xx} y - x^2\partial^2_{xx} y + x\partial_x y - y = 0 \tag{10}$$

Hence

$$-\frac{1}{2}x^2\partial^2_{xx} y + x\partial_x y - y = 0 \tag{11}$$

equivalent to

$$x^2\partial^2_{xx} y - 2x\partial_x y + 2y = 0 \tag{12}$$

which is the differential equation having equation (4) as general solution.

Now, using the third-degree regression result

$$y = a_3x^3 + a_2x^2 + a_1x \tag{13}$$

and calculating the derivatives

$$\partial_x y = 3a_3x^2 + 2a_2x + a_1 \tag{14}$$

$$\partial^2_{xx} y = 6a_3x + 2a_2 \tag{15}$$

$$\partial^3_{xxx} y = 6a_3 \tag{16}$$

we get from equation (16)

$$a_3 = \frac{1}{6}\partial^3_{xxx} y \tag{17}$$

So that, inserting this value into equation (15) leads to

$$\partial^2_{xx} y = x\partial^3_{xxx} y + 2a_2 \tag{18}$$

giving

$$a_2 = \frac{1}{2}(\partial^2_{xx} y - x\partial^3_{xxx} y) \tag{19}$$

which, inserted into equation (14) gives

$$\partial_x y = -\frac{1}{2}x^2\partial^3_{xxx} y + x\partial^2_{xx} y + a_1 \tag{20}$$

and

$$a_1 = \frac{1}{2}x^2\partial^3_{xxx} y + x\partial^2_{xx} y + \partial_x y. \tag{21}$$

At this point, we can use the values of a_1, a_2 and a_3 respectively found in equations (21), (19) and (17), ad insert them into equation (13), to obtain

$$\begin{aligned} y &= \frac{1}{6}x^3\partial^3_{xxx} y + \frac{1}{2}x^2(\partial^2_{xx} y - x\partial^3_{xxx} y) + \\ &\quad \frac{1}{2}x(x^2\partial^3_{xxx} y - x\partial^2_{xx} y + \partial_x y) \\ &= \frac{1}{6}x^3\partial^3_{xxx} y - \frac{1}{2}x^2\partial^2_{xx} y + x\partial_x y \end{aligned} \tag{22}$$

thus, the differential equation is

$$x^3\partial^3_{xxx} y - 3x^2\partial^2_{xx} y + 6x\partial_x y - 6y = 0. \tag{23}$$

Now, let us take into accounts the fourth-degree regression model:

$$y = a_4x^4 + a_3x^3 + a_2x^2 + a_1x \tag{24}$$

with the usual procedure, we calculate

$$\partial_x y = 4a_4x^3 + 3a_3x^2 + 2a_2x + a_1 \tag{25}$$

$$\partial^2_{xx} y = 12a_4x^2 + 6a_3x + 2a_2 \tag{26}$$

$$\partial^3_{xxx} y = 24a_4x + 6a_3 \tag{27}$$

$$\partial^4_{xxxx} y = 24a_4. \tag{28}$$

Thus, after a little algebra, and avoiding intermediate steps:

$$a_4 = \frac{1}{24}\partial^4_{xxxx} y \tag{29}$$

$$a_3 = \frac{1}{6}(\partial^4_{xxxx} y - x\partial^3_{xxx} y) \tag{30}$$

$$a_2 = \frac{1}{4}(x^2\partial^4_{xxxx} y - 2x\partial^3_{xxx} y + 2\partial^2_{xx} y) \tag{31}$$

$$\begin{aligned} a_1 &= -\frac{1}{6}(x^3\partial^4_{xxxx} y - 3x^2\partial^3_{xxx} y + \\ &\quad 6x\partial^2_{xx} y - 6\partial_x y) \end{aligned} \tag{32}$$

hence, we obtain the differential equation having equation (24) as general solution

$$\begin{aligned} &x^4\partial^4_{xxxx} y - 4x^3\partial^3_{xxx} y + \\ &12x^2\partial^2_{xx} y - 24x\partial_x y + 24y = 0. \end{aligned} \tag{33}$$

A Possible Generalized Method

In finding the differential equation ruling the nonlinear regression model we observed an evident pattern, which can be invoked for whichever degree of the regression model we are dealing with.

Indeed, for the nonlinear regression model of unspecified degree m

$$\begin{aligned} y &= a_mx^m + a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \cdots + \\ &a_3x^3 + a_2x^2 + a_1x = \sum_{n=1}^{m} a_nx^n \end{aligned} \tag{34}$$

the differential equation is

$$\begin{aligned} &(-1)^{2m}\frac{m!}{m!}x^m\frac{\partial^m y}{\partial x^m} + (-1)^{2m-1}\frac{m!}{(m-1)!}x^{m-1}\frac{\partial^{m-1} y}{\partial x^{m-1}} + \\ &(-1)^{2m-2}\frac{m!}{(m-2)!}x^{m-2}\frac{\partial^{m-2} y}{\partial x^{m-2}} + \cdots + (-1)^{m+2}\frac{m!}{2!}x^2\frac{\partial^2 y}{\partial x^2} + \\ &(-1)^{m+1}\frac{m!}{1!}x^1\frac{\partial y}{\partial x} + (-1)^m\frac{m!}{0!}x^0 y = 0, \end{aligned} \tag{35}$$

but, since $\frac{m!}{m!}=1$, and since $\frac{m!}{1!}=\frac{m!}{0!}=m!$, to reduce the mathematical formalism in the differential equation (35), we can rewrite it as

$$x^m \frac{\partial^m y}{\partial x^m} + (-1)^{2m-1} \frac{m!}{(m-1)!} x^{m-1} \frac{\partial^{m-1} y}{\partial x^{m-1}} + (-1)^{2m-2} \frac{m!}{(m-2)!} x^{m-2} \frac{\partial^{m-2} y}{\partial x^{m-2}} + \cdots + (-1)^{m+2} \frac{m!}{2!} x^2 \frac{\partial^2 y}{\partial x^2} + (-1)^{m+1} m! x \frac{\partial y}{\partial x} + (-1)^m m! y = 0, \quad (36)$$

This model works for equations (33), (23) and (12), where m was respectively equal to 4, 3, and 2. We can now try to see what will occur if m = 5: in this case, equation (36) reads

$$x^5 \frac{\partial^5 y}{\partial x^5} - 5x^4 \frac{\partial^4 y}{\partial x^4} + 20x^3 \frac{\partial^3 y}{\partial x^3} - 60x^2 \frac{\partial^2 y}{\partial x^2} + 120x \frac{\partial y}{\partial x} - 120y = 0, \quad (37)$$

which is a Euler-Cauchy linear homogeneous differential equation, which can be solved at least with two methods. Using the simpler one, which is to assume that a solution will be of the form $y = x^k$ (in which $k > n$ is an arbitrary integer), and imposing this substitution in equation (37), where we also assume $x \neq 0$, we will find that, in general,

$$\frac{\partial^n x^k}{\partial x^n} = k(k-1)(k-2)\cdots(k-n+1)x^{k-n} \quad (38)$$

so that we obtain the polynomial in k

$$k^5 - 15k^4 + 85k^3 - 225k^2 + 274k - 120 = (k-1)(k-2)(k-3)(k-4)(k-5) = 0 \quad (39)$$

and the individual solutions $y_1 = a_1 x$, $y_2 = a_2 x^2$, $y_3 = a_3 x^3$, $y_4 = a_4 x^4$, and $y_5 = a_5 x^5$, so that the solution of the differential equation (37) is

$$y = y_5 + y_4 + y_3 + y_2 + y_1 = a_5 x^5 + a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x \quad (40)$$

exactly corresponding to what expected.

If regression model contains a constant term, so that

$$y = \sum_{n=0}^{m} a_n x^n \quad (41)$$

then, if we assume $a_0 = k$, it is immediate to see that the differential equation forming from the model given by equation (41) is

$$x^m \frac{\partial^m y}{\partial x^m} + (-1)^{2m-1} \frac{m!}{(m-1)!} x^{m-1} \frac{\partial^{m-1} y}{\partial x^{m-1}} + (-1)^{2m-2} \frac{m!}{(m-2)!} x^{m-2} \frac{\partial^{m-2} y}{\partial x^{m-2}} + \cdots + (-1)^{m+2} \frac{m!}{2!} x^2 \frac{\partial^2 y}{\partial x^2} + (-1)^{m+1} m! x \frac{\partial y}{\partial x} + (-1)^m m! y + (-1)^{m-1} m! a_0 = 0, \quad (42)$$

For example, if the regression model is

$$y = a_5 x^5 + a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + 1 \quad (43)$$

the differential equation is

$$x^5 \frac{\partial^5 y}{\partial x^5} - 5x^4 \frac{\partial^4 y}{\partial x^4} + 20x^3 \frac{\partial^3 y}{\partial x^3} - 60x^2 \frac{\partial^2 y}{\partial x^2} + 120x \frac{\partial y}{\partial x} - 120y + 120 = 0. \quad (44)$$

Discussion

When dealing with the analysis of a biological system, usually one expects to apply a given dynamical model [6] to the data. Here, the data are used after fitting to retrieve the ordinary differential equation having the fitted equation as general solution: however, despite the differential model arising from the regression, we must point out that a nonlinear autonomous differential equation of the type

$$y\,\partial_x y\,\partial_{xx}^2 y\,\partial_{xxx}^3 y \cdots \partial_{x\ldots x}^{n-2} y\,\partial_{x\ldots x}^{n} y\,\partial_{x\ldots x}^{n+1} y = 0 \quad (45)$$

has a solution (e.g., the highest grade one) equal to

$$y_n(x) = c_n x^n + c_{n-1} x^{n-1} + c_{n-2} x^{n-2} + \cdots + c_2 x^2 + c_1 x + c_0, \quad (46)$$

which is equivalent to equation (41), while other solutions are $y_{-1}(x) = 0$, $y_0(x) = c_0$, $y_1(x) = c_1 x + c_0$, $y_2(x) = c_2 x^2 + c_1 x + c_0$, $y_3(x) = c_3 x^3 + c_2 x^2 + c_1 x + c_0$,..., and so on; indeed, from equation (45) we see that

$$y = \partial_x y = \partial_{xx}^2 y = \partial_{xxx}^3 y = \cdots = \partial_{x\ldots x}^{n-2} y = \partial_{x\ldots x}^{n} y = \partial_{x\ldots x}^{n+1} y = 0 \quad (47)$$

e.g., all are solutions of equation (45).

However, in this paper, the case is limited to a simple nonlinear univariable model, but the differential equation obtained “from scratch” can be used to be compared to the differential equation one expects to rule the system.

Conclusion

When dealing with the analysis of a biological system, usually one expects to apply a given dynamical model [3] to the data. In this paper, the data are used after fitting to retrieve the ordinary differential equation having the fitted equation as general solution. It is mandatory to verify if the differential model retrieved from the fitted regression result may actually apply to the system one is studying, so that it is always advisable to look at possible differential models expected for the, system, based on its expected behavior, and this may include any kind of possibilities (for example, if one should expect any oscillatory behavior), so that the expected potential differential behavior may be compared with the differential equation arising from possible regression models.

References

1. Taubes CH. Modeling differential equations in biology. Prentice Hall, Upper Saddle River, USA. 2001.

2. Fässler A. Fast track to differential equations. Springer Nature Switzerland, Cham, CH. 2019.

3. Timan AF. Theory of approximation of functions of a real variable. McMillan, New York, USA. 1963.

4. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical Recipes: The Art of Scientific Computing, 3rd Ed. Cambridge University Press, New York, USA. 2007.

5. Szederkenyi G, Magyar A, Hangos K. Analysis and control of polynomial dynamic models with biological applications. Academic Press, London, UK. 2018.

6. Gros C. Complex and adaptive dynamical systems, 4th Ed. Springer International Publishing, Cham CH. 2015.