

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Automation 2024: Advances in Automation, Robotics and Measurement Techniques	
Series Title		
Chapter Title	Application of the Polynomial Maximization Method for Estimating Nonlinear Regression Parameters with Non-gaussian Asymmetric Errors	
Copyright Year	2025	
Copyright HolderName	The Author(s), under exclusive license to Springer Nature Switzerland AG	
Corresponding Author	Family Name Particle Given Name Prefix Suffix Role Division Organization Address Email	Zabolotnii Serhii Cherkasy State Business College Cherkasy, Ukraine zabolotniua@gmail.com
Author	Family Name Particle Given Name Prefix Suffix Role Division Organization Address Division Organization Address Email	Tkachenko Oleksandr Uzhhorod National University Uzhhorod, Ukraine Ivan Franko National University of Lviv Lviv, Ukraine tkachenko.ck@gmail.com
Author	Family Name Particle Given Name Prefix Suffix Role Division Organization Address Email	Nowakowski Waldemar Casimir Pulaski Radom University Radom, Poland w.nowakowski@uthrad.pl
Author	Family Name Particle	Warsza

Given Name **Zygmunt L.**
Prefix
Suffix
Role
Division
Organization Polish Metrological Society
Address Radom, Poland
Email zlw1936@gmail.com

Abstract	In the article, an alternative approach to estimating parameters in nonlinear regression models under asymmetric error distributions is examined. A novel approach for adaptive estimation is proposed, which is based on the use of second-order polynomial functions. This enables a straightforward implementation to account for deviations from Gaussian idealization in the form of moments up to the fourth order. It is demonstrated that the overall problem can algorithmically be reduced to the numerical solution of a system of nonlinear stochastic equations. Analytical expressions are obtained, which facilitate the estimation of parameters and the analysis of their asymptotic variance. Statistical modeling using the Monte Carlo method was conducted, and the results indicate that the accuracy of PMM2 estimates is comparable to SLS estimates and significantly so exceeds the accuracy of OLS estimates.
Keywords (separated by '-')	Polynomial Maximization Method - non-linear regression models - asymmetric error distributions - Monte Carlo method - higher order statistics



Application of the Polynomial Maximization Method for Estimating Nonlinear Regression Parameters with Non-gaussian Asymmetric Errors

Serhii Zabolotnii¹(✉), Oleksandr Tkachenko^{2,3}, Waldemar Nowakowski⁴, and Zygmunt L. Warsza⁵

¹ Cherkasy State Business College, Cherkasy, Ukraine

zabolotniua@gmail.com

² Uzhhorod National University, Uzhhorod, Ukraine

³ Ivan Franko National University of Lviv, Lviv, Ukraine

⁴ Casimir Pulaski Radom University, Radom, Poland

w.nowakowski@uthrad.pl

⁵ Polish Metrological Society, Radom, Poland

Abstract. In the article, an alternative approach to estimating parameters in nonlinear regression models under asymmetric error distributions is examined. A novel approach for adaptive estimation is proposed, which is based on the use of second-order polynomial functions. This enables a straightforward implementation to account for deviations from Gaussian idealization in the form of moments up to the fourth order. It is demonstrated that the overall problem can algorithmically be reduced to the numerical solution of a system of nonlinear stochastic equations. Analytical expressions are obtained, which facilitate the estimation of parameters and the analysis of their asymptotic variance. Statistical modeling using the Monte Carlo method was conducted, and the results indicate that the accuracy of PMM2 estimates is comparable to SLS estimates and significantly so exceeds the accuracy of OLS estimates.

AQ1

AQ2

Keywords: Polynomial Maximization Method · non-linear regression models · asymmetric error distributions · Monte Carlo method · higher order statistics

1 Introduction

In the contemporary realm of statistical analysis, significant attention is devoted to the study of complex interrelationships among various variables. Consequently, the development of reliable methods for parameter estimation in nonlinear regression models, particularly in the context of complex and realistic data, is of great importance. The predictability and accuracy of such models are critically vital in many scientific and applied fields, including engineering, ecology, biomedicine, economics, and finance. In these and other applied disciplines, regression analysis has made a substantial contribution to understanding and addressing numerous real-world problems [1–5].

However, traditional estimation methods, especially the widely used ordinary least squares (OLS) approach, which are based on the assumption of Gaussian statistical data, can be inefficient or inadequately precise in conditions where the error distribution is asymmetric or has heavy tails. This presents a challenge for researchers to develop new approaches that better accommodate these conditions.

In response to these challenges, contemporary researchers are developing and implementing novel strategies. One of the most successful is based on a modification of the least squares method, first utilized by Wang [6] to address the problem of measurement error in nonlinear regression models. This modification, known as the “Second order” Least Squares (SLS), has been successfully applied for estimating parameters in nonlinear regression with asymmetric errors [7]. Subsequently, several authors have explored the asymptotic efficiency of this approach [8], developed a series of new optimization criteria [9, 10], and applied it to solve practical problems [11–13] and is possible to use also in the automatic diagnostics of the railway tracks surface [27].

AQ3

2 Purpose of the Study

In this article, we explore another approach to solving the problem of parameter estimation in nonlinear regression, based on the use of the Polynomial Maximization Method (PMM). This novel statistical estimation method is founded on the framework of Kunchenko stochastic polynomials [14]. Its application opens new possibilities for constructing sub-optimal methods for processing non-Gaussian data. Beyond parameter estimation, this can include tasks such as testing statistical hypotheses [15], pattern recognition (“matching template”) [16], posterior [17] and online [18] detection of change points (problem of change point).

This study continues a series of works examining the application of the quadratic variant (at degree $S = 2$) of the polynomial maximization method for estimating various linear-parameter models (linear regression [19], polynomial regression [20], autoregressive [21], and moving average [22]) under the condition of asymmetric distribution of their random components and using moment-cumulant description. One of the main results of these works is the fact that in finite samples, PMM2 significantly outperforms the accuracy of OLS. Only when the non-Gaussian regression errors acquire a property of distribution symmetry (or at least with a zero of 3rd moment) do the PMM2 estimates degenerate into OLS estimates. Moreover, as the results of simulation modeling in work [7] demonstrate, a similar effect is observed for SLS estimates.

Thus, the aim of this work is to synthesize an algorithm for estimating the parameters of nonlinear regression based on the use of the polynomial maximization method at degree $S = 2$; to analyze the properties of the resulting PMM2 estimates and to conduct a comparative analysis of their efficiency against OLS and SLS estimates through Monte Carlo statistical modeling.

3 General Statement of the Problem

Suppose there is a regression model of observations in the form of the sum of deterministic and random components:

$$y_v = R_v(\boldsymbol{\theta}, \mathbf{X}) + \xi_v, v = \overrightarrow{1, N}, \quad (1)$$

describing the dependence of the values of the target variable $\mathbf{Y} = \{y_1, y_2, \dots, y_N\}$ on a set of independent variables \mathbf{X} of independent variables (P-factor regressor).

The deterministic component $R_v(\theta, \mathbf{X})$ is a nonlinear function that depends on the Q -dimensional vector of informative parameters $\theta = \{a_0, a_1, \dots, a_{Q-1}\}$.

The samples of the random component (error) of the regression model ξ_v are a sequence of independent and equally distributed random variables that have the property of centeredness ($E\{\xi\} = 0$) and homoscedasticity $D\{\xi\} = \text{const}$. The probabilistic properties of the regression model error differ significantly from the Gaussian (normal) law. The general task is to find estimates of the components of the informative parameter vector θ on the basis of statistical analysis of statistical data \mathbf{Y} and \mathbf{X} .

4 Application of the PMM in Finding Estimates of a Vector Parameter with Unequally Distributed Data

To solve the problem of regression analysis, one can use a modification of the PMM for statistical estimation of a vector parameter with unequally distributed data [14]. It is based on the property of maximizing a functional in the form of a stochastic polynomial of the general form:

$$L_{SN} = \sum_{v=1}^N \sum_{i=1}^S \phi_i(y_v) \int_a k_{iv}(a) dz - \sum_{i=1}^S \sum_{v=1}^N \int_a \Psi_{iv} k_{iv}(a) dz, \quad (2)$$

It is assumed that random variables y_v are described by a sequence of mathematical expectations:

$$\Psi_{iv} = E\{\phi_i(y_v)\}, i = \overrightarrow{1, S}, v = \overrightarrow{1, N}, \quad (3)$$

and are twice differentiated by the parameter a that is being evaluated.

A stochastic polynomial L_{SN} of the form (2) has two main properties:

- 1) for any order S with asymptotic growth $v \rightarrow \infty$ of the sample size \mathbf{Y} polynomial L_{SN} as a function of the parameter a takes a maximum in the neighborhood of the true value of this parameter,
- 2) with different samples \mathbf{Y} the deviation of the maximum of the polynomial L_{SN} from the true value of the parameter a will have a minimum variance (for the corresponding order of the polynomial S).

By analogy with the maximum likelihood method, the parameter a can be found from solving the equation of the form:

$$\frac{d}{da} L_{SN} \Big|_{a=\hat{a}} = \sum_{i=1}^S \sum_{v=0}^N k_{iv} [\phi_i(y_v) - \Psi_{iv}] \Big|_{a=\hat{a}} = 0, \quad (4)$$

The optimal coefficients k_{iv} maximizing the functional (2) are found from solving the system of linear algebraic equations:

$$\sum_{j=1}^S k_{jv} F_{(i,j)v} = \frac{d}{da} \Psi_{iv}, i = \overrightarrow{1, S}, v = \overrightarrow{1, N}, \quad (5)$$

where: $F_{(i,j)v} = \Psi_{(i,j)v} - \Psi_{iv}\Psi_{jv}$, $\Psi_{(i,j)v} = E\{\phi_i(y_v)\phi_j(y_v)\}$, $i, j = \overrightarrow{1, S}$.

This approach to estimation can be extended to the case of finding estimates of a vector parameter $\theta = \{a_0, a_1, \dots, a_{Q-1}\}$. In this, it is necessary to use Q polynomials $L_{SN}^{(p)}$, $p = \overrightarrow{0, Q-1}$ of the general form (2) for each component a_p of vector parameter.

Each p -stochastic polynomial $L_{SN}^{(p)}$ as a function of the parameter a_p at known values of other components of the vector θ at $n \rightarrow \infty$ also has a maximum in the neighborhood of the true value of the parameter a_p . Thus, the desired parameter estimates can be found as a solution to the system of equations of the form:

$$f_{SN}^{(p)}(y_v, x_v) = \sum_{i=1}^S \sum_{v=1}^N k_{iv}^{(p)} [\phi_i(y_v) - \Psi_{iv}] \Big|_{a_p=\hat{a}_p} = 0, p = \overrightarrow{0, Q-1}. \quad (6)$$

When using polynomials of degree $S \geq 2$ finding the PMM estimates of a vector parameter in the vast majority of cases (similar to the situation with MLE) requires the use of numerical methods for solving systems of nonlinear equations. In particular, an approach based on the iterative Newton-Raphson numerical procedure is often used. It is based on the principle of linearization by expanding the left-hand side of each nonlinear equation of system (6) into a Taylor series in the neighborhood of the true value of the vector θ . Limiting ourselves to the first two terms of the series, we can write the linear system in matrix form:

$$\hat{\theta}_{(S)}^{(k+1)} = \hat{\theta}_{(S)}^{(k)} - \left[\mathbf{Z}_S(Y/\hat{\theta}^{(k)}) \right]^{-1} \mathbf{F}_S(Y/\hat{\theta}^{(k)}), \quad (7)$$

which can be used to iteratively search for estimated values. To obtain system (8), it is necessary to calculate the column matrix $\mathbf{F}_S(Y/\hat{\theta}^{(k)})$ composed of the elements of the left-hand side of each nonlinear equation of system (6) and a square matrix $\mathbf{Z}_S(Y/\hat{\theta}^{(k)}) = \mathbf{H}_S(Y/\hat{\theta}^{(k)}) + \mathbf{J}_S(\hat{\theta}^{(k)})$ with the components

$$H_{SN}^{(p,q)} = \sum_{i=1}^S \frac{\partial}{\partial a_q} k_{iv}^{(p)} \left[\sum_{v=1}^N \phi_i(y_v) - N\Psi_{iv} \right] \quad (8)$$

and elements of the matrix of the amount of information extracted $\mathbf{J}_S(\hat{\theta}^{(k)})$ type:

$$J_{SN}^{(p,q)} = \sum_{v=1}^N \sum_{i=1}^S k_{i,v}^{(p)} \frac{\partial}{\partial a_q} \Psi_{iv}, p, q = \overrightarrow{0, Q-1}. \quad (9)$$

To start the iterative procedure, it is assumed that there is some initial approximation $\theta^{(1)} = \{a_0^{(1)}, a_1^{(1)}, \dots, a_{Q-1}^{(1)}\}^T$ which can be chosen as a “rough” estimate found by a simpler method, such as the OLS [14].

5 Finding PMM Estimates of Nonlinear Regression Parameters

5.1 Application of the PMM at Degree S = 1

To solve the problem using the method of polynomial maximization in the formation of the stochastic polynomial (2), we use the power transformations of the sample values as basic functions Y , i.e.

$$\phi_i(y_v) = y_v^i. \quad (10)$$

In this case, the sequence of mathematical expectations (3) is a set of initial moments of the corresponding order:

$$\Psi_{iv} = E\{y_v^i\} = \alpha_{iv}, i = \overrightarrow{1, S}, v = \overrightarrow{1, N}, \quad (11)$$

and the functions $F_{(i,j)v} = \alpha_{(i+j)v} - \alpha_{iv}\alpha_{jv}$.

Thus, when using a stochastic polynomial of order $S = 1$ PMM, the estimates of the elements of the vector parameter θ of the deterministic component of the nonlinear regression model (1) can be found from the solution of the system Q of equations:

$$\sum_{v=1}^N \left\{ k_{1,v}^{(p)} [y_v - R_v(\theta, X)] \right\} = 0, p = \overrightarrow{0, Q-1}. \quad (12)$$

where the optimal coefficients $k_{1,v}^{(p)}$ are found as solutions to a system of linear algebraic equations, and which can be represented as

$$k_{1,v}^{(p)} = \frac{1}{\mu_2} \frac{\partial}{\partial a_p} R_v(\theta, X), p = \overrightarrow{0, Q-1} \quad (13)$$

We use an assumption similar to the one used in the conventional OLS for estimating the parameters of a nonlinear regression [10]. It consists in the fact that the deterministic component $R_v(\theta, X)$ of the nonlinear model (1) can be approximated by a linear function based on the first-order Taylor series:

$$R_v(\theta^{(k)}, X) + \sum_p \frac{\partial}{\partial a_p} R_v(\theta^{(k)}, X) [\theta - \theta^{(k)}] \approx R_v(\theta^{(k)}, X) \sum_p H_{v,p} \Delta \theta, \quad (14)$$

where $H_{v,p}$ – are the elements of the Jacobi matrix H that depend on the components a_p of the vector parameters θ at fixed values of the regressors X .

$$\sum_v R_v(\theta, X) \frac{\partial}{\partial a_p} R_v(\theta, X) = 0, p = \overrightarrow{1, Q}. \quad (15)$$

Then, considering (13), after certain transformations, we can represent (12) in matrix form as a system of linear equations:

$$(H^T H) \Delta \theta^T \approx \frac{1}{N} \sum_{v=1}^N H^T \Delta Y^T, \quad (16)$$

which is equivalent to the corresponding least squares system [23]. Such OLS estimates are obtained by using appropriate numerical iterative procedures (gradient descent, Newton-Hauss method, etc.). They may have a certain bias, but they can be used as a first approximation for more complex estimation methods, in particular, the OLS at higher polynomial degrees.

5.2 Application of the PMM at Degree S = 2

Using the general formula (6), we present a system of equations for finding the PMM estimates $\hat{\theta}_{(2)}$ of the deterministic component of the nonlinear regression model (1):

$$\sum_{v=1}^N \left\{ k_{1,v}^{(p)} [y_v - R_v(\theta, X)] + k_{2,v}^{(p)} \left[(y_v)^2 - \left[(R_v(\theta, X))^2 + \mu_2 \right] \right] \right\} = 0 \quad (17)$$

where, $p = \overrightarrow{0, Q-1}$, and the optimal coefficients $k_{i,v}^{(p)}$, $i = \overrightarrow{1, 2}$ can be written as:

$$k_{1,v}^{(p)} = \frac{\mu_4 - \mu_2^2 + 2\mu_3 R_v(\theta, X)}{\mu_2(\mu_4 - \mu_2^2) - \mu_3^2} \frac{\partial}{\partial a_p} R_v(\theta, X), \quad k_{2,v}^{(p)} = -\frac{\mu_3^2}{\mu_2(\mu_4 - \mu_2^2) - \mu_3^2} \frac{\partial}{\partial a_p} R_v(\theta, X). \quad (18)$$

Substituting the coefficients of (18) into (17), after certain transformations, the system of equations for finding the estimates can be written in the form:

$$\sum_{v=1}^N \left\{ \frac{\partial}{\partial a_p} R_v(\theta, X) \left[A_2(R_v(\theta, X))^2 + B_{2,v} R_v(\theta, X) + C_{2,v} \right] \right\} = 0, \quad p = \overrightarrow{0, Q-1} \quad (19)$$

where:

$$A_2 = \mu_3, \quad B_{2,v} = \mu_4 - \mu_2^2 - y_v \mu_3, \quad C_{2,v} = y_v^2 \mu_3 - y_v (\mu_4 - \mu_2^2) - \mu_2 \mu_3. \quad (20)$$

The analysis of expressions (20) shows that in the case of symmetry of distribution (or at least identity $\mu_3 = 0$) of the random component of errors of the regression model, the use of the PMM at $S = 2$ is inappropriate, since system (19) degenerates into a system of linear equations of the form (13), equivalent to the system for a conventional OLS. The elements of the matrix $Z_2(Y/\hat{\theta}^{(k)})$ necessary for the numerical solution of system (19) can be represented as follows:

$$Z_{2N}^{(p,q)} = \sum_{v=1}^N \sum_{v=1}^N \left\{ \frac{\partial}{\partial a_p} R_v(\theta, X) \frac{\partial}{\partial a_q} R_v(\theta, X) [2A_2 R_v(\theta, X) + B_{2,v}] \right\} \quad (21)$$

Obviously, when the degree of the polynomial $S = 2$ PMM estimates can only be found numerically, for example, using the Newton-Raphson iterative procedure. As an initial approximation, we can use the ODE estimates.

6 Analyzing the Theoretical Efficiency of the PMM Estimates of Nonlinear Regression Parameters

It is known that one of the main criteria for the effectiveness of estimates of informative parameters is their variances. It has been shown [14] that estimates of vector parameters found using the method of maximizing polynomials generally have the property of correctness and are asymptotically unbiased. To obtain the analytical expressions describing the variances of the PMM estimates, we use the matrix of the amount of information extracted $J_S(\hat{\theta}^{(k)})$ with elements of the form (9). In the moment description, such a matrix consists of elements:

$$J_{SN}^{(p,q)} = \sum_{v=1}^N \sum_{i=1}^S k_{i,v}^{(p)} \frac{\partial}{\partial a_q} \alpha_{iv}, \quad p, q = \overrightarrow{0, Q-1}. \quad (22)$$

In the statistical sense, the amount of extracted information is conceptually close to the amount of Fisher information. It is shown that the variances $\sigma_{(a_p)S}^2$ of the PMM estimates of the components of the vector parameter $\boldsymbol{\theta}$ in the asymptotic case (at $N \rightarrow \infty$) can be obtained as the elements of the main diagonal of the variation matrix, which is the inverse of the matrix composed of the elements:

$$\mathbf{V}_{PMM S}(\boldsymbol{\theta}) = \begin{bmatrix} J_{SN}^{(1,q)} & J_{SN}^{(2,q)} & \dots & J_{SN}^{(1,q)} \\ J_{SN}^{(2,1)} & J_{SN}^{(2,2)} & \dots & J_{SN}^{(2,q)} \\ \dots & \dots & \dots & \dots \\ J_{SN}^{(p,1)} & J_{SN}^{(p,2)} & \dots & J_{SN}^{(p,q)} \end{bmatrix}^{-1} \quad (23)$$

Another important property of the PMM is that when the number of S of the stochastic polynomial, the variance of the estimates decreases, since the amount of extracted information asymptotically (at $S \rightarrow \infty$) tends to the Fisher information [14].

Using the obtained relations (13) for the linear variant of the polynomial maximization method (with degree $S = 1$), we can write the expression for the amount of information extracted in the form:

$$J_{1N}^{(p,q)} = \frac{1}{\mu_2} \sum_{v=1}^N \left[\frac{\partial}{\partial a_p} R_v(\boldsymbol{\theta}, \mathbf{X}) \frac{\partial}{\partial a_q} R_v(\boldsymbol{\theta}, \mathbf{X}) \right], p, q = \overline{0, Q-1}, \quad (24)$$

where μ_2 - is the 2nd order central moment (variance) of random errors ξ_v of the corresponding regression model.

Since the expressions for finding the OLS estimates of the parameters of nonlinear regression models at degree $S = 1$ are equivalent to the OLS estimates, the accuracy of the estimates of both methods is obviously the same. Thus, the expressions (24) defining the accuracy of the OLS estimators at $S = 1$ can be used for a comparative analysis of the efficiency of the OLS estimators obtained using polynomials of higher order. Using expressions (18) for the coefficients $k_{i,v}^{(p)}$, $i = \overline{1, 2}$ which minimize the variance of the estimates of the components of the desired parameter at degree $S = 2$, we obtain the corresponding expressions for the amount of information extracted:

$$J_{2N}^{(p,q)} = \frac{\mu_2^3 - \mu_2\mu_4}{\mu_2(\mu_2^3 - \mu_2\mu_4 + \mu_3^2)} \sum_{v=1}^N \left[\frac{\partial}{\partial a_p} R_v(\boldsymbol{\theta}, \mathbf{X}) \frac{\partial}{\partial a_q} R_v(\boldsymbol{\theta}, \mathbf{X}) \right], p, q = \overline{0, Q-1}, \quad (25)$$

where: μ_2, μ_3, μ_4 - are the central moments of random errors ξ_v of the regression model.

A comparative analysis of expressions (24) and (25) allows us to write down the relations for the corresponding variation matrices containing the asymptotic variances of the parameter estimates in the form:

$$\mathbf{V}_{(PMM 2)} = g_{\frac{PMM 2}{PMM 1}} \mathbf{V}_{(PMM 1)} = g_{\frac{PMM 2}{OLS}} \mathbf{V}_{(PMM 1)}. \quad (26)$$

where the coefficient of variance reduction of estimates, which is the same for all components of the vector parameter $\boldsymbol{\theta} = \{a_0, a_1, \dots, a_{Q-1}\}$ can be represented as:

$$g_{\frac{PMM 2}{OLS}} = \frac{\mu_2^3 - \mu_2\mu_4}{\mu_2(\mu_2^3 - \mu_2\mu_4 + \mu_3^2)} = 1 - \frac{\gamma_3^2}{2 + \gamma_4}. \quad (27)$$

The transition in expression (27) from a momentary to a cumulative description is that the deviation of the values of the cumulative coefficients of higher orders γ_r due to the fact from zero shows the degree of difference from the Gaussian distribution. In addition, cumulative coefficients of higher orders cannot take arbitrary values. In particular, for the cumulative skewness and kurtosis coefficients, the following inequality is known $\gamma_4 + 2 \geq \gamma_3^2$. Considering this inequality, based on the analysis of (3.6), we can conclude that the coefficient of variance reduction g_{PMM2}^{OLS} is a dimensionless value belonging to the range $(0; 1]$. In general, it can be stated that the relative efficiency of the quadratic version of PMM depends on the degree of difference between the distribution of regression errors and the Gaussian law, which is numerically expressed by the difference in the absolute value (module) of the skewness coefficient γ_3 . With its increase, the magnitude of the relative decrease in variance can be quite significant and even asymptotically tend to zero when the absolute value of the skewness coefficient approaches the boundary of the region of acceptable values, i.e. $|\gamma_3| \rightarrow \sqrt{2} + \gamma_4$.

7 Adaptive Estimation of Regression Parameters Using the PMM

It is obvious that the expediency of using the quadratic version (at $S = 2$) of the PMM is only appropriate when the distribution of the random component of regression models differs from the Gaussian law and its asymmetry. We should also note the fact, which is important from a practical point of view, that in order to obtain the PMM estimates, additional a priori information about the error properties of the regression model is required, which is usually absent in real conditions.

It is understood that an adaptive approach can be applied under the assumption that model errors are independent of the regressors [24]. Adaptability is understood in the sense that a number of successive refinement steps are applied. The first step is to estimate the parameters of the deterministic component of the regression model using a conventional OLS that does not consider the specifics of the probability distribution of errors. After removing the deterministic component, the type of random component of the regression model is identified and its parameters are estimated. At the third stage, the refined (adaptive) estimates of the informative parameters of the regression model are found, considering the probabilistic properties of the errors. This approach is based on the property that the inadequacy of the Gaussian model is not a critical factor from the point of view that the OLS estimates of the regression parameters still remain unbiased and valid, although they are no longer optimal. And since the least squares method is inherently linear, the probabilistic properties of the regression residuals are virtually indistinguishable from those of the original random component of the regression model.

We recall the above-mentioned fact that the task of finding the estimated values of higher-order statistics is methodologically much simpler than the task of identifying and estimating the parameters of the distribution density. In this sense, the task of overcoming a priori uncertainty about the probability characteristics of the model is much simpler than the approach based on maximum likelihood. In addition, the degree of uncertainty in the estimates of the moments (depending on their order) may also be less than the uncertainty in the estimates of the parameters of the distribution density. And as will be shown in Sect. 4, this factor can lead to an additional improvement (relative to the

PMM) of the resulting adaptive PMM estimates of the informative regression parameters. It should be noted that the obtained a posteriori estimates of higher-order statistics (in particular, cumulative skewness and kurtosis coefficients) can be used to test the hypothesis of Gaussian distribution and symmetry of regression errors based on Jarque -Bera statistical tests [25].

8 Static Modeling

To verify the theoretical results, we modified a set of functions written in R that implemented a Monte Carlo simulation procedure for finding estimates of linear regression, autoregressive, and moving average models using the polynomial maximization method [19–22]. Now, in addition, it allows for a comparative analysis of the accuracy of three methods (PMM2, SLS, and OLS) when estimating parameters of nonlinear regression models with random component of a non-Gaussian asymmetric distribution.

To find the estimates of the ordinary least squares (OLS) method, the built-in *nls* function of nonlinear regression analysis is used. The implementation of the determination of SLS estimates by minimizing the corresponding functional is based on the use of functions of the *optimx* library containing various quasi-Newton optimization functions, in particular, “BFGS” (Broyden-Fletcher-Goldfarb-Shanno). To solve the systems of nonlinear equations that occur when finding PMM2 estimates, the library of functions *nleqslv* was used, which implements numerical Broyden or Newton methods with a choice of global strategies, such as line search and confidence region [26].

In this section, we present the results of a comparative analysis using two types of nonlinear regression models, which were also studied in [7]. The first regression model has an exponential deterministic function

$$y_v = \theta_1 \exp(\theta_2 X) + \xi_v, v = \overrightarrow{1, N} \quad (28)$$

with the true values of the parameters $\theta_1 = 10$ i $\theta_2 = -0.6$;

and the second is a model of species growth:

$$y_v = \frac{\theta_1}{[1 + \exp(\theta_2 + \theta_3 x)]} + \xi_v, v = \overrightarrow{1, N} \quad (29)$$

with the true values of the parameters $\theta_1 = 10$, $\theta_2 = 1.5$ and $\theta_3 = -0.8$. The model of the independent variable is used $X \sim \text{Uniform}(0,20)$.

The centered chi-squared distribution $\xi = (\chi(3) - 3)/\sqrt{3}$ with the scale parameter $\sigma^2 = 2$ was used as the error model. For such a random variable, the degree of non-Gaussianity can be numerically described by the values of the skewness coefficients $\gamma_3 = 1.4$ and kurtosis $\gamma_4 = 2$. According to (27), the expected value of the decrease in the variance of estimates will be $g_{\frac{\text{PMM2}}{\text{OLS}}} \approx 0.56$. The mathematical formulation of the estimation assumes that the stochastic component has a significantly different asymmetric distribution from the Gaussian, but no a priori information about its parameters. Three criteria are used to compare the effectiveness of different methods: the mean of the estimates (Mean), their variance (Var), and the standard deviation from the true value (MSE). By repeated (for $m = 1000$) Monte Carlo tests with different sample sizes $n =$

30, 50, 100, 200, the results of, the totality of which is presented in Tables 1 and 2. The data presented in Tables 1 and 2 generally correlate with the analogous results of work [7]. They demonstrate that for all the analyzed methods, the resulting estimates are asymptotically unbiased, as their averaged (Mean) values asymptotically (as $n \rightarrow \infty$) converge to the true values of the informative parameters. Furthermore, there is a significant advantage in efficiency (reduction in the magnitude of variance and MSE) of the SLS and PMM2 estimates compared to OLS estimates.

This is particularly illustrated by the bar charts (constructed based on the data from Tables 1 and 2) presented in Fig. 1. The charts depict the dynamics of the ratios $\frac{MSE_{SLS}}{MSE_{OLS}}$ and $\frac{MSE_{PMM2}}{MSE_{OLS}}$ and for the estimates of various parameters of the studied regression models, depending on the sample size of statistical data.

AQ4

Table 1. Results of statistical modeling of the regression model parameters estimation $Y = \theta_1 \exp(\theta_2 X) + \xi$

n		OLS			SLS			PMM2		
		Mean	Var	MSE	Mean	Var	MSE	Mean	Var	MSE
200	$\theta_1 = 10$	9.9099	0.3128	0.3205	9.9835	0.1847	0.1847	9.9845	0.1873	0.1873
	$\theta_2 = -0.6$	-0.5955	0.0026	0.0026	-0.6001	0.0016	0.0016	-0.5997	0.0016	0.0016
100	$\theta_1 = 10$	9.8506	0.6514	0.6729	9.9795	0.4136	0.4135	9.9816	0.4142	0.4140
	$\theta_2 = -0.6$	-0.5959	0.0055	0.0055	-0.6026	0.0035	0.0035	-0.6023	0.0035	0.0035
50	$\theta_1 = 10$	9.8446	1.3942	1.4167	10.0858	1.1170	0.9998	10.1466	0.9812	1.0015
	$\theta_2 = -0.6$	-0.5961	0.0127	0.0127	-0.6047	0.0116	0.0084	-0.6050	0.0081	0.0081
30	$\theta_1 = 10$	9.8111	2.7453	2.7775	10.1339	0.2657	2.2352	10.2456	2.2298	2.2873
	$\theta_2 = -0.6$	-0.6011	0.0208	0.0208	-0.6132	0.0029	0.0158	-0.6111	0.0154	0.0155

The analysis of the presented results allows us to conclude that, in general, the accuracy of the SLS and PMM2 estimates is practically the same. For most cases, as the sample size n increases, the relative accuracy of PMM2 gradually increases and tends to the theoretically determined (27) almost twofold advantage (in the RMS metric) over OLS. The only exceptions are estimates of the scaling parameter θ_1 of the regression growth model with the relative improvement in accuracy is only 10–15%.

Table 2. Results of statistical modeling of the regression model parameters estimation $Y = \frac{\theta_1}{[1+\exp(\theta_2+\theta_3X)]} + \xi$

n		OLS			SLS			PMM2		
		Mean	Var	MSE	Mean	Var	MSE	Mean	Var	MSE
200	$\theta_1 = 10$	10.003	0.0133	0.0133	10.0003	0.0116	0.0116	9.9981	0.0118	0.0118
	$\theta_2 = 1.5$	1.5145	0.0449	0.0451	1.5038	0.0236	0.0235	1.5039	0.02389	0.0239
	$\theta_3 = 0.8$	-0.8061	0,0081	0,0082	-0,8031	0,0041	0,0041	-0,804	0,0043	0,0043
100	$\theta_1 = 10$	9.9979	0.02534	0.0253	9.9912	0.0224	0.0225	9.9866	0.0225	0.0226
	$\theta_2 = 1.5$	1.5128	0.0723	0.0724	1.4904	0.0415	0.0415	1.4936	0.0426	0.0426
	$\theta_3 = 0.8$	-0.8032	0.0135	0.0134	-0.7982	0.0075	0.0074	-0.801	0.0079	0.0079
50	$\theta_1 = 10$	10.019	0.057	0.0573	10.0091	0.0497	0.04968	10.001	0.04869	0.0486
	$\theta_2 = 1.5$	1.5404	0.1531	0.1545	1.4983	0.0973	0.0972	1.4819	0.1035	0.1037
	$\theta_3 = 0.8$	-0.8131	0.0281	0.0283	-0.8029	0.0168	0.0168	-0.8	0.0171	0.0171
30	$\theta_1 = 10$	10.038	0.1003	0.1017	10.0061	0.0826	0.0825	9.9979	0.0837	0.0836
	$\theta_2 = 1.5$	1.5961	0.2717	0.2806	1.5107	0.1848	0.1847	1.4727	0.2149	0.2154
	$\theta_3 = 0.8$	-0.8241	0.0474	0.0479	-0.8093	0.0317	0.0318	-0.805	0.0372	0.0372

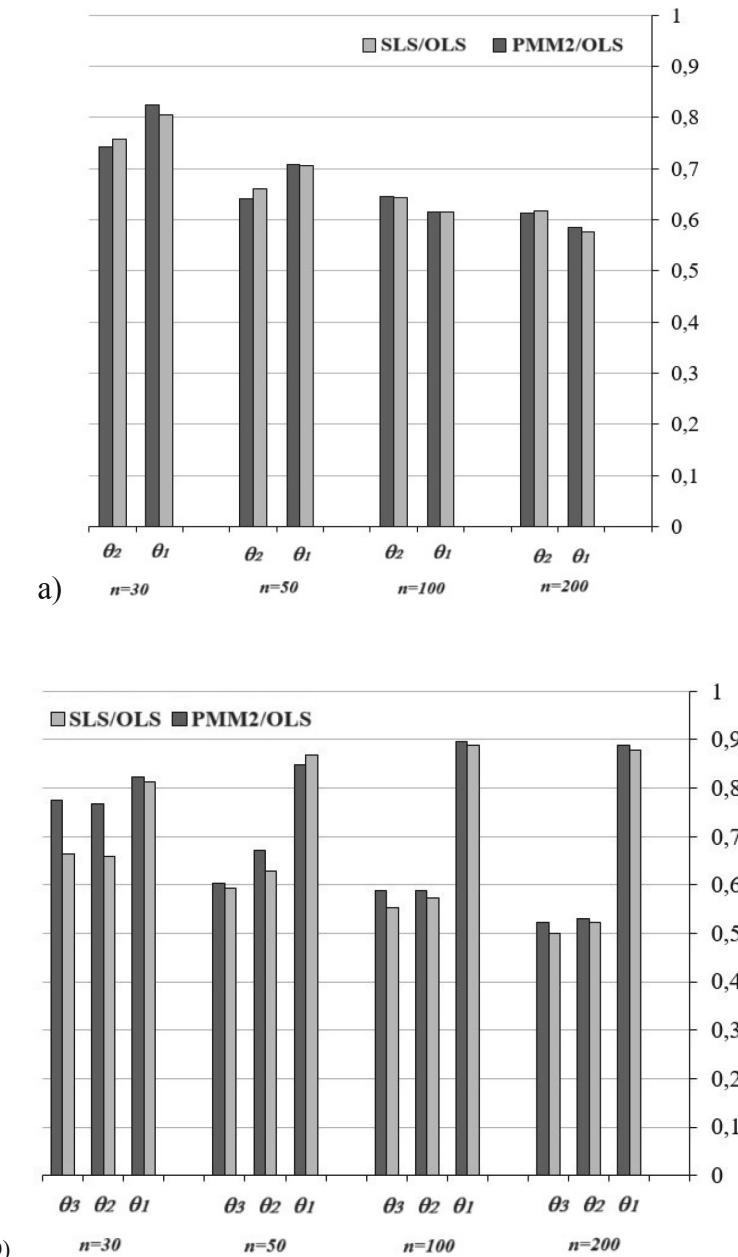


Fig. 1. Experimental values of the MSE ratio of estimates *a) model* $Y = \theta_1 \exp(\theta_2 X) + \xi$, *b) model* $Y = \frac{\theta_1}{[1 + \exp(\theta_2 + \theta_3 X)]} + \xi$

9 Conclusions

Based on the apparatus of Kunchenko stochastic polynomials, a new approach to adaptive finding estimates of nonlinear regression parameters in the asymmetric distribution of random errors is proposed. The approach is based on the use of power polynomials, which makes it possible to consider deviations from the Gaussian idealization in a realistic way in the higher-order statistics (moments or cumulants). It is shown that the general problem can be algorithmically reduced to solving a system of nonlinear stochastic equations using the Newton-Raphson numerical iteration procedure.

A theoretical analysis of the accuracy of polynomial estimates at degree $S = 2$ is carried out under the condition of asymmetric distribution of non-Gaussian errors. It is shown that the application of the proposed approach potentially reduces the variance of polynomial estimates (compared to the known least squares methods), the value of which depends on the cumulative coefficients of skewness and kurtosis. When the distribution is symmetric, PMM2 estimates become equivalent to OLS.

A comparative analysis of the accuracy of parameter estimates for two types of nonlinear regression models (exponential and growth) is carried out by means of simulation efficiency of PMM2 and SLS estimates, which significantly outperform OLS estimates. The similarity of the accuracy of SLS and PMM2 can be explained by the fact that both estimators use the same additional information about the probabilistic nature of the random component of regression models in the form of moments of regression residuals up to the 4th order.

It should be noted that the polynomial maximization method has an additional potential. It consists in the possibility of scaling the proposed approach by increasing the degree of the forming stochastic polynomials. This makes it possible to use the moment-cumulative description of higher orders algorithmically simply. Another interesting aspect is the possibility of using alternative (not only power) basis functions to form stochastic polynomials. The basis of such a basis, for example, can be trigonometric transformations, which will allow us to operate with a probabilistic description in the form of a characteristic function. All this should be the direction of future research.

References

1. Gallant, A.R.: Nonlinear statistical models. Wiley, New York (1987)
2. Wood, S.N.: Statistical inference for noisy nonlinear ecological dynamic systems. *Nature* **466**, 1102–1107 (2010)
3. Moon, K.-H., Han, S.W., Lee, T.S., Seok, S.W.: Approximate mpa-based method for performing incremental dynamic analysis. *Nonlinear Dyn.* **67**, 2865–2888 (2012)
4. Paula, D., Linard, B., Andow, D.A., Sujii, E.R., Pires, C.S.S., Vogler, A.P.: Detection and decay rates of prey and prey symbionts in the gut of a predator through metagenomics. *Mol. Ecol. Resour.* **15**, 880–892 (2015)
5. Santos, J.D.A., Barreto, G.A.: An outlier-robust kernel RLS algorithm for nonlinear system identification. *Nonlinear Dyn.* **90**, 1707–1726 (2017)
6. Wang, L.: Estimation of nonlinear Berkson-type measurement error models. *Stat. Sin.* **13**, 1201–1210 (2003)
7. Wang, L., Leblanc, A.: Second-order nonlinear least squares estimation. *Ann. Inst. Stat. Math.* **60**, 883–900 (2008). <https://doi.org/10.1007/s10463-007-0139-z>

8. Kim, M., Ma, Y.: The efficiency of the second-order nonlinear least squares estimator and its extension. *Ann. Inst. Stat. Math.* **64**, 751–764 (2012). <https://doi.org/10.1007/s10463-011-0332-y>
9. Gao, L.L., Zhou, J.: New optimal design criteria for regression models with asymmetric errors. *J. Stat. Plann. Inference* **149**, 140–151 (2014). <https://doi.org/10.1016/j.jspi.2014.01.005>
10. He, L., Yue, R.X.: R-optimality criterion for regression models with asymmetric errors. *J. Stat. Plann. Inference* **199**, 318–326 (2019). <https://doi.org/10.1016/j.jspi.2018.07.008>
11. Salamh, M., Wang, L.: Second-order least squares estimation in nonlinear time series models with ARCH errors. *Econometrics* **9**(4), 41 (2021). <https://doi.org/10.3390/econometrics9040041>
12. Zhang, J., Cui, L.: Exponential parametric distortion nonlinear measurement errors Models. *Commun. Stat. Theor. Methods* (2022). <https://doi.org/10.1080/03610926.2022.2111526>
13. He, L., Yue, R.X., Du, A.: Optimal designs for comparing curves in regression models with asymmetric errors. *J. Stat. Plann. Infer.* **228**, 46–58 (2024). <https://doi.org/10.1016/j.jspi.2023.06.005>
14. Kunchenko, Y.: Polynomial Parameter Estimations of Close to Gaussian Random variables. Shaker Verlag, Germany, Aachen (2002)
15. Zabolotnii, S.W., Martynenko, S.S., Salypa, S.V.: Method of verification of hypothesis about mean value on a basis of expansion in a space with generating element. *Radio Electron. Comm. Syst.* **61**, 222–229 (2018). <https://doi.org/10.3103/S0735272718050060>
16. Chertov, O., Slipets, T.: Epileptic seizures diagnose using kunchenko polynomials template matching. *Progress in Industrial Mathematics at ECMI 2012*. Springer International Publishing, Cham, pp. 245–248 (2014). https://doi.org/10.1007/978-3-319-05365-3_33
17. Zabolotnii, S.V., Warsza, Z.L.: Semi-parametric polynomial method for retrospective estimation of the change-point of parameters of non-Gaussian sequences. In: *Advanced Mathematical and Computational Tools in Metrology and Testing X*, pp. 400–408 (2015). https://doi.org/10.1142/9789814678629_0048
18. Zabolotnii, S.W., Warsza, Z.L.: Semi-parametric polynomial modification of CUSUM algorithms for change-point detection of non-Gaussian sequences. In: *Electronic Proceedings of XXI IMEKO World Congress “Measurement in Research and Industry” September 2015 Prague Czech Republic*, vol. 30, pp. 2088–2091 (2015)
19. Zabolotnii, S., Warsza, Z.L., Tkachenko, O.: Polynomial estimation of linear regression parameters for the asymmetric PDF of errors. In: *Advances in Intelligent Systems and Computing*, vol. 743 (2018). https://doi.org/10.1007/978-3-319-77179-3_75
20. Zabolotnii, S., Tkachenko, O., Warsza, Z.L.: Application of the polynomial maximization method for estimation parameters in the polynomial regression with non-gaussian residuals. In: *Advances in Intelligent Systems and Computing*, vol. 1390. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-74893-7_36
21. Zabolotnii, S., Tkachenko, O., Warsza, Z.L.: Application of the polynomial maximization method for estimation parameters of autoregressive models with asymmetric innovations. In: *Advances in Intelligent Systems and Computing*, vol. 1427. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-03502-9_37
22. Zabolotnii, S., Tkachenko, O., Warsza, Z.L.: Polynomial maximization method for estimation parameters of asymmetric non-gaussian moving average models. *Lecture Notes in Networks and Systems*, vol. 630. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-25844-2_21
23. Seber, G.A.F., Wild, C.J.: *Nonlinear Regression*. Wiley, New York (1989)
24. Cook, R.D., Weisberg, S.: *Residuals and influence in regression. Monographs on Statistics and Applied Probability* (1982). <https://doi.org/10.2307/1269506>
25. Jarque, C.M., Bera, A.K.: A tests of observations and regression residuals. *Int. Stat. Rev.* **55**, 163–172 (1987)

26. Dennis, J.E. Jr., Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Siam (1996)
27. Bojarczak, P., Nowakowski, W.: Application of deep learning networks to segmentation of surface of railway tracks. Sensors **21**, 4065 (2021). <https://doi.org/10.3390/s21124065>, ISSN1424-8220

Author Queries

Chapter 30

Query Refs.	Details Required	Author's response
AQ1	Per Springer style, both city and country names must be present in the affiliations. Accordingly, we have inserted the city and country names in affiliations. Please check and confirm if the inserted city and country names are correct. If not, please provide us with the correct city and country names.	
AQ2	Please be aware that your name and affiliation and if applicable those of your co-author(s) will be published as presented in this proof. If you want to make any changes, please correct the details now. Please note that after publication corrections won't be possible. Due to data protection we standardly publish professional email addresses, but not private ones, even if they have been provided in the manuscript and are visible in this proof. If you or your co-author(s) have a different preference regarding the publication of your mail address(es) please indicate this clearly. If no changes are required for your name(s) and affiliation(s) and if you agree with the handling of email addresses, please respond with "Ok".	
AQ3	Please confirm if the whether section headings are correctly identified.	
AQ4	Caption of Tables 1 and 2 seems to be identical. Please check and correct if necessary.	