



(ARTICLE)

Convergence and Statistical Performance of NLS and SNLS-VP under Non-Normal Errors and Challenging Starting Values

* Nawal Mohamed Othman Mustafa

Department of Statistics, Faculty of Science, Al-Marj, University of Benghazi, Al-Marj, Libya

التقارب والأداء الإحصائي لطريقتي المربعات الصغرى غير الخطية (NLS) والمربعات الصغرى غير الخطية القابلة للفصل مع الإسقاط المتغير (SNLS-VP) في ظل الأخطاء غير الطبيعية وقيم البدء الصعبة

نوال محمد عثمان مصطفى

قسم الإحصاء، كلية العلوم المرج، جامعة بنغازي، المرج، ليبيا

Publication history: Received on 03 October 2026; Accepted on 09 October 2026; Published on 10 October 2026

Abstract:

The study examines the convergence and performance of the standard Nonlinear Least Squares (NLS) and the Separable Nonlinear Least Squares with Variable Projection Method (SNLS-VP) under non-normal errors and poor starting values. A Monte Carlo study with 500 replications per simulation scenario was carried out with three levels of starting-value difficulty (easy, medium, and hard), three error distributions (Normal, Student's t with 3 degrees of freedom, and Log-normal), and three sample sizes ($n = 50, 100, \text{ and } 200$). We compared the two methods in terms of Bias, Root Mean Squared Error (RMSE), Average Number of Iterations (ANI), and Convergence Rate (CR). Results show that NLS generally yields smaller Bias and RMSE, especially for small sample sizes and under Normal and t_3 errors. In contrast, SNLS-VP shows more stable convergence in the considered scenarios, with a 100% convergence rate across all simulation setups. SNLS-VP is also more robust to difficult starting values, whereas NLS is generally faster when it converges but requires better starting values. The differences in statistical performance between the two procedures diminish as the sample size increases, especially under Log-normal errors. In sum, the results suggest that the relative performance of NLS and SNLS-VP depends on sample size, error distribution, and starting-value conditions, indicating a balance between statistical accuracy and computational convergence.

Keywords: Nonlinear least squares; separable nonlinear least squares; variable projection; convergence; non-normal errors; starting values.

تتناول هذه الدراسة تقارب وأداء طريقة المربعات الصغرى غير الخطية التقليدية (NLS) وطريقة المربعات الصغرى غير الخطية القابلة للفصل باستخدام أسلوب الإسقاط المتغير (SNLS-VP)، وذلك في ظل أخطاء غير طبيعية وقيم ابتدائية ضعيفة. أُجريت دراسة محاكاة بطريقة مونت كارلو باستخدام 500 تكرار لكل سيناريو محاكاة، مع ثلاثة مستويات لصعوبة القيم الابتدائية (سهلة، ومتوسطة، وصعبة)، وثلاثة توزيعات للأخطاء (التوزيع الطبيعي، وتوزيع Student's t بدرجات حرية تساوي 3، والتوزيع اللوغاريتمي الطبيعي)، وثلاثة أحجام للعينات (200 و 50,100 و n). وتمت مقارنة الطريقتين من حيث التحيز (Bias)، والجذر التربيعي لمتوسط مربع الخطأ (RMSE)، ومتوسط عدد التكرارات (ANI)، ومعدل التقارب (CR). أظهرت النتائج أن طريقة NLS تحقق عمومًا قيمًا أقل لكل من التحيز و RMSE، ولا سيما عند أحجام العينات الصغيرة وتحت أخطاء التوزيع الطبيعي وأخطاء t3، وفي المقابل أظهرت طريقة SNLS-VP استقرارًا أكبر في التقارب ضمن السيناريوهات المدروسة، حيث حققت معدل تقارب بلغ 100% في جميع سيناريوهات المحاكاة. كما تبين أن SNLS-VP أكثر متانة تجاه القيم الابتدائية الصعبة، في حين أن NLS تكون عمومًا أسرع عند تحقيق التقارب، لكنها تتطلب قيمًا ابتدائية أفضل. وتتضاءل الفروق في الأداء الإحصائي بين الطريقتين مع زيادة حجم العينة، ولا سيما تحت أخطاء التوزيع اللوغاريتمي الطبيعي. وبوجه عام، تشير النتائج إلى أن الأداء النسبي لطريقتي NLS و SNLS-VP يتأثر بحجم العينة، وتوزيع الأخطاء، وظروف القيم الابتدائية، مما يعكس وجود توازن بين الدقة الإحصائية واستقرار التقارب الحسابي.

الكلمات المفتاحية: المربعات الصغرى غير الخطية؛ المربعات الصغرى غير الخطية القابلة للفصل؛ الإسقاط المتغير؛ التقارب؛ الأخطاء غير الطبيعية؛ القيم الابتدائية.

1. Introduction

Nonlinear regression models are also widely used when linear regression is insufficient to model the relationship between the response and predictor variables (Bates & Watts, 1988). The most widely used approach in order to estimate nonlinear models' parameters is Nonlinear Least Squares (NLS) (Bates & Watts, 1988; Dennis & Schnabel, 1996). However, its numerical behavior can be significantly influenced by the initial parameter estimates, and its statistical behavior may be influenced by the error distributions, particularly for heavy-tailed or skewed errors (Bates & Watts, 1988). In the special, but still highly relevant, situation where the parameters of a nonlinear regression model can be partitioned into nonlinear and linear parameters, Separable Nonlinear Least Squares (SNLS) and Variable Projection (VP) apply (Golub & Pereyra, 1973). Variable Projection reduces the dimension of the nonlinear optimization problem by eliminating the linear parameters analytically (Golub & Pereyra, 1973; Dennis & Schnabel, 1996). The comparative advantage of NLS with respect to SNLS-VP, on the other hand, may vary with the sample size and the initial parameter vector and the error distribution may be such that they are very adverse. A considerable amount of literature has investigated NLS and Variable Projection, but their behavior in the presence of a combination of these factors is unknown. In order to fill this gap, this paper employs Monte Carlo simulation to analyze NLS and its counterpart SNLS-VP for two types of separable nonlinear models. The procedures are evaluated through bias, RMSE, ANI, and CR

2. Materials and Methods

1. Nonlinear Regression Model

Nonlinear regression models are a class of models in which the response variable is a nonlinear function of unknown parameters and one or more explanatory variables. Unlike the linear regression model, the nonlinear regression model can have parameters that are involved nonlinearly in the model. As a result, estimation of parameters is done through numerical iterative procedures and can be computationally intensive (Bates & Watts, 1988; Seber & Wild, 2003).

The nonlinear regression model can be written as

$$y_i = f(X_i, B) + \varepsilon_i \quad i = 1, 2, \dots, n \quad (1.1)$$

y_i This is the response variable

x_i This is the vector of regressors.

$f(\cdot)$ The nonlinear function of explanatory variables to the dependent variable

B This is the vector of parameters, where ρ is the parameter of interest

The most prevalent approach for obtaining an estimator of unknown parameters is termed NLS estimation. The NLS estimator is defined as a solution to the following minimization problem of squared residuals. Since the objective function is nonlinear in terms of the unknown parameters, the optimization problem has no analytic solution. Therefore, numerical methods are applied to estimate the parameter vector. the efficiency of nonlinear regression estimation method depends not only on the assumed structure of the errors, but also on numerical issues such as the choice of the initial value to the parameter vector and the numerical optimization procedure. Such difficulties are made more pronounced when the initial guesses for the parameter values are far from their true values and when the errors are non-normally distributed.

2. Nonlinear Least Squares (NLS)

Unlike linear least squares, NLS usually does not produce closed-form solutions for the estimates of parameters. But the estimates of the parameters are found by minimizing the sum of squared residuals numerically (Bates & Watts, 1988):

$$\hat{\mathbf{B}} = \underset{\mathbf{B}}{\operatorname{argmin}} \mathbf{S}(\mathbf{B}) = \underset{\mathbf{B}}{\operatorname{argmin}} \sum_{i=1}^n [y_i - f(x_i, \mathbf{B})]^2 \quad (1.2)$$

Since the objective function is a nonlinear function of parameters, the NLS technique usually needs numerical iterative optimization methods. Typical algorithms include Gauss–Newton and Levenberg–Marquardt algorithms (Bates & Watts, 1988; Mohammad et al., 2019). A key feature of NLS estimation is that it is sensitive to initial parameter values. These values specify the initial point for the iterative optimization procedure, and we have seen how the initial point can have an impact on convergence and on the number of iterations needed, but also on the solution to which the algorithm converges. For challenging nonlinear models, bad starting values may cause the algorithm to converge slowly, to different solutions, or not to converge. Thus, the determination of starting values is an important issue for evaluating numerical convergence and NLS estimation in terms of statistical performance (Bates & Watts, 1988).

3. Numerical Optimization for NLS

1 Gauss-Newton Algorithm

The Gauss–Newton method is a widely used approach for estimating the parameters of a nonlinear model (Bates & Watts, 1988; Seber & Wild, 2003). It operates by linearizing the model around the current parameter estimates using a first-order Taylor expansion and iteratively refining them to achieve the best fit. The objective is to minimize the sum of squared residuals, which quantifies the discrepancy between the model's predictions and the observed data.

The algorithm begins with an initial parameter vector θ^0 and approximates the nonlinear model function around the current estimate θ^t via a first-order Taylor expansion:

$$f(x, \theta) \approx f(x, \theta^{(t)}) + J^{(t)}(\theta - \theta^{(t)}) \quad (1.3)$$

where $J^{(t)}$ represents the Jacobian matrix containing the partial derivatives of the model function with respect to the parameters:

$$J_{ij}^{(t)} = \frac{\partial f(x_i, \theta)}{\partial \theta_j} \Big|_{\theta = \theta^{(t)}}$$

Based on this linear approximation, the parameter update vector $\Delta\theta^{(t)}$ is computed by solving the normal equations:

$$(J^{(t)T} J^{(t)}) \Delta\theta^{(t)} = J^{(t)T} r^{(t)}$$

Where $r^{(t)}$ denotes the residual vector representing the difference between the actual observations and the model predictions. The parameters are updated iteratively as follows:

$$\theta^{(t+1)} = \theta^{(t)} + \Delta\theta^{(t)}$$

This process continues until a convergence criterion is satisfied—typically when the change in the sum of squared errors or parameter estimates falls below a predefined tolerance threshold ε .

2. Levenberg–Marquardt Algorithm

This method is a numerical procedure based on iterative techniques that combine the Gauss–Newton method and gradient descent when estimating parameter values in nonlinear least-squares models. A damping factor is also applied in this procedure to balance the strengths of both approaches while influencing the convergence and stability of the algorithm

$$B^K = B^{(K+1)} + \Delta B \quad (1.4)$$

4. Starting Values in Nonlinear Regression

Starting values are crucial in nonlinear regression, as numerical methods of estimation (e.g., Nonlinear Least Squares (NLS) and Separable Nonlinear Least Squares (SNLS)) must start from an initial parameter vector, on which the subsequent iterative estimation process is based. The quality of initial estimates can also influence how fast the algorithm converges, how many computational resources are needed, and sometimes even whether the algorithm reaches an acceptable solution (Bates & Watts, 1988; Seber & Wild, 2003).

1. Suitable Initial Values

Initial values are suitable if they are not too far from the true parameter values. With good initial values, the algorithm's estimation process starts effectively. For example, suppose that the true parameter vector is:

$$B = (2, 1, 0.5)^T$$

A reasonable set of initial values is given by:

$$B^0 = (1.9, 1.1, 0.45)^T$$

In this case, the initial parameter vector is close enough to the true vector to allow for a meaningful estimation process.

2. Challenging Initial Values

Challenging initial values are those intentionally chosen to be very far from the true parameter vector. Such initial values are used to test the sensitivity and robustness of the nonlinear estimation algorithms to bad initial values. For example, when the true parameter vector is

$$B = (2.1, 0.5)^T$$

Some difficult initial values are of the form

$$B^0 = (8, -4)^T$$

Thus the initial point is off by a large distance from the true parameter vector. This is a more demanding numerical problem and demonstrates how the NLS and SNLS methods behave in the presence of bad initial estimates.

3. Separable Nonlinear Least Squares (SNLS)

Separable nonlinear models refer to a category of nonlinear regression model where the model can be expressed such that part of the parameters are linear while others are nonlinear in nature. This aspect of reparability evident in this class of models enables the estimation of linear parameters by means of least squares and the nonlinear parameter by nonlinear optimization, thereby reducing the dimensionality of the optimization problem and increasing efficiency and stability of the numerical approach compared to estimating nonlinear parameters directly (Ruhe & Wedin, 1980; Golub & Pereyra, 2003).

Separable nonlinear models can be expressed in the following form:

$$y_i = \sum_{j=1}^p B_j g_j(x_i, \alpha)$$

or in matrix form

$$y = X(\alpha)B + \varepsilon \quad (1.5)$$

B : Linear parameters

α : Non-linear parameters

For a fixed value of α It is possible to estimate B directly.

$$\hat{B}(\alpha) = [X(\alpha)'X(\alpha)]^{-1}X(\alpha)'y$$

Then, a numerical search is conducted for:

$$\hat{\alpha} = \operatorname{argmin} ||y - X(\alpha)\hat{B}(\alpha)||^2$$

4. Results

Structure of Separable Nonlinear Models

Variable Projection

The Modified Newton approach is a technique with the use of Newton methods in the field of nonlinear optimization where the optimization is posed using separable conditions. This method is known because it focuses on the peculiar properties of the developed model, where we see dividing the parameters between the linear and non-linear components; thus, the method can be applied directly to the separable non-linear least-squares (SNLS) problem. Given the specific real value for the non-linear parameters, we first have the solution for linear least-squares, which helps find out linear parameter estimates, and we can substitute them back into the objective function, thus decreasing our original optimization problem to a lower-dimensional system where only non-linear parameters are present. The method thus decreases the number of parameters used in our numerical optimization, thus improving computation speed and reliability in the case we have a considerable number of linear parameters in our model (Golub & Pereyra, 2003; O'Leary & Rust, 2013).

The modified Newton method is a Newton-type scheme for nonlinear

$$y = x(\theta)B + \varepsilon$$

Where β Linear parameters θ Non-linear parameters.

Estimating linear parameters when θ is fixed

$$\min_{\theta} ||y - X(\theta)\hat{B}(\theta)||^2$$

Or equivalently:

$$\min_{\theta} = ||[I - P_{X(\theta)}]y||^2$$

$$P_X = XX^+$$

The Simulation Study

A Monte Carlo simulation study was carried out to compare the numerical performance of nonlinear least squares (NLS) and separable nonlinear least squares with variable projection (SNLS-VP). The study was designed to examine how different data and estimation conditions affect the accuracy and convergence of these two methods. The simulation looked at three factors: sample size, difficulty of starting values, and error distribution. Three sample sizes were used: $n = 50$, $n = 100$, and $n = 200$. To test how sensitive the methods are to starting values, three levels of difficulty were considered: Easy, Moderate, and Difficult. For the error distributions, three types were included: the Normal distribution, Student's t-distribution with 3 degrees of freedom (t3), and the Log-normal distribution. These distributions were chosen to represent the normal case and deviations from normality, including heavy tails and skewness. The two methods under comparison were NLS and SNLS-VP. For each combination of the settings, 500 Monte Carlo replications were run. Performance was measured using four criteria: Bias, Root Mean Squared Error (RMSE), Number of Iterations (ANI), and Convergence Rate (CR). Bias and RMSE were used to evaluate estimation accuracy. ANI and CR were used to assess how well the methods converged numerically. The experimental factors and their levels are summarized in Table 1.

Table 1: Examples of some regression functions used in the simulation study

Factor	Levels/ Values
Sample size (n)	50, 100, 200
Regression model	$y = 2 + 1.3x_i + 3e^{0.8x_i} + \varepsilon$ $y = 2 + 1.5x_i + 3e^{(0.8x_i)}(x_i - 2) + \varepsilon$
Starting-value difficulty	Easy, Moderate, Difficult
Error distribution	Normal, t3, Lognormal
Estimation method	NLS, SNLS-VP
Monte Carlo replications	500 per cell
Performance measures	Bias, RMSE, ANI, Convergence Rate CR)

2.1 Numerical Summary for the Simulation Study

The simulation results are presented in Tables 2–5, which include a comparison of NLS with SNLS-VP for different sample sizes ($n = 50, 100, 200$), levels of starting values (Easy, Moderate, Difficult), and error distributions (Normal, t3, and Log-normal). The methods are contrasted with respect to Bias, RMSE, ANI, and CR. Specifically, smaller Bias and RMSE indicate better estimation accuracy, smaller ANI indicates greater computational efficiency, and a higher CR indicates better numerical stability and convergence. These indices reflect the statistical and numerical performance of the two methods

Table 2. Bias and Root Mean Squared Error (RMSE) of NLS and SNLS-VP under Different Sample Sizes, Initial Values, and Error Distributions Based on 500 Monte Carlo Replications: First Regression Function

	NLS					SNLS-VP		
	n	Starting values	Normal	t3	Log-normal	Normal	t3	Log-normal
Bias	50	Easy	0.2125	0.1856	0.2194	0.4709	0.5629	0.2992
		Moderate	0.3461	0.3249	0.2860	0.4830	0.6442	0.2921
		Difficult	0.2358	0.3265	0.0810	0.5248	0.6890	0.1893
	100	Easy	0.2032	0.2371	0.1155	0.2786	0.4679	0.1313
		Moderate	0.3925	0.2184	0.1208	0.4789	0.3252	0.1086
		Difficult	0.2063	0.2062	0.0660	0.3158	0.4439	0.1240
	200	Easy	0.1751	0.1514	0.0544	0.2132	0.2373	0.0616
		Moderate	0.1652	0.1435	0.0053	0.1607	0.2160	0.0009-
		Difficult	0.0829	0.1514	0.0353	0.1745	0.2241	0.0607
RMSE	50	Easy	1.0876	1.0574	0.9967	1.5804	1.7060	1.1951
		Moderate	1.2639	1.2090	1.0359	1.5757	1.7802	1.1335
		Difficult	1.2303	1.2913	0.8652	1.6838	1.8034	1.0586
	100	Easy	1.0486	1.1256	0.7741	1.2454	1.5498	0.8355
		Moderate	1.1805	1.1117	0.8433	1.3748	1.3758	0.8706
		Difficult	1.1136	1.0932	0.7441	1.2964	1.4715	0.8364
	200	Easy	0.8491	0.9458	0.6089	0.9764	1.1251	0.6403
		Moderate	0.9140	0.9333	0.6064	0.9498	1.1408	0.6262
		Difficult	0.8603	0.9898	0.5420	0.9885	1.1213	0.6170

Table 3. Average Number of Iterations (ANI) and Convergence Rate (CR) of NLS and SNLS-VP under Different Sample Sizes, Initial Values, and Error Distributions Based on 500 Monte Carlo Replications: First Regression Function

	NLS					SNLS-VP		
	n	Starting values	Normal	t3	Log-normal	Normal	t3	Log-normal
ANI	50	Easy	45.07	42.21	41.17	39.77	40.91	35.27
		Moderate	60.68	59.88	57.28	40.32	42.30	35.41
		Difficult	88.04	80.98	79.85	39.23	41.12	35.98
	100	Easy	41.36	44.40	34.45	33.21	34.59	28.87
		Moderate	59.94	62.48	51.67	34.94	36.21	30.94
		Difficult	84.20	81.89	79.27	33.69	35.80	29.27
	200	Easy	35.63	39.59	30.35	29.28	31.87	25.53
		Moderate	48.92	54.02	42.76	28.60	31.12	26.94
		Difficult	80.00	90.74	69.53	29.93	29.86	27.22
Convergence Rate(CR)	50	Easy	90.0	87.0	96.4	100	100	100
		Moderate	93.6	89.4	98.6	100	100	100
		Difficult	89.4	89.8	93.2	100	100	100
	100	Easy	95.2	92.0	98.2	100	100	100
		Moderate	96.2	95.0	99.6	100	100	100
		Difficult	92.4	91.2	93.8	100	100	100
	200	Easy	97.4	95.2	99.0	100	100	100
		Moderate	98.2	96.6	99.8	100	100	100
		Difficult	93.2	95.8	96.6	100	100	100

Table 4. Bias and Root Mean Squared Error (RMSE) of NLS and SNLS-VP under Different Sample Sizes, Initial Values, and Error Distributions Based on 500 Monte Carlo Replications: Second Regression Function

	NLS					SNLS-VP		
	n	Starting values	Normal	t3	Log-normal	Normal	t3	Log-normal
Bias	50	Easy	-0.3162	-0.1562	-0.2308	0.5624	0.5883	0.5402
		Moderate	-0.1806	-0.1782	-0.0609	0.7025	0.6549	0.5266
		Difficult	0.1238	0.0915	0.0609	0.9197	1.0317	0.4763
	100	Easy	-0.2313	-0.2724	-0.2478	0.5866	0.6658	0.3694
		Moderate	0.0478	0.0335	0.1217	0.6836	0.6023	0.2879
		Difficult	0.0498	0.1122	0.1109	0.5912	0.8762	0.3929
	200	Easy	-0.2400	-0.2705	-0.1122	0.4703	0.4271	0.3108
		Moderate	0.0886	0.0639	0.0577	0.4480	0.4413	0.1041
		Difficult	0.2284	0.1396	0.2375	0.5153	0.6794	0.1041
RMSE	50	Easy	1.2723	1.2202	1.2877	2.0710	2.0715	1.8970
		Moderate	1.2499	1.1302	1.1467	2.1467	2.1171	1.9533
		Difficult	1.1055	1.1990	0.9771	2.2244	2.3448	1.8672
	100	Easy	1.2404	1.2977	1.2944	1.9739	2.0651	1.5333
		Moderate	1.2087	1.1823	1.1592	2.0580	1.9982	1.6072
		Difficult	1.1023	1.2404	1.1181	1.9689	2.2177	1.6922
	200	Easy	1.4604	1.3174	1.3095	1.8054	1.8143	1.4071
		Moderate	1.1178	1.1541	0.9779	1.7867	1.7681	1.2976
		Difficult	1.2307	1.2140	1.0179	1.7807	1.9714	1.3172

Table 5. Average Number of Iterations (ANI) and Convergence Rate (CR) of NLS and SNLS-VP under Different Sample Sizes, Initial Values, and Error Distributions Based on 500 Monte Carlo Replications: Second Regression Function

	NLS					SNLS-VP		
	n	Starting values	Normal	t3	Log-normal	Normal	t3	Log-normal
ANI	50	Easy	19.07	16.11	10.39	31.68	32.30	25.10
		Moderate	15.86	15.11	9.71	30.87	31.72	25.58
		Difficult	10.43	11.72	8.05	33.59	35.82	29.83
	100	Easy	12.30	11.28	5.83	25.08	27.62	21.24
		Moderate	12.40	15.21	6.64	26.33	27.24	23.84
		Difficult	11.44	10.06	6.12	29.68	30.84	26.17
	200	Easy	6.66	8.60	4.12	20.39	22.25	19.78
		Moderate	7.55	10.25	4.18	22.61	24.31	22.74
		Difficult	8.62	7.99	5.12	26.44	26.19	25.58
Convergence Rate(CR)	50	Easy	80.6	82.2	83.0	100	100	100
		Moderate	87.4	84.8	86.0	100	100	100
		Difficult	81.6	80.8	86.8	100	100	100
	100	Easy	83.2	82.8	88.4	100	100	100
		Moderate	87.4	87.0	94.0	100	100	100
		Difficult	89.2	82.4	92.0	100	100	100
	200	Easy	86.8	86.4	93.2	100	100	100
		Moderate	89.2	91.4	96.6	100	100	100
		Difficult	90.4	88.0	97.2	100	100	100

5. Discussion

The Results and Discussion section presents and discusses the results. Under various sample sizes, initial-value conditions, and error distributions, Tables 2 and 3 depict the statistical accuracy and the convergence of NLS and SNLS-VP for the first regression function, based on 500 Monte Carlo replications.

Bias and RMSE

The results of Table 2 demonstrate that sample size and error distribution both noticeably influence the accuracy of the estimation for each of the two methods. In general, when the sample size increases from $n=50$ to $n=200$, Bias and RMSE are both reduced for NLS and SNLS-VP, which suggests that the accuracy of the estimation improves when the sample information increases. This diminution is even more pronounced in the log-normal error case, where the RMSE dramatically decreases as the sample size increases. For $n=50$, under the Normal and t_3 error distributions, the Bias and RMSE of NLS estimation are typically lower than those of SNLS-VP. Under the Easy initial values, for example, the RMSE of NLS is 1.0876 under Normal errors and 1.0574 under t_3 errors, compared to 1.5804 and 1.7060 for SNLS-VP. A similar trend is observed with the Moderate and Difficult initial values. Yet, under log-normal errors, the distinction between the two methods is far less, and SNLS-VP achieves a smaller RMSE with the Moderate and Difficult starting values.

As the sample size increases to $n=100$, the overall accuracy of both methods in estimation is better. Under the Normal and t_3 distributions, NLS generally maintains lower RMSE values, although SNLS-VP is competitive in some cases. SNLS-VP performs particularly well under log-normal errors, with RMSE values of 0.8355, 0.8706, and 0.8364 for Easy, Moderate, and Difficult initial values, respectively. These are close to, or in a few cases even lower than, the equivalent NLS results. The improvement becomes even more pronounced at $n=200$. Under both procedures, the biases become fairly small, in particular under log-normal errors. For instance, with Moderate initial values, Bias=0.0053 for NLS and almost zero for SNLS-VP under log-normal errors. These RMSE values are among the lowest in the table. This trend indicates that, with increasing sample size, the magnitude of estimation error is significantly lessened, which in turn guides both procedures to better estimates of the underlying regression parameters.

The influence of initial values can also be seen. Tougher initial values tend to increase Bias and RMSE in general, especially for small samples. This behavior is most pronounced for SNLS-VP when $n=50$ under the t_3 distribution, with the Bias and RMSE increasing up to 0.6890 and 1.8034, respectively. However, as the sample size increases, the detrimental effect of difficult starting values becomes much weaker. At $n=200$, the differences between the initial-value conditions are minimal, and this indicates that increasing the sample size resulted in more stable results.

Average Number of Iterations and Convergence Rate

Table 3 offers additional insight into the computational performance of the two estimation methods. Across the settings examined, NLS typically takes far more iterations than SNLS-VP, especially when the initial values are Moderate or Difficult. For instance, with a sample size of 50 and Difficult starting values, NLS needs between 79.85 and 88.04 iterations across the three error distributions. SNLS-VP, by comparison, uses 35.98–41.12.

This difference also appears with larger samples: at $n=200$ and Difficult starting values, NLS requires 80.00 iterations under Normal errors, 90.74 under t_3 errors, and 69.53 under log-normal errors. The corresponding figures for SNLS-VP are just 29.93, 29.86, and 27.22. Overall, SNLS-VP's computational advantage persists as the sample size grows. The initial values have a stronger impact on the time complexity of NLS. For instance, at $n=50$, the ANI of NLS under Normal errors grows from 45.07 for Easy initial values to 88.04 for Difficult initial values. On the other hand, its respective SNLS-VP values are 39.77 and 39.23. This relatively minimal difference implies that SNLS-VP is not as sensitive to the quality of the starting values in terms of the number of iterations required to converge. CR results in (b) further support this finding. SNLS-VP has an excellent convergence rate (100%) in all simulated situations. NLS also has high convergence rates, but its performance is more sensitive to the initial values and error distribution, especially when $n=50$. Under Difficult initial values, the NLS convergence rates are 89.4%, 89.8%, and 93.2% for Normal, t_3 , and log-normal errors, respectively. The convergence rates increase as the sample size increases, and they are greater than 93% for all NLS cases at $n=200$.

The performance in terms of statistical accuracy and convergence of NLS and SNLS-VP for the second regression function is summarized in **Tables 4 and 5** under varying sample sizes, initial points, and error distributions.

Statistical Accuracy

From Table 4, we observe that SNLS-VP has absolute Bias and RMSE larger than those of NLS in most simulation scenarios. The contrast is especially clear for $n=50$, with SNLS-VP having substantially greater Bias and RMSE across the three error distributions. For instance, in the Difficult initial values case, the RMSE of NLS is between 0.9771 and 1.1990, while the RMSE of SNLS-VP is between 1.8672 and 2.3448. The gap in performance decreases as the sample size grows in some cases, in particular under log-normal errors. The RMSE values for the two methods are generally smaller under log-normal errors when $n=200$, although the NLS method still provides better estimates in general. These findings imply that NLS has better statistical accuracy for the second regression function, but the differences tend to decrease as the sample size increases.

Performance of the Convergence Process

A pronounced superiority in the convergence reliability of SNLS-VP is observed in Table 5. SNLS-VP has a 100% convergence rate for all sample sizes, initial-value conditions, and error distributions. On the other hand, the convergence rates for NLS ranged from 80.6% to 97.2%, although they increased with larger sample sizes. This trend suggests that SNLS-VP is far more robust with respect to convergence, especially with small sample sizes or difficult starting values. In terms of computational complexity, our numerical results show that NLS is more efficient than SNLS-VP in all experiments for the second regression function, in the sense that it requires fewer iterations. The gap is most significant when $n=50$, and NLS needs around 8–19 iterations, while SNLS-VP needs 25–36 iterations. The number of iterations for both methods tends to decrease as the sample size grows. Therefore, SNLS-VP achieves significantly more reliable convergence at the expense of more iterations.

6. Conclusion

This study compared NLS and SNLS-VP under different sample sizes, initial values, and error distributions using 500 Monte Carlo replications. The results showed that NLS generally achieved better statistical accuracy, with lower Bias and RMSE, whereas SNLS-VP demonstrated superior convergence reliability, achieving a 100% convergence rate across all considered scenarios. For the second regression function, NLS also required fewer iterations, indicating greater computational efficiency in this setting. Overall, neither method consistently outperformed the other across all evaluation criteria. Therefore, the choice between NLS and SNLS-VP should depend on the characteristics of the nonlinear regression problem and the primary estimation objective, with NLS being preferable when statistical accuracy is the main concern and SNLS-VP when reliable convergence is prioritized.

7. References

- [1] Bates, D. M., & Watts, D. G. (1988). *Nonlinear regression analysis and its applications*. Wiley.
- [2] Dennis, J. E., Jr., & Schnabel, R. B. (1996). *Numerical methods for unconstrained optimization and nonlinear equations*. Society for Industrial and Applied Mathematics.
- [3] Golub, G. H., & Pereyra, V. (1973). The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, 10(2), 413–432.
- [4] Golub, G. H., & Pereyra, V. (2003). Separable nonlinear least squares: The variable projection method and its applications. *Inverse Problems*, 19(2), R1–R26.
- [5] Mohammad, H., Waziri, M. Y., & Santos, S. A. (2019). A brief survey of methods for solving nonlinear least-squares problems. *Numerical Algebra, Control and Optimization*, 9(1), 1–13.
- [6] O’Leary, D. P., & Rust, B. W. (2013). Variable projection for nonlinear least squares problems. *Computational Optimization and Applications*, 54(3), 579–593.
- [7] Ruhe, A., & Wedin, P.-Å. (1980). Algorithms for separable nonlinear least squares problems. *SIAM Review*, 22(3), 318–337.
- [8] Seber, G. A. F., & Wild, C. J. (2003). *Nonlinear regression*. Wiley Interscience.