

CENTRAL ASIAN JOURNAL OF INNOVATIONS ON TOURISM MANAGEMENT AND FINANCE

https://cajitmf.centralasianstudies.org/index.php/CAJITMF

Volume: 07 Issue: 04 | October 2026 ISSN: 2660-454X

Article

Differential Identification and Stable Polynomial Reconstruction from Nonuniform Experimental Data: The 5% Criterion, Operator Norm and Iterative Correction

Sherali Baratovich Ochilov[1], Oisha Kurbanovna Khudayberdieva[2]

1. Associate Professor, Department of Economics and Management, Bukhara State Technical University, Bukhara, Uzbekistan
2. Assistant, Department of Economics and Management, Bukhara State Technical University, Bukhara, Uzbekistan

*Correspondence: ochilovshb@mail.ru[1], oyshaxudayberdiyeva66@gmail.com[2]

Abstract: It is not always possible to describe an experimental process in advance by one of the standard analytical functions. Model selection becomes particularly difficult when the measurement points are nonuniformly spaced, the data contain random errors, and the functional form of the process is unknown. A sequential differential identification algorithm based on scaled divided differences is formalized. The model order m is taken as the first order at which the m-th order scaled divided differences satisfy a prescribed relative stability criterion; a polynomial approximation is then reconstructed by integrating the relation $y(m) \approx C$. For nonuniform grids, a differential operator matrix Wm is introduced, and the amplification of measurement errors is estimated through its operator norm. A deterministic sufficient stability condition is derived for a general tolerance γ; for $\gamma = 0.05$ it takes the form $\beta + \|Wm\|_\infty \varepsilon_0 \le |C|/41$. An explicit upper bound is also given for the reconstruction error when the function is smooth but not polynomial. A contractive Landweber-type iterative correction is added to damp the computational error of the initial coefficient estimates. A 100,000-run Monte-Carlo test on a nonuniform cubic example confirmed that the theoretical condition is sufficient and conservative. For a nonpolynomial synthetic process, the 5% criterion stopped at the third order and produced a cubic approximation with a range-normalized RMSE (NRMSE) of 0.103%. The results show that differential identification can be studied as a practical model-selection tool for nonuniform experimental data, although noise amplification in high-order numerical differentiation remains its main limitation. The framework does not treat finite differences or Landweber iteration as new; its contribution is the combined identification–stability–reconstruction workflow for nonuniform experimental data.

Keywords: Differential identification; divided differences; nonuniform grid; polynomial approximation; numerical stability; operator norm; Monte-Carlo simulation; Landweber iteration.

Citation: Ochilov S. B. Differential Identification and Stable Polynomial Reconstruction from Nonuniform Experimental Data: The 5% Criterion, Operator Norm and Iterative Correction. Central Asian Journal of Innovations on Tourism Management and Finance 2026, 7(4), 886-896

Received: 17th Jul 2026
Revised: 20th Aug 2026
Accepted: 15th Sep 2026
Published: 30th Oct 2026

Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

When an empirical function is constructed from experimental data, the model type is often chosen in advance: linear, quadratic, exponential, logistic, hyperbolic, and so on. Such an approach is natural if the mechanism of the process justifies the chosen form. However, when the functional form is unknown, the experimental points are nonuniformly spaced, and none of the standard functions represents the entire interval with sufficient accuracy, model selection itself becomes a separate problem.

The general possibility of polynomial approximation is guaranteed by the Weierstrass approximation theorem [1]. However, the theorem does not indicate which polynomial degree should be chosen, how this degree should be determined from noisy, nonuniform data, or how the reliability of high-order differentiation should be assessed.

Finite and divided differences are classical tools of numerical analysis. The recursive definition of divided differences, their connection with Lagrange interpolation, and the fact that the m-th order divided difference equals the m-th derivative at some intermediate point are standard results [2]. The efficient computation of weights for high-order derivatives on nonuniform grids has been studied extensively in the work of Fornberg [3, 4]. There are also closely related studies on determining the polynomial degree by finite differences and on estimating coefficients in the space of differences. The authors' earlier studies used finite differences to construct correlation dependences for functions of many variables [5] and to build correlation models on exact data. Therefore, this paper makes no claim of novelty regarding finite differences themselves or the classical indicator of polynomial degree.

Obtaining numerical derivatives from noisy data is, in principle, a delicate problem: a small measurement error can turn into a large error after differentiation. In regularization theory, this issue is treated as ill-posed numerical differentiation [6]. Practical remedies include local least-squares polynomial smoothing, regularized and variational differentiation, and comparative studies of local and global differentiation methods [7]. Error propagation and the stability of numerical algorithms in general have been systematically described by Higham [8].

The iterative correction used here belongs to the classical Landweber iteration, which dates back to 1951 [9] and is a standard regularizing iteration for ill-posed problems. In this paper it is employed as an auxiliary, stable correction stage following differential identification.

The aim of this study is to formalize and investigate a unified scheme that consists of the following steps:

- computing successive scaled divided differences on nonuniform data;
- selecting the first order that satisfies a relative stability criterion;
- obtaining a deterministic stability certificate for the result through the operator norm;
- reconstructing a polynomial approximation by integrating the relation $y^{(m)} \approx C$;
- correcting the coefficients, when necessary, by a contractive iteration.

2. Materials and Methods

Problem statement and the differential algorithm

Let the following experimental data be given:

$$x_0 < x_1 < \cdots < x_n, \quad y_i = f(x_i) + \varepsilon_i, \quad i = 0, 1, \ldots, n. \qquad (1)$$

Here f is an unknown smooth process and ε_i is a measurement or computational error. The steps $h_i = x_{i+1} - x_i$ are, in general, not equal.

For order zero, $D_i^{(0)} = y_i$, and higher orders are defined recursively by

$$D_i^{(r)} = r(D_{i+1}^{(r-1)} - D_i^{(r-1)}) / (x_{i+r} - x_i), \quad r = 1, 2, \ldots \qquad (2)$$

At the first order, (2) gives the ordinary local slope:

$$D_i^{(1)} = (y_{i+1} - y_i) / (x_{i+1} - x_i). \qquad (3)$$

From the recursion it follows that

$$D_i^{(m)} = m!\cdot[x_i, x_{i+1}, \ldots, x_{i+m}]y. \qquad (4)$$

Thus, the proposed computation is a table of scaled divided differences on a nonuniform grid. By the classical mean value theorem, if $f \in C^m$, then for every block there exists $\xi_i \in (x_i, x_{i+m})$ such that, for exact data,

$$D_i^{(m)} = f^{(m)}(\xi_i) \qquad (5)$$

This result forms the basis of the differential interpretation [1].

The 5% identification criterion

At order m, $N_m = n - m + 1$ values remain. Their mean is

$$\bar{D}^{(m)} = (1/N_m) \sum_{i=0}^{N_m-1} D_i^{(m)}, \quad (6)$$

and the relative dispersion criterion is

$$\delta_m = (100/N_m) \sum_{i=0}^{N_m-1} |D_i^{(m)} - \bar{D}^{(m)}| \,/\, |D_i^{(m)}|. \quad (7)$$

In this work, $\gamma = 5\ \%$ is adopted. The algorithm stops at the first order for which

$$\delta_m \le 5\ \%. \quad (8)$$

If m = 1, the local slopes are stable and the model is linear; if m = 2, it is quadratic; if m = 3, it is cubic; and so on. If the true process is not a polynomial, condition (8) does not imply that "the process is exactly a polynomial of degree m"; it means that the m-th derivative is nearly constant on the given interval and that a polynomial approximation of degree m can be tested.

In practice, to avoid the spurious conclusion "only one number remains, therefore it is a constant" at high orders, it is recommended that at least four values of $D_i^{(m)}$ remain. This rule is not a theorem but a practical safeguard for identification.

2.3. Stability theory on a nonuniform grid

Expression (4) can be written as a linear operator:

$$\bar{D}^{(m)} = W_m y. \quad (9)$$

For the i-th block, the exact weights are

$$w_{ij}^{(m)} = m! \,/\, \prod_{k=0,\, k\neq j}^{m} (x_{i+j} - x_{i+k}), \quad j = 0, 1, \ldots, m. \quad (10)$$

This formula follows from the explicit Lagrange form of the divided difference [10]. On a nonuniform grid, excessively close placement of some points reduces the denominators in (10) and may sharply increase the weights and, consequently, the noise amplification. Fornberg's algorithms are devoted precisely to computing such weights on arbitrary grids [11].

Theorem 1 (deterministic stability estimate of differential identification). Let $f \in C^m[a, b]$ and

$$|f^{(m)}(x) - C| \le \beta, \quad x \in [a, b], \quad C \neq 0. \quad (11)$$

Let the experimental error satisfy

$$\|\varepsilon\|_\infty \le \varepsilon_0. \quad (12)$$

Introduce the operator characteristic

$$\kappa_m = \|W_m\|_\infty = \max_i \sum_j |w_{ij}^{(m)}|. \quad (13)$$

Then for every local estimate

$$|D_i^{(m)} - C| \le r_m, \quad r_m = \beta + \kappa_m \varepsilon_0. \quad (14)$$

If $r_m < |C|$, then for criterion (7)

$$\delta_m/100 \le 2r_m \,/\, (|C| - r_m). \quad (15)$$

Proof. For exact data, (5) and (11) give $|D_{i,f^{(m)}} - C| \le \beta$. By (9), the error component equals $W_m\varepsilon$. From the properties of the matrix ∞-norm, $\|W_m\varepsilon\|_\infty \le \|W_m\|_\infty \|\varepsilon\|_\infty \le \kappa_m\varepsilon_0$. Hence (14) follows. Write $D_i^{(m)} = C + r_i$ with $|r_i| \le r_m$. For the mean value, $|\bar{D}^{(m)} - C| \le r_m$ also holds. Therefore $|D_i^{(m)} - \bar{D}^{(m)}| \le 2r_m$, and since $r_m < |C|$, $|D_i^{(m)}| \ge |C| - r_m > 0$. Applying these bounds to each term and averaging yields (15).

Corollary 1 (the general criterion γ and the 1/41 condition for 5%). If the relative threshold is expressed as a fraction γ, i.e. $\delta_m/100 \le \gamma$, then (15) yields the following sufficient condition:

$$\beta + \kappa_m\varepsilon_0 \le \gamma |C| \,/\, (2 + \gamma). \quad (16)$$

Indeed, with $q = r_m/|C|$, the inequality $2q/(1 - q) \le \gamma$ implies $q \le \gamma/(2 + \gamma)$. For $\gamma = 0.05$, $0.05/2.05 = 1/41$; therefore

$$\beta + \|W_m\|_\infty\varepsilon_0 \le |C|/41. \quad (17)$$

This is a **sufficient but not necessary** condition derived analytically from the 5% criterion. It is therefore natural for it to be conservative in practice.

Noise amplification for random errors

If the ε_i are independent with $E\varepsilon_i = 0$ and $\mathrm{Var}(\varepsilon_i) = \sigma^2$, then for each block

$$\mathrm{Var}(D_{i,\varepsilon}^{(m)}) = \sigma^2 \sum_j (w_{ij}^{(m)})^2. \quad (18)$$

Hence the local standard deviation is

$$\sigma_i^{(m)} = \sigma \cdot (\Sigma_j (w_{ij}^{(m)})^2)^{1/2}. \quad (19)$$

On a uniform grid with step h, (18) reduces, through the classical binomial formula, to

$$\mathrm{Var}(\Delta^m \varepsilon / h^m) = C(2m, m) \cdot \sigma^2 / h^{2m}, \quad (20)$$

where C(2m, m) is the binomial coefficient. Formula (20) shows explicitly why noise grows rapidly in high-order differentiation. This is consistent with the classical ill-posed nature of numerical differentiation of noisy data [12].

Polynomial reconstruction and approximation error

Let m be the first accepted order at the identification stage and

$$C = D^{(m)}. \quad (21)$$

The basic differential model is taken as

$$y^{(m)} = C. \quad (22)$$

Integrating it m times gives a polynomial of the form

$$P_m(x) = (C/m!)(x - x_0)^m + \Sigma_{j=0}^{m-1} (s_j/j!)(x - x_0)^j. \quad (23)$$

If the initial derivative conditions of the process are known, then $s_j = f^{(j)}(x_0)$. If they are not directly known, m independent calibration conditions can be taken from the experimental points; the subsequent iterative correction is applied to stabilize this first estimate.

Theorem 2 (error of differential polynomial approximation). Let $f \in C^m[a, b]$ and let (11) hold. Let the polynomial P_m satisfy

$$P_m^{(m)}(x) = C \quad (24)$$

and

$$P_m^{(j)}(x_0) = f^{(j)}(x_0), \quad j = 0, 1, \ldots, m - 1. \quad (25)$$

Then

$$|f(x) - P_m(x)| \le (\beta/m!)|x - x_0|^m. \quad (26)$$

In particular, if $L = b - a$, then

$$\|f - P_m\|_\infty \le \beta L^m/m!. \quad (27)$$

Proof. Let $g = f - P_m$. From (24)–(25), $g^{(m)}(x) = f^{(m)}(x) - C$ and $g^{(j)}(x_0) = 0$, with $|g^{(m)}| \le \beta$. Integrating m times gives

$$g(x) = (1/(m - 1)!) \int_{x_0}^{x} (x - t)^{m-1} g^{(m)}(t)\, dt. \quad ()$$

Taking absolute values and applying $|g^{(m)}| \le \beta$ yields (26). Inequality (27) follows from $|x - x_0| \le L$. ■

Corollary 2 (errors in the initial conditions). If $e_j = P_m^{(j)}(x_0) - f^{(j)}(x_0)$, then the same integral representation gives

$$|f(x) - P_m(x)| \le \Sigma_{j=0}^{m-1} (|e_j|/j!)|x - x_0|^j + (\beta/m!)|x - x_0|^m. \quad (28)$$

Thus, the reconstruction error comes from two sources: the error in the integration constants and the deviation of the m-th derivative from a constant.

Contractive iterative correction

Let the differential stage provide the model order m and an initial coefficient estimate $a^{(0)}$. To refine the coefficients, it is recommended to first rescale the argument to $z \in [-1, 1]$. The polynomial

$$P_m(z) = a_0 + a_1 z + \cdots + a_m z^m \quad (29)$$

is expressed with the design matrix

$$\Phi_{ij} = z_i^j, \quad i = 0, \ldots, n, \quad j = 0, \ldots, m. \quad (30)$$

The following iteration is applied:

$$a^{(k+1)} = a^{(k)} + \tau \Phi^T (y - \Phi a^{(k)}). \quad (31)$$

Iteration (31) is a classical Landweber/gradient-type iteration [9]. In this work, its role is not to select a new model type but to stably correct the initial coefficients obtained by the differential method.

Theorem 3 (contraction condition of the iterative correction). Let Φ have full column rank and

$$0 < \tau < 2/\lambda_{max}(\Phi^T \Phi). \quad (32)$$

Then the least-squares solution

$$a_{LS} = (\Phi^T\Phi)^{-1}\Phi^T y \qquad (33)$$

is the unique fixed point of (31), and

$$\|a^{(k)} - a_{LS}\|_2 \le q^k \|a^{(0)} - a_{LS}\|_2, \qquad (34)$$

where

$$q = \|I - \tau\Phi^T\Phi\|_2 = \max_j |1 - \tau\lambda_j| < 1. \qquad (35)$$

Proof. The vector a_{LS} satisfies the normal equation $\Phi^T(y - \Phi a_{LS}) = 0$. Subtracting (33) from (31) gives $a^{(k+1)} - a_{LS} = (I - \tau\Phi^T\Phi)(a^{(k)} - a_{LS})$. By (32), $|1 - \tau\lambda_j| < 1$ for all positive eigenvalues of $\Phi^T\Phi$. Hence (35) holds, and (34) follows by induction. ■

If $y = \Phi a^* + \varepsilon$, the iteration does not "remove" the experimental noise; it converges to a_{LS}. The residual coefficient error caused by the noise is bounded by

$$\|a_{LS} - a^*\|_2 \le \|\Phi^+\|_2 \|\varepsilon\|_2. \qquad (36)$$

This fundamentally separates the computational error from the data error: the former decays as q^k, whereas the latter is limited by the quality of the data and the conditioning of the matrix.

Practical algorithm

For hand calculation or a simple spreadsheet, the algorithm can be written as follows.

Input: (x_i, y_i), $x_0 < \cdots < x_n$; criterion $\gamma = 0.05$; minimum number of remaining local estimates $N_{min} = 4$; measurement error bound ε_0, if known.

Step 1. Compute $D_i^{(1)} = (y_{i+1} - y_i)/(x_{i+1} - x_i)$.

Step 2. If $\delta_1 \le 5$ %, accept m = 1; otherwise compute $D^{(2)}$, $D^{(3)}$, … successively by (2).

Step 3. At the first order for which $\delta_m \le 5$ % and $N_m \ge N_{min}$, select m as the candidate order and set $C = D^{(m)}$.

Step 4. Compute W_m and $\kappa_m = \|W_m\|_\infty$ from the nonuniform-grid weights. If reliable estimates of β and ε_0 are available, certify stability by (16), or by (17) for the 5% case.

Step 5. Integrate $y^{(m)} = C$ m times and construct P_m from physical initial conditions or calibration points.

Step 6. If higher coefficient accuracy is required, rescale x to [−1, 1] and perform the iterative correction (31) under condition (32).

Step 7. Validate the model on points not used for identification or in a repeated experiment. This step is necessary to control the risk of interpolation/overfitting.

Numerical experiments and computational setup

Three computational experiments were performed: (i) a nonuniform cubic example with a 100,000-run Monte-Carlo test of the stability condition; (ii) a nonpolynomial synthetic process; and (iii) a test of the contractive iterative correction. All computations were carried out in Python 3 using NumPy 2.3.5, SciPy 1.17.0 and Matplotlib 3.10.8. The main Monte-Carlo series consisted of 100,000 realizations for each noise level, and the base seed of the pseudorandom generator was set to 20260903. All figures were recomputed for this paper; the synthetic examples are not to be interpreted as real fermentation data.

3. Results and discussion

Experiment 1: nonuniform cubic and a 100,000-run Monte-Carlo test

The following nonuniform grid was used:

$$x = (0, 0.4, 1.1, 1.7, 2.6, 4.0, 5.2, 7.0). \qquad (37)$$

The test function was

$$f(x) = 2x^3 - x^2 + 3x + 1. \qquad (38)$$

Its third derivative is $f''' = 12$. The results of the recursive computation are given in Table 1.

Table 1. Successive differential estimates on the nonuniform grid

Order	D^(m) values	δ_m, %
m = 1	2.92; 5.12; 12.14; 26.84; 62.72; 121.48; 215.68	559.531
m = 2	4.00; 10.80; 19.60; 31.20; 45.20; 62.80	156.006
m = 3	12.00; 12.00; 12.00; 12.00; 12.00	0.000

The algorithm stops for the first time at m = 3 and gives C = 12. The operator norm computed for this grid is

$$\|W_3\|_\infty = 42.0168067. \quad (39)$$

For an exact cubic, β = 0. Hence, by (17), the error bound that deterministically guarantees the 5% criterion is

$$\varepsilon_{crit} = 12 / (41 \cdot 42.0168067) = 0.00696585. \quad (40)$$

In the Monte-Carlo test, independent errors

$$\varepsilon_i \sim U[-a, a] \quad (41)$$

were added to each point, and 100,000 realizations were computed for each value of a (Table 2, Figure 1).

Table 2. Fulfilment of the 5% criterion in 100,000 Monte-Carlo runs

a/ε_crit	a	δ_3 ≤ 5%, %	Median δ_3, %	95th percentile, %
1.0	0.006966	100.000	0.245	0.543
2.0	0.013932	100.000	0.491	1.088
5.0	0.034829	100.000	1.230	2.730
7.5	0.052244	98.946	1.842	4.101
10.0	0.069659	91.795	2.455	5.510
15.0	0.104488	70.331	3.695	8.442
20.0	0.139317	51.030	4.931	11.489

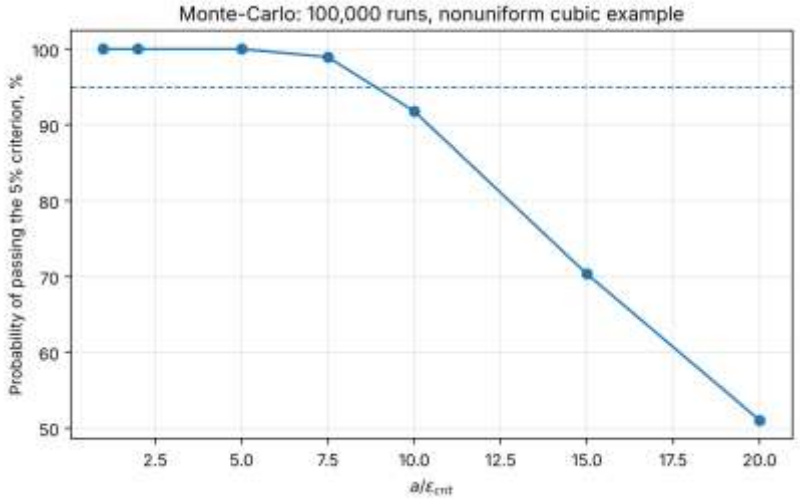


Figure 1. Monte-Carlo results for the nonuniform cubic example. The dashed line indicates the 95% level.

The result confirms that condition (17) is sufficient: at a = ε_{crit}, all 100,000 runs passed the 5% criterion. At the same time, the condition is conservative: on this grid, all realizations passed even up to $5\varepsilon_{crit}$. Therefore, the value 1/41 should be interpreted not as an empirical "exact threshold" but as a safe bound that provides a mathematical guarantee.

Experiment 2: a nonpolynomial synthetic process

To test a smooth process that is not exactly equal to any standard function, the following synthetic function was used:

$$f(x) = 20 + 5x - 0.8x^2 + 0.06x^3 + 0.2\sin(0.4x), \quad (42)$$

with the nonuniform points

$$x = (0, 1.2, 2.8, 4.5, 6, 8, 10.5, 13, 15.5, 18). \quad (43)$$

This example does not represent real biotechnological data; it describes a conditional process with a nonuniform experimental design and a small harmonic component not accounted for in the model.

The computation gave

$$\delta_1 = 182.712\ \%, \quad \delta_2 = 541.874\ \%, \quad \delta_3 = 1.872\ \%. \quad (44)$$

The algorithm therefore stops at m = 3. The mean of the third-order estimates is

$$C_3 = 0.361205. \quad (45)$$

When the integration constants were determined from the first three calibration points, the following polynomial was obtained:

$$P_3(x) = 20 + 5.087141x - 0.808722x^2 + 0.060201x^3. \quad (46)$$

This polynomial is not identical to the original function (42). Nevertheless, at the experimental points the range-normalized RMSE was

$$NRMSE = 0.10285\ \%, \quad (47)$$

and the ratio of the maximum absolute error to the range was 0.19912% (Figure 2).

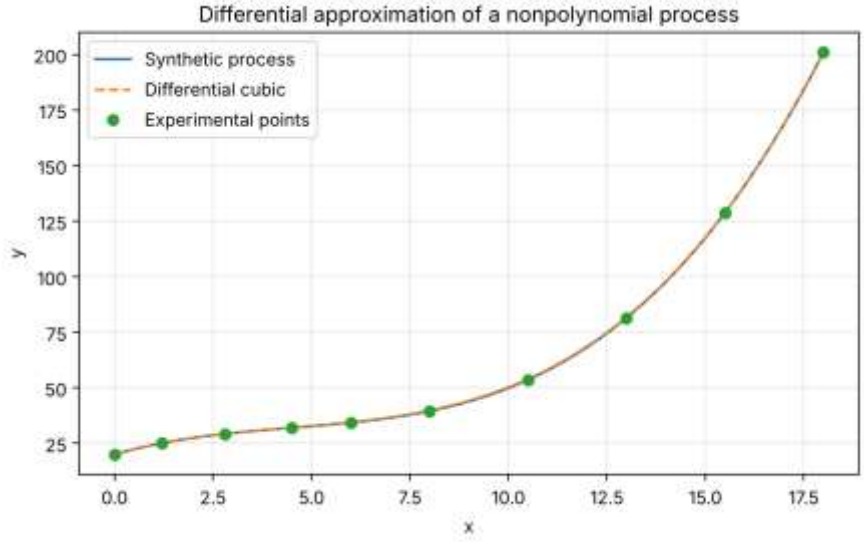


Figure 2. Nonpolynomial synthetic process and the cubic reconstructed by the differential method.

For comparison, several standard functions were fitted to the same 10 points by ordinary and nonlinear least squares (Table 3). This comparison does not claim universal superiority of any method; its sole purpose is to show the importance of the a priori choice of model form in this conditional example.

Table 3. Model errors for the synthetic process

Model	Range-normalized NRMSE, %
Linear	13.656
Quadratic	4.716
Exponential + const	1.113
Logistic + const	1.047
Differential cubic	0.103

3.3. Experiment 3: contractive iteration

In the cubic example (37)–(38), the argument x was normalized to [−1, 1]. In this basis, the exact coefficients are

$$a^* = (85, 243.25, 245, 85.75)^T.(48)$$

The eigenvalues of the matrix $\Phi^T\Phi$ are

$$0.190803,\ 0.886041,\ 5.145307,\ 11.205082.(49)$$

As a symmetric choice that accelerates contraction,

$$\tau = 2 / (\lambda_{min} + \lambda_{max}) = 0.175502(50)$$

was taken. Then

$$q = \| I - \tau\Phi^T\Phi \|_2 = 0.966514 < 1.(51)$$

The initial coefficients were deliberately perturbed by about 20–30%:

$$a^{(0)} = a^* \odot (1.3, 0.7, 1.25, 0.8).(52)$$

Table 4. Decay of the iteration error

k	Exact data: error w.r.t. a_LS	Noisy data: error w.r.t. a_LS	Noisy data: error w.r.t. true coefficients
0	100.106	100.106	100.106
10	59.797	59.792	59.797
25	35.621	35.618	35.632
50	15.200	15.199	15.207
100	2.769	2.768	2.776
200	0.09185	0.09184	0.10519
300	0.00305	0.00305	0.03543

In the noisy test, one fixed realization of $\varepsilon_i \sim U[-5\varepsilon_{crit}, 5\varepsilon_{crit}]$ was used. At step 300, the iteration error fell to 0.00305 relative to the fixed a_{LS}, but remained at 0.03543 relative to the true coefficients (Figure 3). This illustrates the physical meaning of (36): the computational error decays, whereas the error originating from data noise remains.

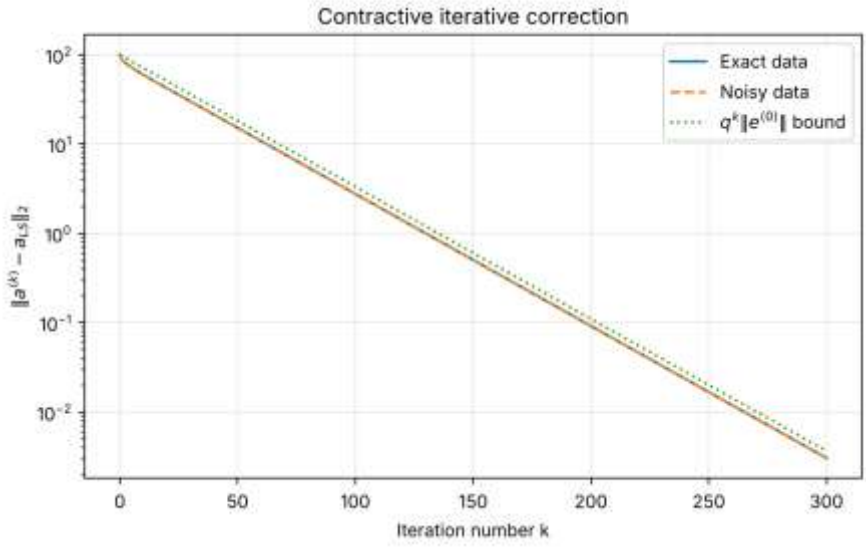


Figure 3. Geometric decay of the coefficient error in the iterative correction.

4. Discussion

Several fundamental conclusions follow from the results obtained.

First, a nonuniform step is not an obstacle in itself. The recursive formula (2) works for any strictly increasing x_i and coincides with classical divided differences. The problem

lies not in the nonuniformity of the steps as such but in the excessive closeness of individual points, which enlarges the weights (10). Therefore, $\|W_m\|$ serves as a natural stability indicator of the grid geometry.

Second, the 5% criterion should be interpreted as the "true polynomial degree" only when the function is itself a polynomial. For a general smooth process, the criterion indicates differential stability on the given interval and is linked to the polynomial approximation error through Theorem 2. Experiment 2 illustrates the meaning of Theorem 2 in practice: the true function need not be a polynomial; the essential condition is that the derivative of a certain order remains sufficiently stable around a constant on the chosen interval.

Third, high-order differentiation amplifies noise. Consequently, m cannot be increased without limit on the basis of δ_m alone. In practice, it is recommended to keep at least four local estimates, to estimate ε_0 from instrument specifications or repeated experiments, and to compute κ_m. If condition (16) is not satisfied, this does not mean that the result is "rejected"; it means only that no deterministic guarantee is available.

Fourth, the iterative correction does not replace differential identification. Iteration (31) is equivalent to the gradient iteration of the least-squares functional, so it is scientifically important to present it as a separate classical stage. The differential part provides the model order and the first approximation, whereas the Landweber part damps the initial coefficient error and fits the model to the data within the given model class.

Fifth, the 100,000 Monte-Carlo runs in this study do not prove the theoretical condition; the proof is given in Theorem 1. The Monte-Carlo test only shows how conservative the condition is in practice and how the acceptance probability decreases when the errors exceed the bound.

Comparison with previous work

Banerjee proposed a method for selecting the best polynomial order on equally spaced noisy data using finite differences and a t-test, followed by re-estimation of the coefficients. Omelchenko and Fedorov showed that, for equally spaced observations, finite differences of a single order can provide sufficient statistics for polynomial regression coefficients, and they studied the influence of rounding error. Fornberg [14] provided general algorithms for computing finite-difference weights on arbitrary grids. Classical information criteria such as that of Akaike choose the model order by a penalized likelihood, whereas the criterion used here relies on the differential stability of the data. Local smoothing and regularized differentiation [15] address the noise problem but do not by themselves select a polynomial order. For this reason, the possible novelty of the present paper does not lie in any of these individual elements. The present paper continues the authors' earlier work [16] by treating nonuniformly spaced data with random errors and by supplying a deterministic stability estimate.

The distinguishing construction of this work is the following chain: selecting the first stable order of scaled divided differences on a nonuniform grid by the γ-criterion; deriving a deterministic stability certificate based on the operator norm and directly tied to that criterion; reconstructing the approximation by integrating $y^{(m)} \approx C$; and, when necessary, adding a contractive iterative correction. A targeted literature search did not reveal a formulation equivalent to this complete combination; however, no claim of absolute priority is made.

Limitations and future research

This study has several limitations. First, the 5% value is a practical threshold rather than a universal physical constant; in other fields, γ should be chosen according to data accuracy and the cost of a wrong decision. For this reason, Theorem 1 is formulated for a general γ.

Second, the bound β on the deviation of the true m-th derivative from a constant may not be directly known for real data. Estimating it from local estimates, repeated experiments or domain knowledge is a separate statistical problem.

Third, on a nonuniform grid, κ_m can become very large at high orders. In such cases, instead of blindly differentiating the data, one should redesign the experimental points, average repeated measurements, or use regularized differentiation [17] or local smoothing methods [18].

Fourth, a historical test on real biotechnological or fermentation data was not included in this paper because the original table is not available. If it is found, it can be added as an external validation separate from the synthetic examples.

5. Conclusions

A differential identification scheme based on successive scaled divided differences has been formalized for nonuniform experimental data. Its main results are as follows:

- the recursion $D_i^{(m)} = m![x_i, \ldots, x_{i+m}]y$ provides local estimates of the m-th derivative on a nonuniform grid;
- for a general criterion γ, the deterministic sufficient stability condition $\beta + \|W_m\|_\infty \varepsilon_0 \le \gamma|C|/(2+\gamma)$ has been obtained;
- for the 5% case, this condition becomes $\beta + \|W_m\|_\infty \varepsilon_0 \le |C|/41$;
- it has been shown that, when the m-th derivative is close to a constant, the error of the polynomial reconstructed by integration is bounded by $\beta|x - x_0|^m/m!$;
- the contractive Landweber-type correction geometrically damps the computational and initial coefficient errors but does not remove experimental noise;
- the 100,000-run Monte-Carlo test showed that the theoretical condition is sufficient and conservative;
- for the nonpolynomial synthetic process, the cubic model selected by the differential criterion gave an NRMSE of 0.103%.

Thus, the main value of the method lies not in "declaring any process a polynomial" but in the ability, when the functional form is not known in advance, to select a local polynomial model based on differential stability, to assess its reliability through the operator norm, and to correct its coefficients stably.

REFERENCES

[1] NIST Digital Library of Mathematical Functions, "§3.3 Interpolation: Divided Differences and Newton's Interpolation Formula," National Institute of Standards and Technology. [Online]. Available: https://dlmf.nist.gov/3.3. [Accessed: Sep. 3, 2026].

[2] F. B. Hildebrand, *Introduction to Numerical Analysis*, 2nd ed. New York, NY, USA: Dover Publications, 1974.

[3] B. Fornberg, "Generation of finite difference formulas on arbitrarily spaced grids," *Mathematics of Computation*, vol. 51, no. 184, pp. 699–706, 1988, doi: 10.1090/S0025-5718-1988-0935077-0.

[4] B. Fornberg, "Calculation of weights in finite difference formulas," *SIAM Review*, vol. 40, no. 3, pp. 685–691, 1998, doi: 10.1137/S0036144596322507.

[5] A. Banerjee, "The method of finite difference regression," *Open Journal of Statistics*, vol. 8, pp. 49–68, 2018, doi: 10.4236/ojs.2018.81005.

[6] A. V. Omelchenko and O. V. Fedorov, "Polynomial regression coefficients estimation in finite differences space," in *Proc. 25th Int. Conf. Radioelektronika*, Pardubice, Czech Republic, 2015, pp. 257–260, doi: 10.1109/RADIOELEK.2015.7129024.

[7] A. G. Ramm and A. B. Smirnova, "On stable numerical differentiation," *Mathematics of Computation*, vol. 70, no. 235, pp. 1131–1153, 2001, doi: 10.1090/S0025-5718-01-01307-2.

[8] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed. Philadelphia, PA, USA: SIAM, 2002, doi: 10.1137/1.9780898718027.

[9] L. Landweber, "An iteration formula for Fredholm integral equations of the first kind," *American Journal of Mathematics*, vol. 73, no. 3, pp. 615–624, 1951, doi: 10.2307/2372313.

[10] L. N. Trefethen, *Approximation Theory and Approximation Practice*, extended ed. Philadelphia, PA, USA: SIAM, 2019, doi: 10.1137/1.9781611975949.

[11] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964, doi: 10.1021/ac60214a047.

[12] K. Ahnert and M. Abel, "Numerical differentiation of experimental data: Local versus global methods," *Computer Physics Communications*, vol. 177, no. 10, pp. 764–774, 2007.

[13] R. Chartrand, "Numerical differentiation of noisy, nonsmooth data," *ISRN Applied Mathematics*, vol. 2011, Art. no. 164564, 2011, doi: 10.5402/2011/164564.

[14] I. Knowles and R. J. Renka, "Methods for numerical differentiation of noisy data," *Electronic Journal of Differential Equations, Conference*, vol. 21, pp. 235–246, 2014.

[15] M. Hanke and O. Scherzer, "Inverse problems light: Numerical differentiation," *The American Mathematical Monthly*, vol. 108, no. 6, pp. 512–521, 2001.

[16] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996.

[17] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974, doi: 10.1109/TAC.1974.1100705.

[18] S. B. Ochilov, G. D. Khasanova, and O. K. Khudayberdieva, "Method for constructing correlation dependences for functions of many variables used finite differences," *The American Journal of Management and Economics Innovations*, vol. 3, no. 5, pp. 46–52, 2021.