Finite order time-series representations of discrete data from any finite degree univariate polynomial

Professor Asoke K Nandi

FREng FIEEE PhD (Cambridge)

asoke.nandi@brunel.ac.uk

Polynomials

- The word polynomial was first used in 17th century (source: OED and Wikipedia).
- An example of a polynomial of a single variable (indeterminate) is $(x^2 - 4x + 7)$. The degree is 2.
- An example with three indeterminates is $x^3 + 2xyz^2 - yz + 1$. The degree is 4.
- **In this presentation, we are only considering univariate (single variable) polynomials.**

Summary

- Every polynomial of degree q can be represented perfectly as a time-series of order q.
- A q-degree polynomial requires (q + 1) unknowns – q coefficients and one constant.
- A q-order time-series requires only one unknown – one constant.
- From polynomial data, the degree of the polynomial, the additive noise, and the coefficient corresponding to the highest degree can be deduced without using a polynomial model.

Journal Papers

A K Nandi, “Data modelling with polynomial representations and autoregressive representations, and their connections”, IEEE Access, vol. 8, pp. 110412-110424, DOI: 10.1109/ACCESS.2020.3000860, 2020.

A K Nandi, “Model order selection from noisy polynomial data without using any polynomial coefficients”, IEEE Access, vol. 8, pp. 130417-130430, DOI: 10.1109/ACCESS.2020.3008527, 2020.

A K Nandi, “Degree and noise power estimation from noisy polynomial data via AR modelling”, Digital Signal Processing, vol. 114, 103071 (13 pages), DOI: 10.1016/j.dsp.2021.103071, 2021.

A K Nandi, “Theoretical and experimental results on time-series representations of polynomials”, Frontiers in Signal Processing, vol. 5, 1582043 (13 pages), DOI: 10.3389/frsip.2025.1582043, 2025.

Introduction

An accurate estimation of the degree of a polynomial from data is important in many types of applications.

This significantly improves the performance in

- Modelling
- Detection
- Estimation
- Prediction

It is particularly important in Data Science.

Introduction

For more than a century, determining the degree of a polynomial has been based on fitting data directly to polynomial representations.

A pioneering paradigm has been recently developed for uniformly sampled polynomial data. This allows the determination of the degree of a polynomial by fitting data to a corresponding “time-series” representation.

It needs no polynomial coefficients, deviating from the convention that existed for more than a century.

Polynomial – with noise

A polynomial of degree q in uniformly sampled discrete time can be written as

$$y(n) = \sum_{i=0}^{q} c(i)\, n^{i}$$

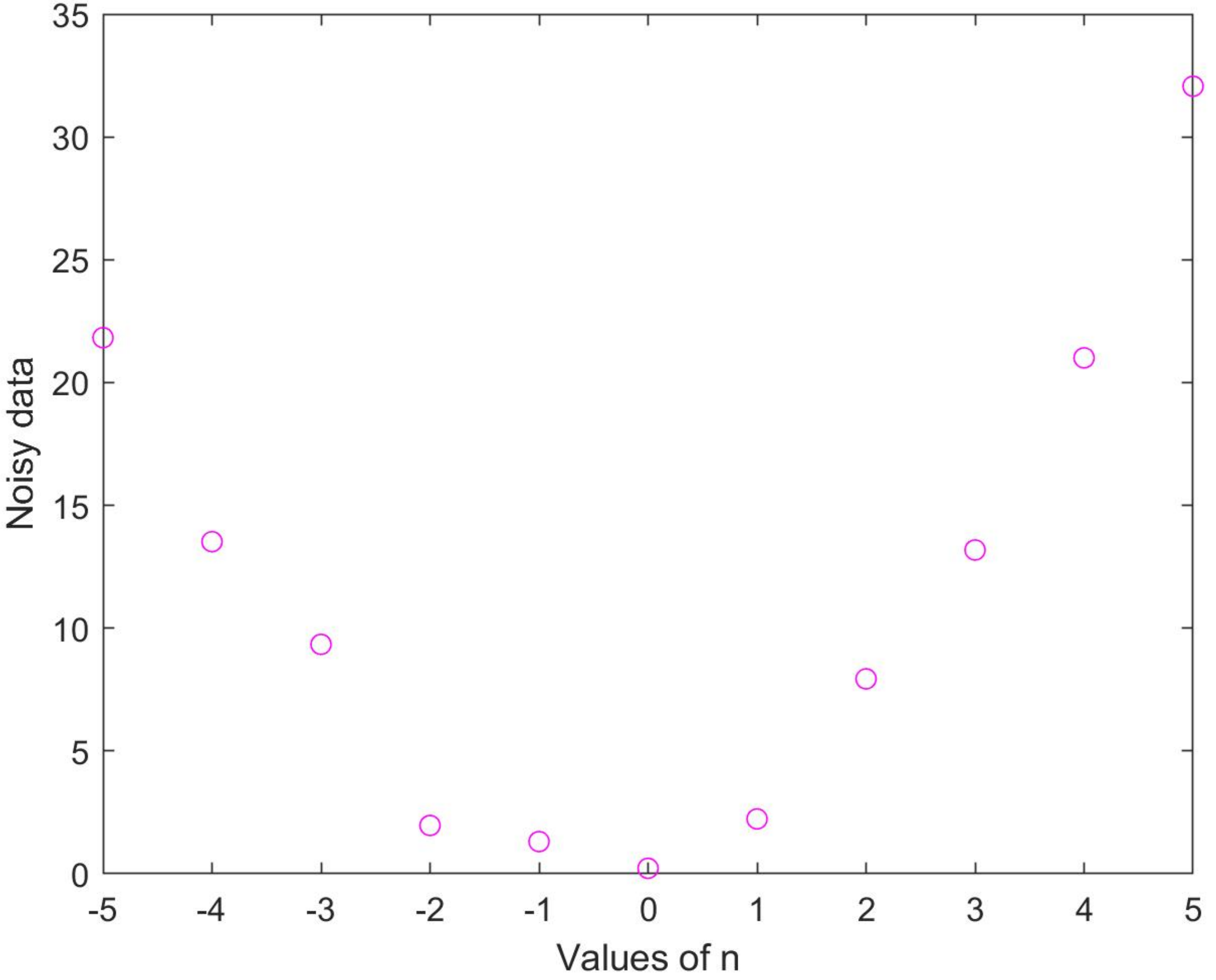
Thus, a set of real-valued noisy data from polynomials in uniformly sampled discrete "time", can be represented by

$$x(n) = \sum_{i=0}^{q} c(i)\, n^{i} + e(n)$$

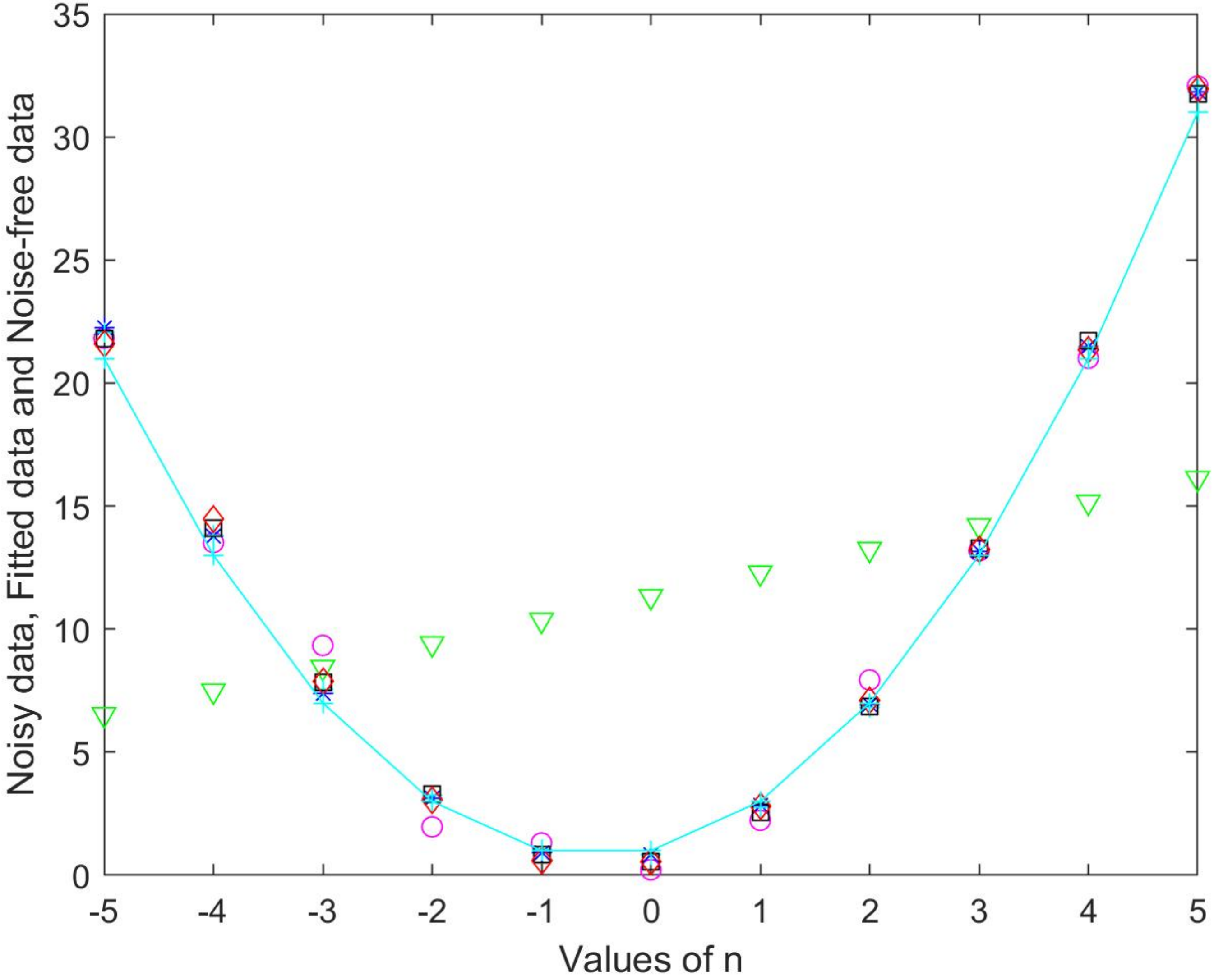
where $e(n)$ represents noise.

For this presentation, noise is assumed to additive, zero-mean, and i.i.d.

Quadratic Polynomial – with noise



Quadratic Polynomial – with noise



Polynomial – with noise

Observations (in this case of a correct polynomial degree is 2):

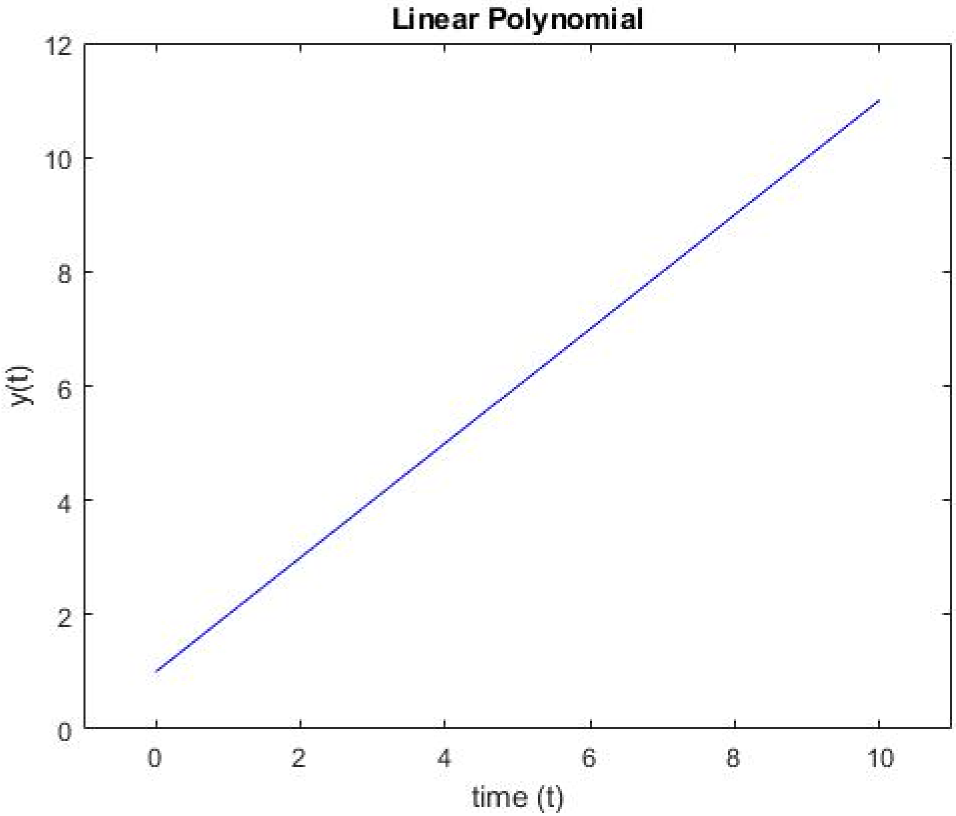
1) All the fits are carried out with Least-squares method.
2) A polynomial of degree 1 did not fit the data well.
3) Polynomials of degree 2, 3, 4, and 5 fitted the data well.
4) How do we tell the degree of the underlying polynomial? One way is to look at the root mean square error (RMSE) of the fits and to choose the degree corresponding to the smallest RMSE.
5) RMSE(degree = 1) = 9.29; RMSE(degree = 2) = 0.82; RMSE(degree = 3) = 0.81; RMSE(degree = 4) = 0.78; and RMSE(degree = 5) = 0.74.
6) The smallest RMSE corresponds to degree = 5.

   But this is not the same as the true degree of 2.

Generally, RMSE gets smaller as the degree of the fitting polynomial increases.

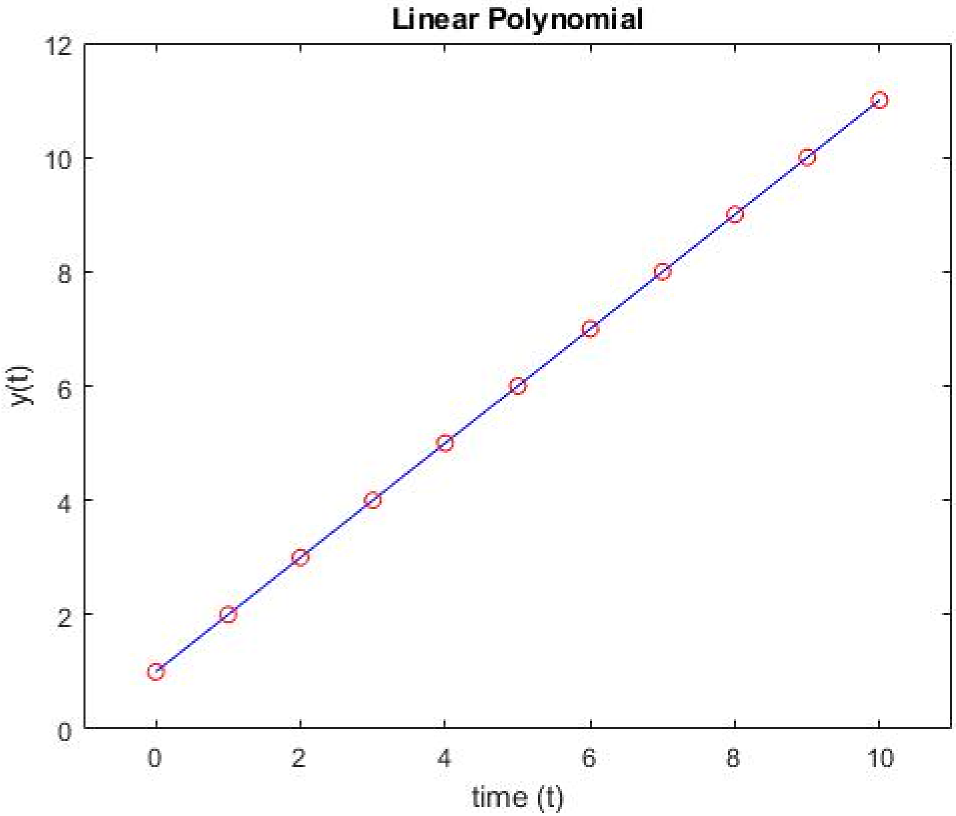
Introduction

$$y(t) = t + 1$$



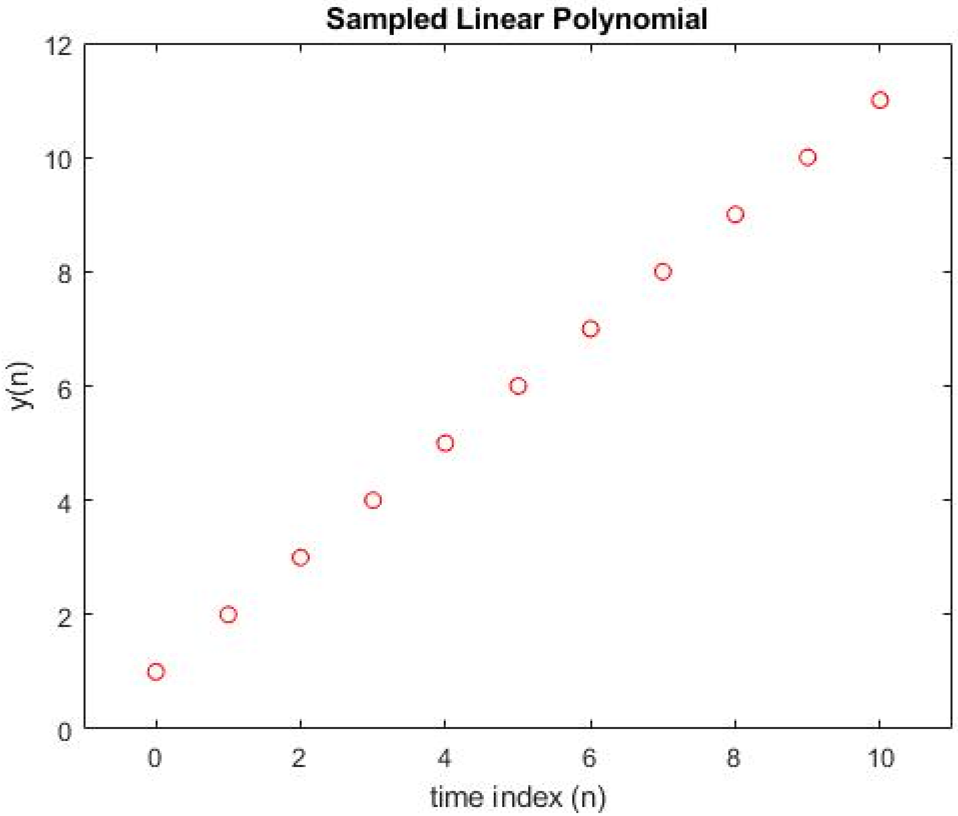
Introduction

$$y(t) = t + 1$$



No noise

$$y(n) = n + 1$$



Introduction

Polynomial of degree 1:

$$y(n) = c(1)\, n + c(0)$$

AR time-series of order 1:

$$y(n) = a(1)\, y(n-1) + \mu$$

$c(1)$ and $c(0)$ can take any real values.

Yet $a(1) = 1$ and $\mu = c(1)$ for every linear polynomial.

Polynomial – no noise

Example for q = 1:

Polynomial of degree 1:

$$y(n) = c(1)\,n + c(0)$$

PROOF: $y(n-1) = c(1)\,(n-1) + c(0) = c(1)n - c(1) + c(0)$

$$= c(1)n + c(0) - c(1) = y(n) - c(1)$$

Therefore, $y(n) = y(n-1) + c(1)$

AR time-series of order 1:

$$y(n) = a(1)\,y(n-1) + \mu$$

$a(1) = 1$ and $\mu = c(1)$

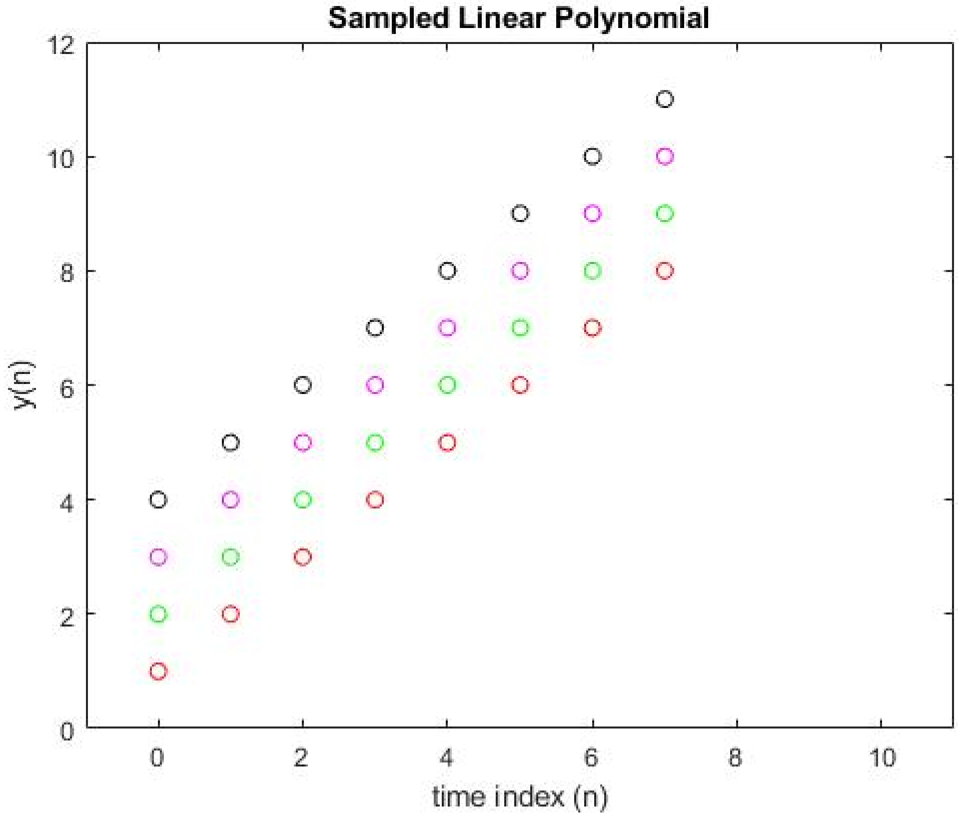
Therefore, in time-series formulation there is only one unknown (μ),

but in polynomial formulation there are 2 unknowns.

Introduction

$y(n) = n + 1$ $y(n) = n + 2$

$y(n) = n + 3$ $y(n) = n + 4$



$$y(n) = y(n-1) + 1$$

Introduction

Polynomial of degree 2:

$$y(n) = c(2)\, n^2 + c(1)\, n \; + c(0)$$

AR time-series of order 2:

$$y(n) = a(2) y(n-2) + a(1) y(n-1) + \mu$$

$c(2), c(1)$, and $c(0)$ can take any real values.

Yet $a(2) = -1$, $a(1) = 2$, and $\mu = 2c(2)$ for every quadratic polynomial.

Polynomials and AR parameters

Information for polynomials of different degrees

Degree of polynomial *q*	AR parameters			
	$a(1)$	$a(2)$	$a(3)$	μ
1	1			$c(1)$
2	2	-1		$2c(2)$
3	3	-3	1	$6c(3)$

Polynomial – no noise

A polynomial of degree q in continuous time can be written as

$$y(t) = \sum_{i=0}^{q} b(i)\, t^i$$

For uniformly sampled discrete time, the continuous time, t, is represented as $t = nT$, where n is an integer and T is the sampling period. So, we can write

$$y(nT) = \sum_{i=0}^{q} b(i)\, (nT)^i$$

This can be written as

$$y(n) = \sum_{i=0}^{q} c(i)\, n^i\,,$$

where $c(i) = b(i)\, T^i$.

Without loss of generality, let $T = 1$.

Polynomial – no noise

$$y(n) = \sum_{i=1}^{q} c(i)\, n^{i} + c(0),$$

All noise-free data from uniformly sampled polynomials of finite degree *q* can be perfectly represented by an autoregressive “time-series” model of order *q* and a constant, such that

$$y(n) = \sum_{i=1}^{q} a(i)\, y(n-i) + \mu$$

with $\boldsymbol{a(i) = (-1)^{i+1}\binom{q}{i}}$, for $i = 1, 2, \ldots, q$, and $\mu = \boldsymbol{c(q)\,(q!)}$

Therefore, in “time-series” formulation there is only one unknown (μ),

but in polynomial formulation there are (q + 1) unknowns.

Summary of four cases.

Data source	Number of data	Degree of polynomial	Polynomial representation	Time-series representation	Comments
Case I Polynomial	35	3	At $(p+1)=4$, RMS prediction error of 10^{-12}.	At $q=4$, RMS prediction error of 10^{-9}.	Both offer very low prediction errors.
Case II Sine wave	35	∞	At $(p+1)=1$, RMS prediction error of 3.8.	At $q=2$, RMS prediction error of 10^{-15}.	Time-series offers significantly better prediction.
Case III Non-polynomial	35	∞	At $(p+1)=3$, RMS prediction error of 10^{6}.	At $q=4$, RMS prediction error of 10^{-8}.	Time-series offers significantly better prediction.
Case IV Inverse polynomial	35	∞	At $(p+1)=3$, RMS prediction error of 6.7.	At $q=7$, RMS prediction error of 10^{-8}.	Time-series offers significantly better prediction.

Four lessons

- All polynomials of degree q can be represented perfectly by an all-pole filter with q repeated roots (or poles) at $z = +1$.
- Data representable by finite order all-pole filters, independent of whether they are from finite degree or infinite degree polynomials, can be described by a finite order AR time-series.
- A finite order time-series can represent perfectly infinitely many polynomials of infinite degree.
- Time-series are more general than polynomials.

Computer experiments

- 5 polynomial degrees (q = 1, 2, 3, 4, and 5)
- 5 noise standard deviations (σ = 1, 2, 4, 7, and 10)
- 3 types of additive noise (Gaussian, Uniform, and Exponential)
- 10,000 realisations of each of these 75 combinations
- **A total of 750,000 realisations**

5 polynomials (for data generation)

$$y(n) = 2n + 1$$

$$y(n) = 4n^2 + 2n + 1$$

$$y(n) = 8n^3 + 4n^2 + 2n + 1$$

$$y(n) = 16n^4 + 8n^3 + 4n^2 + 2n + 1$$

$$y(n) = 32n^5 + 16n^4 + 8n^3 + 4n^2 + 2n + 1$$

$$x(n) = y(n) + e(n)$$

Degree = q (for estimation)

Degree = 1, 2, 3, 4, and 5

$$x(n) = x(n-1) + \mu$$

$$x(n) = 2x(n-1) - x(n-2) + \mu$$

$$x(n) = 3x(n-1) - 3x(n-2) + x(n-3) + \mu$$

$$x(n) = 4x(n-1) - 6x(n-2) + 4x(n-3) - x(n-4) + \mu$$

$$x(n) = 5x(n-1) - 10x(n-2) + 10x(n-3) - 5x(n-4) + x(n-5) + \mu$$

$$x(n) = y(n) + e(n)$$

Degree = q (estimation)

$$\mu = c(q)(q!)$$

$$< \sigma_{\mu(q)} >^2 = \left((1)^2 + \sum_{i=1}^{q} a(i)^2 \right) \sigma^2$$

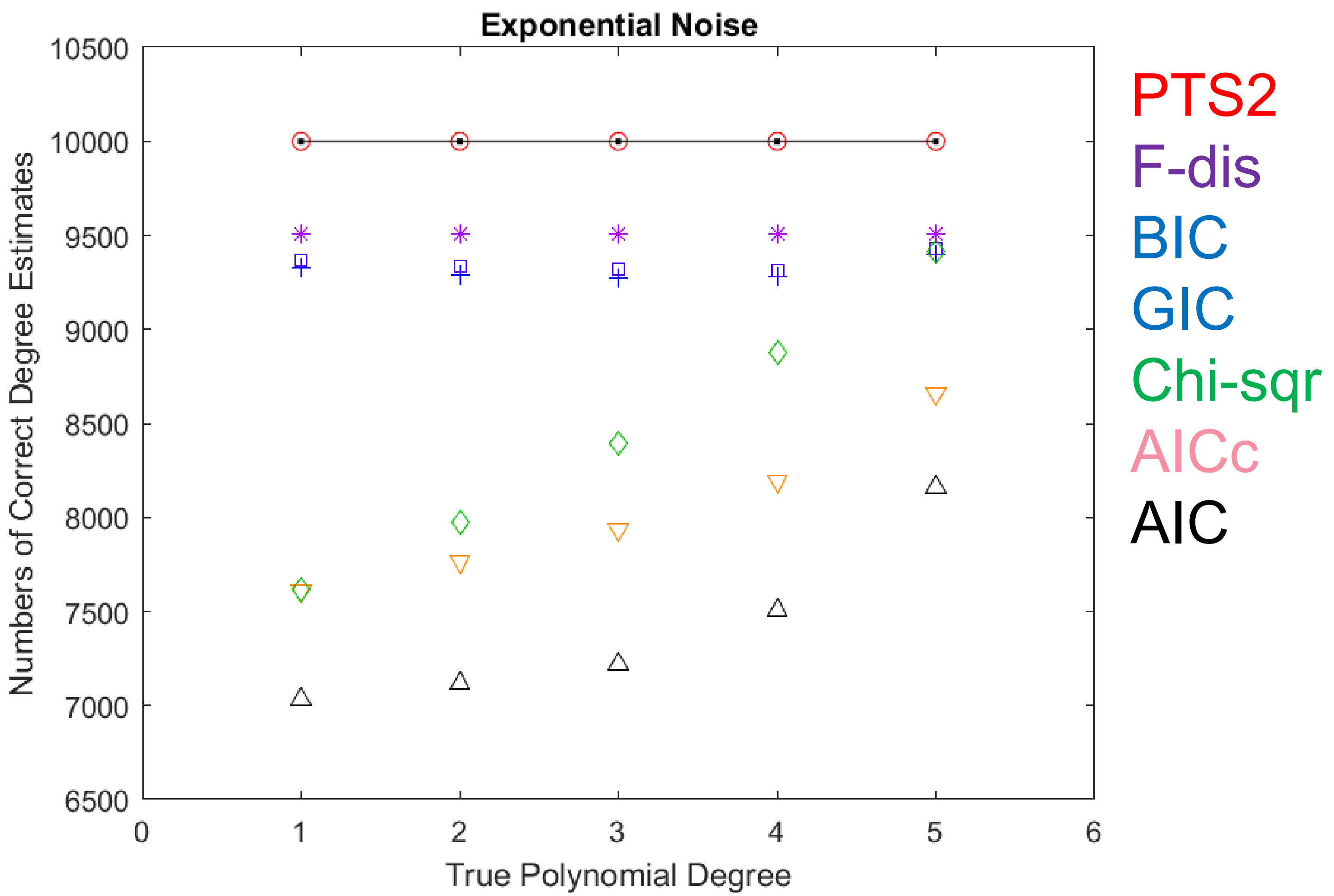
Estimate μ and $\sigma^2_{\mu(q)}$ from data $\boldsymbol{x}(\boldsymbol{n})$

$$PTS2(q) = minimise\ (\sigma^2_{\mu(q)})$$

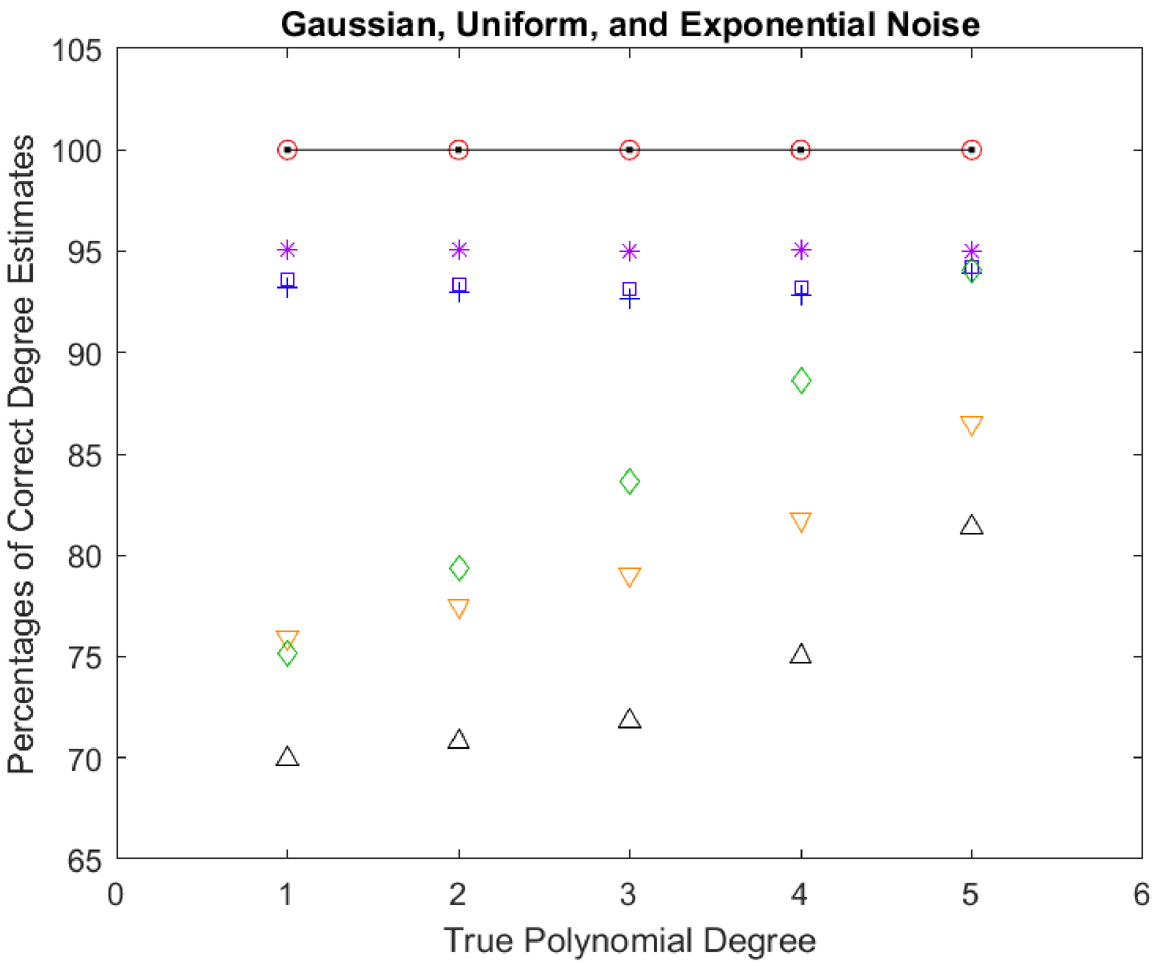
$$c(q) = \mu/(q!)$$

$$\sigma^2 = < \sigma_{\mu(q)} >^2 / \left((1)^2 + \sum_{i=1}^{q} a(i)^2 \right)$$

Number of correct degree estimates

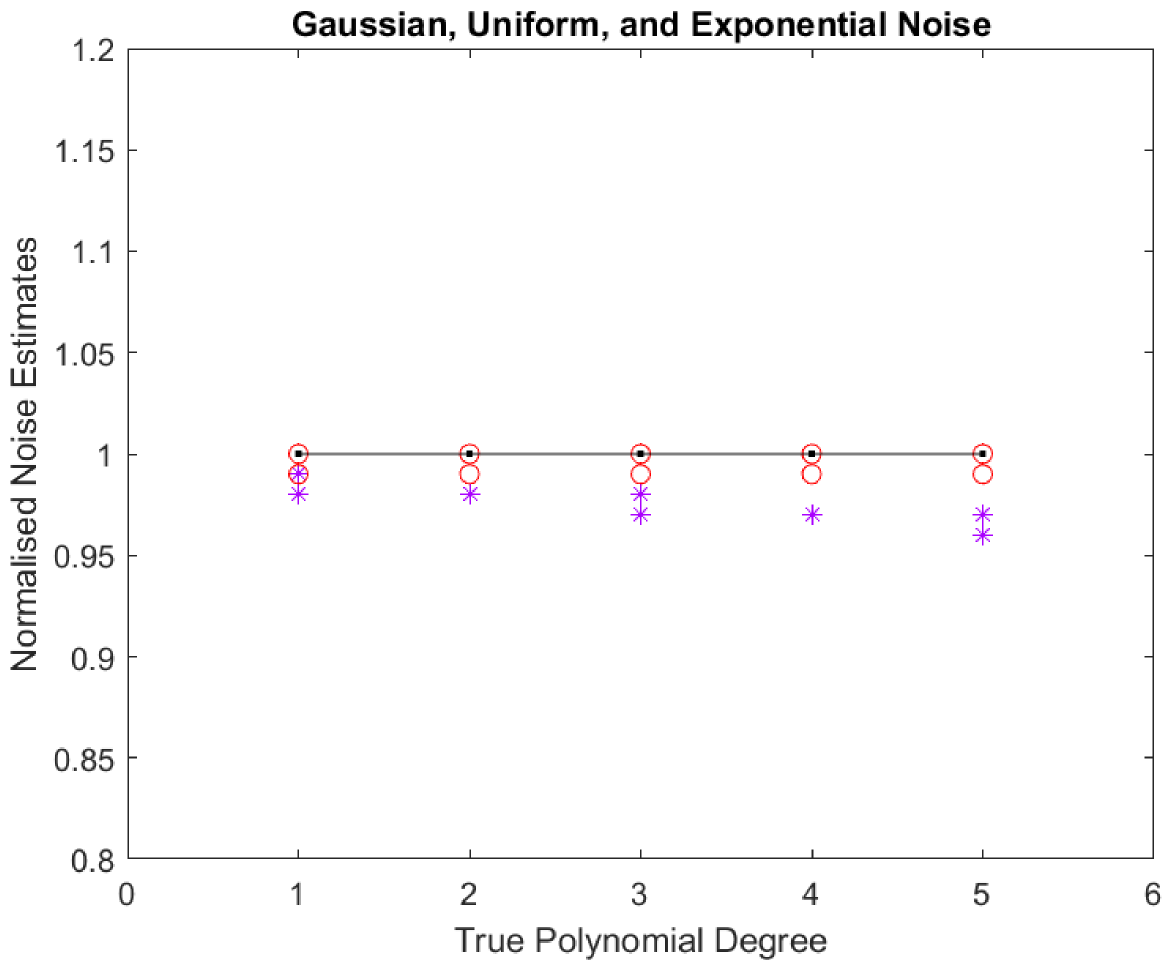


Percentage of correct degree estimates



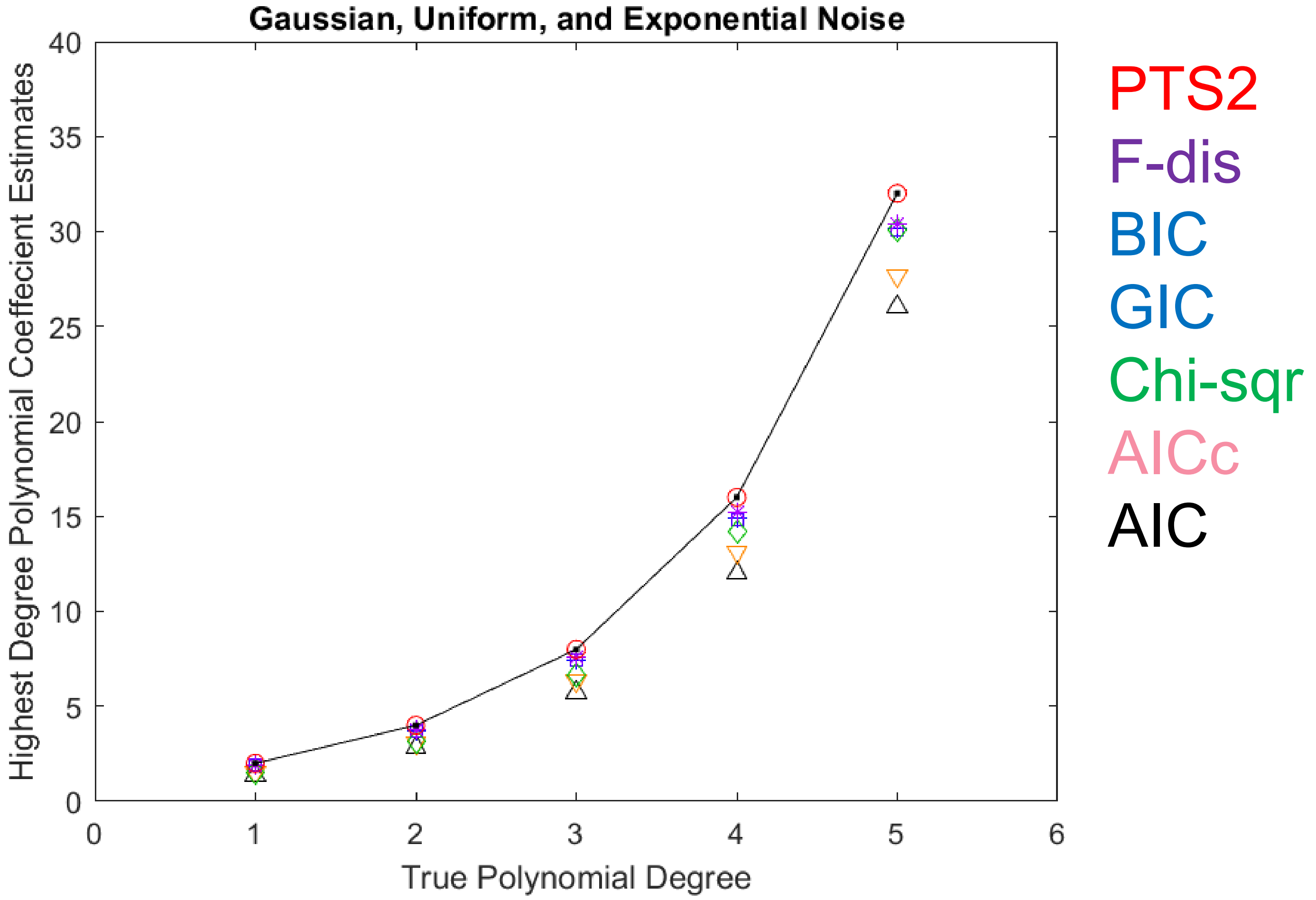
PTS2
F-dis
BIC
GIC
Chi-sqr
AICc
AIC

Normalised noise estimates

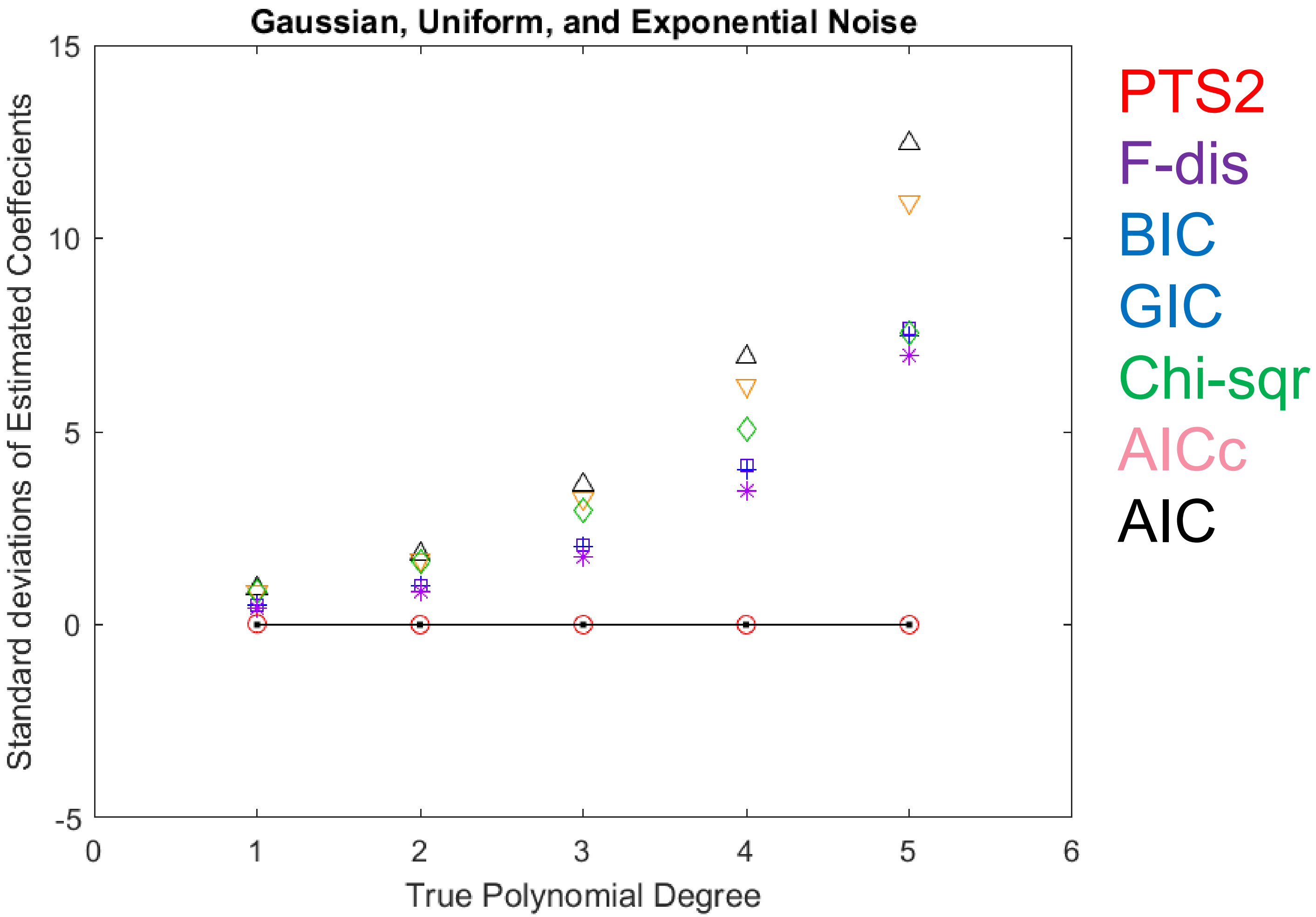


PTS2
F-dis

Highest degree polynomial coefficient estimates



Standard deviation of estimated coefficients



Closing remarks

- Determining the degree of a polynomial so far has been based on fitting noisy polynomial data directly using polynomial representations.

- Polynomials of degree q with (q + 1) unknowns can be represented by "time-series" models of order q with only one unknown constant μ, with $\mu = c(q)\,(q!)$.

- For uniformly sampled polynomial data, this paradigm allows the estimation of polynomial degree, additive noise, and highest degree coefficient with a "time-series" representation, without using any polynomial model.

Journal Papers

A K Nandi, “Data modelling with polynomial representations and autoregressive representations, and their connections”, IEEE Access, vol. 8, pp. 110412-110424, DOI: 10.1109/ACCESS.2020.3000860, 2020.

A K Nandi, “Model order selection from noisy polynomial data without using any polynomial coefficients”, IEEE Access, vol. 8, pp. 130417-130430, DOI: 10.1109/ACCESS.2020.3008527, 2020.

A K Nandi, “Degree and noise power estimation from noisy polynomial data via AR modelling”, Digital Signal Processing, vol. 114, 103071 (13 pages), DOI: 10.1016/j.dsp.2021.103071, 2021.

A K Nandi, “Theoretical and experimental results on time-series representations of polynomials”, Frontiers in Signal Processing, vol. 5, 1582043 (13 pages), DOI: 10.3389/frsip.2025.1582043, 2025.

Publishing Opportunity in Signal Processing

Frontiers in Signal Processing

https://www.frontiersin.org/journals/signal-processing

First Impact Factor 2.6

CiteScore 4.7