Fitting Sum Of Exponentials Model Using Differential Linear Regression

Molise Mokhomo[1], Naleli Jubert Matjelo[2]

[1]*Mukuru, Cape Town, Western Cape, South Africa.*
[2]*National University of Lesotho, Department of Physics and Electronics, P. O. Roma 180, Lesotho*

---***---

Abstract – *In this paper, we present the problem of fitting a sum of exponentials (growth or decay) to some data using linear regression. We start by transforming the problem from the commonly adopted nonlinear regression formulation and present it as a linear regression problem using differential formulation. This allows the problem to be solved in closed form. We demonstrate here by fitting a dataset to a two-component sum of exponentials and show its performance in both noisy and noise-free datasets. The method is shown to be sensitive to noise and this due to the noise amplification inherent in any differential formulation.*

***Key Words*: Exponential Growth, Exponential Decay, Regression, Noise Amplification, Model Fitting, Parameter Estimation.**

1. INTRODUCTION

Fitting models to data comes frequently in data-based research. Model fitting makes extensive use of parameter estimation methods to obtain the optimal model parameters associated with the given data set. Some of the disciplines making use of parameter estimation include system identification, characterization, behavioral analysis, model-based control, state estimation, forecasting, smoothing, and filtering [1], [2], [3]. One of the models used to estimate almost many smooth continuous functions is the sum of the exponentials model (SEM), composed of the weighted average of exponentials [4]. Fitting SEM to data is often considered as a nonlinear regression problem from which there exists no closed-form solution to the problem hence the need for iterative optimization algorithms for solving this problem [4], [5].

In this paper, we adopt a differential formulation of the SEM which allows the parameter estimation problem to posed as a linear regression problem, solvable in closed-form. The approach is based on exploiting the differential equation satisfied by the SEM when formulating the regression problem. In this work we consider the two-component SEM to demonstrate the differential-based linear regression formulation of the SEM fitting problem. This method is related to the method based on integral equations in [6].

The rest of this paper is organized as follows. Section 2 presents a two-component SEM and the resulting differential linear regression problem formulation. Section 3 presents model-fitting simulation results and discussion. Section 4 concludes this work with a summary of major findings and some remarks.

2. LINEAR REGRESSION MODEL FOR SEM

2.1 Sum of Exponentials Model

A general one-dimensional M-component SEM $y(x)$ can be represented as follows,

$$y(x) = \sum_{i=1}^{M} a_i e^{b_i x} \tag{1}$$

with a_i as the i^{th} component weight, and b_i being related to the i^{th} component decay or growth rate. We restrict ourselves to $M = 2$ however, the same concept applies to higher values of M. With $M = 2$ we have the following model with four unknown parameters to be determined from the data,

$$y(x) = a_1 e^{b_1 x} + a_2 e^{b_2 x} \tag{2}$$

2.2 Sum of Exponentials Model In Differential Form

Differentiating equation (2) twice and relating it with its first two derivatives we obtain the following homogeneous ordinary differential equation with constant coefficients,

$$y'' + \alpha_1 y' + \alpha_2 y = 0 \tag{3}$$

with parameters α_1 and α_2 given by,

$$\alpha_1 = -b_1 - b_2 \tag{4}$$

$$\alpha_2 = b_1 b_2 \tag{5}$$

2.3 Differential Linear Regression Model

Given a data $y_i(x_i)$ of size N (i.e. $i = 1, 2, 3, \dots, N$) to fit an SEM, we can formulate the linear regression problem using equations (3-5) as shown in the cost function below,

$$J(\alpha_n) = \sum_{i=1}^{N} (y'' + \alpha_1 y' + \alpha_2 y)^2 \tag{6}$$

The resulting solution to this linear regression problem is given by the following matrix equation,

$$\begin{bmatrix}\alpha_1\\ \alpha_2\end{bmatrix} = -\sum_{i-1}^{N}\begin{bmatrix}y'^2 & yy'\\ yy' & y^2\end{bmatrix}^{-1}\sum_{i-1}^{N}\begin{bmatrix}y'y''\\ yy''\end{bmatrix} \quad (7)$$

from which the estimated parameters are computed as follows,

$$b_1 = -\frac{\alpha_1}{2} \mp \frac{1}{2}\sqrt{\alpha_1^2 - 4\alpha_2} \quad (8)$$

$$b_2 = -\frac{\alpha_1}{2} \pm \frac{1}{2}\sqrt{\alpha_1^2 - 4\alpha_2} \quad (9)$$

Based on the solutions in equations (8 – 9) we continue to solve for the weights a_1 and a_2 from the following linear regression constructed based on equation (2),

$$J(a_m) = \sum_{i=1}^{N}(a_1 e^{b_1 x_i} + a_2 e^{b_2 x_i} - y_i) \quad (10)$$

which leads to the following solution for two weights a_1 and a_2,

$$\begin{bmatrix}a_1\\ a_2\end{bmatrix} = \sum_{i-1}^{N}\begin{bmatrix}e^{2b_1 x_i} & e^{(b_1+b_2)x_i}\\ e^{(b_1+b_2)x_i} & e^{2b_2 x_i}\end{bmatrix}^{-1}\sum_{i=1}^{N}\begin{bmatrix}y_i e^{b_1 x_i}\\ y_i e^{b_2 x_i}\end{bmatrix} \quad (11)$$

In the next section, we present simulation results for fitting this SEM using this linear regression approach.

3. SIMULATION RESULTS AND DISCUSSION

3.1 SEM Fitting Using Noiseless Data

The data points for simulation were generated by a two-component SEM shown in equation (2) with parameters $a_1 = 6.50$, $a_2 = -8.10$, $b_1 = -2.25$, and $b_2 = -0.75$ discretization parameter $h = 10^{-2}$. We then subjected this generated data to the linear regression outlined in the previous section to obtain the parameter estimates. Fig. 1 below shows the plot of the generated data and the SEM model estimation based on the linear regression.

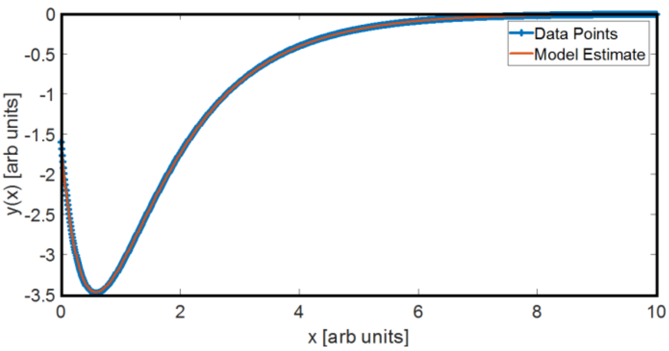

Fig. 1 *SEM fitting with noiseless data.*

The parameters estimated from this data using the linear regression model when the data is noiseless are shown in Table 1 below.

Table 1: *Estimated parameters using noiseless data.*

Parameter	Exact Value	Estimated Value	% Error
a_1	6.50	6.474	0.400%
a_2	−8.10	−8.099	0.012%
b_1	−2.25	−2.225	1.111%
b_2	−0.75	−0.747	0.400%

These percentage errors are mostly constituted by the discretization as shown in [7]. Fig. 2 below shows the squared error between the data points and the model-generated points based on these estimated parameters.

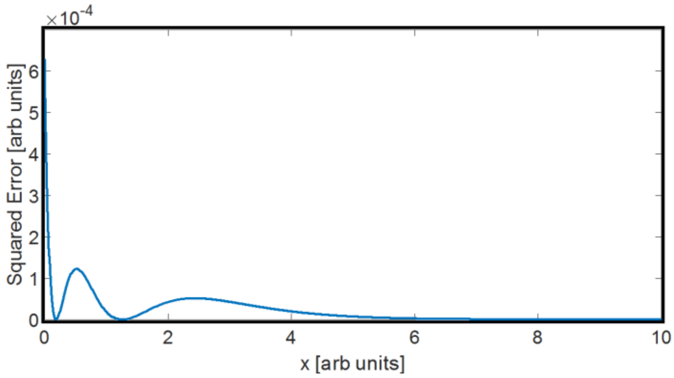

Fig. 2 *Error between noiseless data and model.*

Overall the errors seem acceptable for practical estimation purposes. We next show the impact of adding some noise to the data as done in [8], [9].

3.2 SEM Fitting Using Noiseless Data

Fig. 3 below shows the same generated data plotted together with the model estimate when the uniform noise (with amplitude 10^{-3}) is added to data.

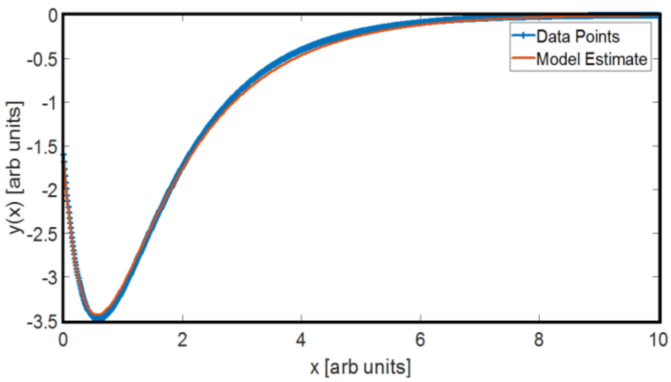

Fig. 3 *SEM fitting with noisy data.*

Table 2 below shows the parameters estimated when the noise has been added to the fitting data.

Table 2: *Estimated parameters using noisy data.*

Parameter	Exact Value	Estimated Value	% Error
a_1	6.50	5.279	18.78%
a_2	−8.10	−6.933	14.41%
b_1	−2.25	−2.481	10.27%
b_2	−0.75	−0.676	9.87%

Comparing the errors from Table 1 with those in Table 2, it is evident that the differential linear regression formulation is sensitive to noise. Fig. 2 below shows the squared error between the data points and the model points generated based on these estimated parameters.

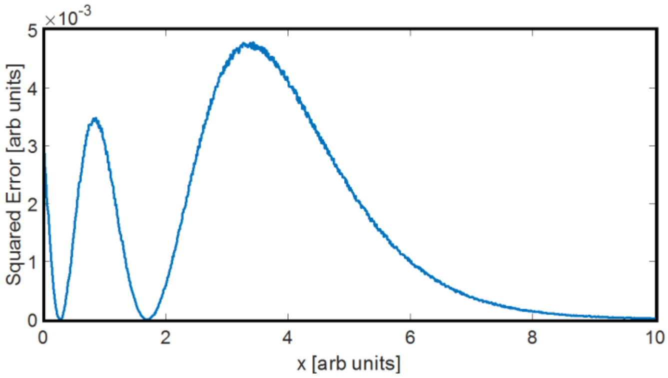

Fig. 4 *Error between noisy data and model.*

In [8], [9] it is pointed out that one way to improve the differential linear regression model and make it robust against noise is to adopt the autoregressive formulation approach. This approach is based on integration as opposed to the differential formulation which is based on differentiation. As such the differential formulation amplifies high-frequency noise (high pass filter) while the autoregressive formulation attenuates high-frequency noise (low pass filter) hence is more robust against noise.

4. CONCLUSIONS

In this work, we have successfully shown how the two-component SEM data-fitting problem can be presented as a linear regression problem with a closed-form form solution. This formulation is shown to be sensitive to noise due to its noise amplification effect inherent in its formulation. As part of future work, both the differential and autoregressive formulations will be applied to artificial neural networks as a way to find a closed-form solution to the network training problem.

REFERENCES

[1] Beck M.B. Applications of System Identification and Parameter Estimation in Water Quality Modeling. (Proceedings of the Oxford Symposium): IAHS-AISH Publ, no. 129, 1980.

[2] Ding R., Zhuang L. Parameter and State Estimator for State Space Models. The Scientific World Journal, vol. 2014, Article ID 106505, 10 pages, 2014.

[3] Luengo D., Martino L., Bugallo M., Elvira V., Sarkka S. A Survey of Monte Carlo Methods for Parameter Estimation. EURASIP J. Adv. Signal Process, 25, 2020,

[4] de Groen P., de Moor B. The Fit of a Sum of Exponentials to Noisy Data, Journal of Computational and Applied Mathematics 20, 175-187, 1987.

[5] Xu C. Nonlinear Least Squares: Newton-type Methods. In: Floudas C., Pardalos P. (eds) Encyclopedia of Optimization. Springer, Boston, MA., 2008.

[6] Jacquelin J. Regressions et Equations Integrales, 2009.

[7] Matjelo N.J., Mokhomo M. Gaussian Mixture Model Fitting Using Differential Linear Regression. International Research Journal of Engineering and Technology (IRJET), 08(07), 2021.

[8] Matjelo N.J. Differential Linear Regression Model For Frequency Estimation. International Research Journal of Engineering and Technology (IRJET), 08(07):1-7, 2021.

[9] Matjelo N.J. Autoregressive Linear Regression Model For Frequency Estimation. International Research Journal of Engineering and Technology (IRJET), 08(07):243-248, 2021.