MODEL SELECTION AND PARAMETER ESTIMATION FOR EXPONENTIAL SIGNALS

ANAND RAMASWAMI
Department of Physics, Washington University
1 Brookings Drive,
St. Louis, Missouri 63130

AND

G. LARRY BRETTHORST
Department of Chemistry, Washington University
1 Brookings Drive,
St. Louis, Missouri 63130

Abstract. In this paper, probability theory is applied to the problem of estimating the decay rate constants in data that are known to contain sums of exponentials. In many instances the number of exponential components is unknown. One might naively believe that the correct strategy, to estimate the parameters, is to first determine the number of exponentials and then estimate the parameters from a model containing that number of exponentials. However, this is not what the rules of probability theory indicate should be done. Probability theory indicates that the best parameter estimates are obtained from the probability for the parameter of interest independent of the number of exponentials in the data. This probability density function is a weighted average. It is a sum over the probability for the parameter given the number of exponentials weighted by the probability that this number of exponentials is the correct value. When the number of exponentials in the data are well determined, this reduces to the problem of determining the number of exponentials and then estimating the parameters given that number of exponentials. However, when the data fail to strongly support a single value for the number of exponentials the result from probability can differ significantly from this intuitive procedure. In this paper a sketch of these calculations is presented, and numerical examples are used to illustrate the calculations.

Key words: Exponential, Model Selection, Parameter Estimation

K. M. Hanson and R. N. Silver (eds.), Maximum Entropy and Bayesian Methods, 9–14.

1. Introduction

Exponential models are ubiquitous throughout all branches of science and engineering. Thus, there is considerable interest in the measurement and analysis of multiexponential data. In chemistry, the concentrations of reactants often follow exponential kinetic schemes. Radioactive nuclear decay is exponential and in nuclear magnetic resonance (NMR) the rates of polarization and decay of magnetization are also generally exponential in nature. In all of these examples, the values of the amplitudes and decay rate constants contain the information of interest. Because such importance is attached to the amplitudes and decay rate constants, it is desirable to have available mathematical methods with strong rigorous foundations.

In this paper probability theory as generalized logic will be used to solve this problem. The problem of estimating the decay rate constants independent of the number of exponentials will be used to illustrate the techniques needed. Other problems, such as estimating the amplitudes, and noise variance are solved similarly. To estimate the decay rate constants, first, a data set $D \equiv \{d_1, \ldots, d_N\}$ is postulated. This data has been sampled from a time series $y(t)$ at discrete times t_i $(1 \leq i \leq N)$; uniform sampling is not assumed. The time series $y(t)$ is assumed to be the sum of two terms, a signal plus noise:

$$d_i = y(t_i) = f(t_i) + e_i \quad (1 \leq i \leq N), \tag{1}$$

where the signal $f(t)$ is taken to be of the form

$$f(t_i) = \sum_{j=1}^{m} B_j \exp\{-\alpha_j t_i\} \tag{2}$$

where e_i represents the value of the noise at time t_i, B_j is the amplitude of jth exponential, α_j is the decay rate constant, and m is the number of exponentials. From Eq. (2), the multiexponential model is invariant under permutations of the labels of the decay rate constants and amplitudes. This invariance manifests itself in the joint posterior probability density for the decay rate constants and can be visualized graphically as multiple peaks of equal probability corresponding to exchange of the decay rate constants. Therefore, a convention must be adopted that distinguishes one component from another. Here the exponential rate constants will be ordered such that $\alpha_1 < \alpha_2 < \cdots < \alpha_m$.

2. Estimating The Decay Rate Constant Independent Of The Number Of Exponentials

In probability theory all of the information relevant to estimating a decay rate, α_j, is summarized in a probability density function. For α_j this is denoted by $P(\alpha_j|DI)$: the probability that decay rate "j" had value α. This quantity is computed from the joint probability for α_j and the number of decay rate constants,

m, by using the sum rule:

$$P(\alpha_j|DI) = \sum_{m=1}^{max} P(\alpha_j m|DI) = \sum_{m=1}^{max} P(m|DI)P(\alpha_j|mDI) \tag{3}$$

where $P(m|DI)$ is the probability for the number of exponentials given the data and the prior information, and $P(\alpha_j|mDI)$ is the probability for the jth decay rate constant given the number of exponentials, the data, and the prior information. We define max as the upper bound on the number of signal components (which we take as three for practical computational reasons). To obtain the optimal estimate of the decay rate constant requires three steps. First, the probability for the number of exponentials, $P(m|DI)$, must be computed. Second, the probability for the decay rate constant given the number of exponentials, $P(\alpha_j|mDI)$, must be computed. Last, the sum must be computed. From a practical standpoint this last step need only be done if the data do not strongly indicate one particular value of m.

To apply Eq. (3) one must obtain both the probability for the number of exponentials and the probability for the parameter given the number of exponentials. Thus, we have a model selection problem and a parameter estimation problem. Both of these types of problems are well studied in probability theory. For specific examples see [1,2]. For a tutorial on parameter estimation see [3] and for a tutorial on model selection see [4]. Here we use the results described in those papers to illustrate how to estimate parameters independent of the model order.

Before the problem is addressed one might ask: What does it mean to estimate the low decay rate constant, α_1, the medium decay rate constant, α_2, or the high decay rate constant, α_3 given a one exponential model? Here we adopt the convention that if there is a single exponential model the low, medium, and high decay rate constants are all the same. Similarly for a bi-exponential model we defined the medium and high decay rate constants to be the same.

3. Example

The data set used in the example is shown in Fig. 1(a). The probability for the number of exponentials was computed for this data, Fig. 1(b). The probability for the number of exponentials is equally split between one and two exponentials. Note that while this probability distribution indicates there is probably one or two exponentials present, three is completely ruled out. The probability for a third exponential is approximately 10^{-5}.

To estimate the low decay rate constant we must compute the probability density function for the low decay rate constant given the one, two and three exponential models. These probability density functions are shown in Figures 2(a), 2(b), and 2(c) respectively. The probability for the low decay rate independent of the number of exponentials, the weighted average, is shown in Figure 2(d). Similarly, the probability for the middle decay rate constant, Fig 3(d), is computed from the probability for the middle decay rate constant given the number of exponentials is one, two, or three in Figs 3(a), 3(b), and 3(c) respectively. Finally,

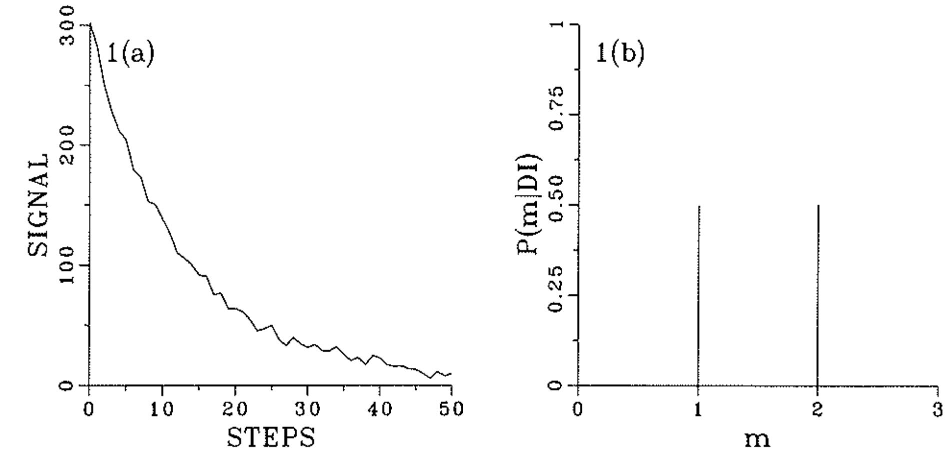

Figure 1. Figure 1(a) is simulated data. It consists of a bi-exponential signal plus Gaussian noise. The exponentials have amplitudes of 100 and 200, and decay rate constants of 0.05 and 0.1 respectively. The noise variance was carefully chosen so that the posterior probability for the number of exponentials given the data, Fig 1(b), would not strongly favor either a one or two exponential model.

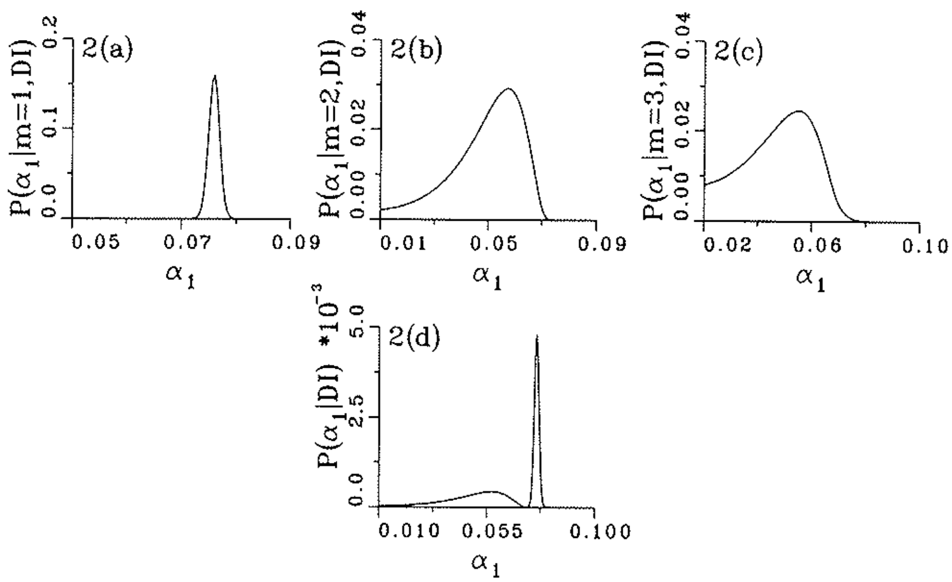

Figure 2. Figure 2(a,b,c) shows the probability distribution for the low decay rate constant given one, two, and three exponential models respectively. Figure 2(d) show the probability distribution for the low decay rate constant independent of the number of exponentials. From Eq. (3) this is a weighted average of Panels 2(a), 2(b), and 2(c); the weights are 0.499, 0.501, and 10^{-5}.

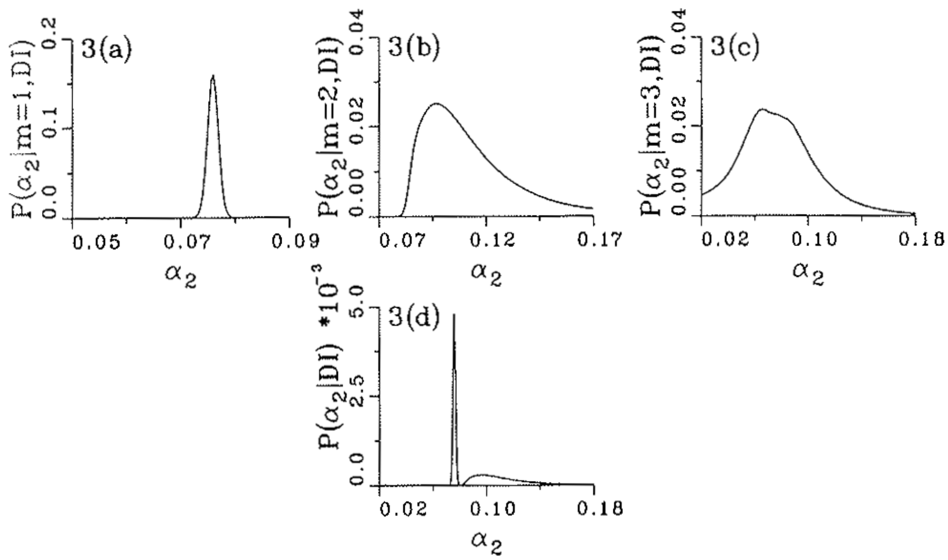

Figure 3. Figure 3(a,b,c) shows the probability distribution for the medium decay rate constant given one, two, and three exponential models respectively. Figure 3(d) show the probability distribution for the medium decay rate constant independent of the number of exponentials. From Eq. (3) this is a weighted average of Panels 2(a), 2(b), and 2(c); the weights are 0.499, 0.501, and 10^{-5}.

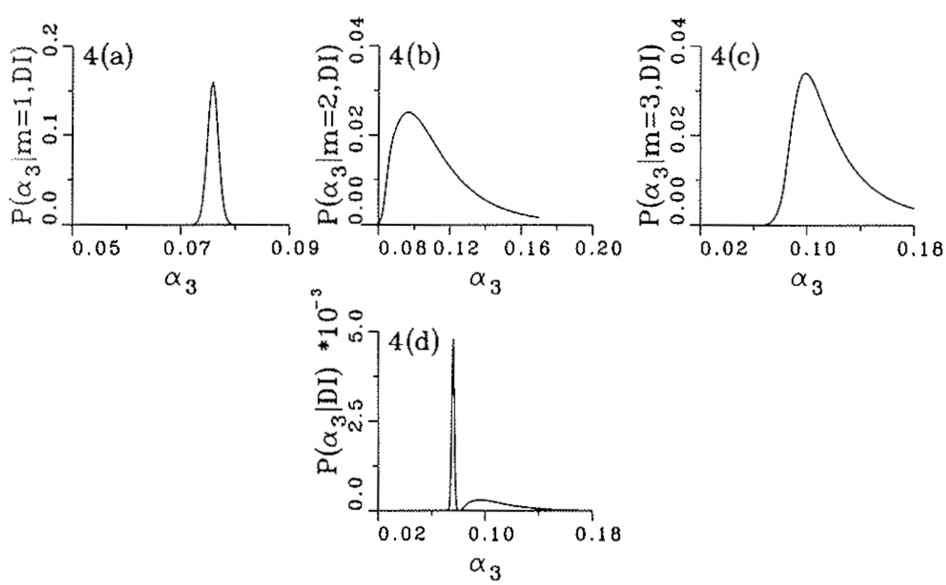

Figure 4. Figure 4(a,b,c) shows the probability distribution for the high decay rate constant given one, two, and three exponential models respectively. Figure 4(d) show the probability distribution for the high decay rate constant independent of the number of exponentials. From Eq. (3) this is a weighted average of Panels 2(a), 2(b), and 2(c); the weights are 0.499, 0.501, and 10^{-5}.

the same procedures may be applied to the problem of estimating the high decay rate constant, Fig 4. Notice that here the probability for the high decay rate constant, Fig 4(d), is the same as the probability for the middle decay rate constant, Fig 3(d). The reason for this is that there is no evidence in the data for more than two exponentials, so the probability for the high decay rate constant is the same as the probability for the middle decay rate constant. This is just what one would expect given the convention we adopted.

4. Summary

Probability theory indicates that all of the information relevant to estimating the decay rate constants independent of the number of exponentials is summarized in a probability density function. This probability density function is a weighted average of the probability for the decay rate constant of interest given that one knows the number of exponentials. The weights are just the probability that the number of exponentials was correct.

References

1. Bretthorst, G. Larry., *Bayesian Analysis. II. Signal Detection and Model Selection*, J. Magn. Reson. **88**, pp. 552-570 1990.
2. Gull Stephen F., *Bayesian Inductive Inference and Maximum Entropy*, G.J. Erickson and C.R.Smith (Eds), in "Maximum Entropy and Bayesian Methods in Science and Engineeringl" Vol 1, pp. 53-74 Kluwer Academic Publishers 1988.
3. Bretthorst, G. Larry., *An Introduction to Parameter Estimation Using Bayesian Probability Theory*, P. Fougere (Ed.), in "Maximum Entropy and Bayesian Methods" p. 53, Kluwer Academic Publishers 1989.
4. Bretthorst, G. Larry., *An Introduction to Model Selection using Probability Theory as Logic*, G. Heidbreder (Ed.), in "Maximum Entropy and Bayesian Methods" Kluwer Academic Publishers, in press.