A Direct Computational Method for Estimating
the Parameters of a Nonlinear Model

by

Stephen W. Cheng^{1,3}

and

Jerome Eisenfeld^{1,2,4}

Technical Report 85

¹Department of Mathematics, University of Texas at Arlington,
Arlington, Texas 76019.

²Department of Medical Computer Science, University of Texas Health Science
Center at Dallas, Dallas, Texas 75235.

³Partially supported by NIH Grant HL 10078.

⁴Partially supported by University of Texas at Arlington Organized Research
Grant 15-681.

A DIRECT COMPUTATIONAL METHOD FOR ESTIMATING THE PARAMETERS OF A NONLINEAR MODEL

by

Stephen W. Cheng and Jerome Eisenfeld

I. Introduction:

We consider the system identification problem which deals with determining information about the unknown constant matrices A , B , and C for the linear dynamical system:

$$\dot{z}(t) = A z(t) + Bu(t), \quad 0 < t < T,$$

$$y(t) = C z(t), \quad 0 < t < T,$$

where $u(t)$ and $y(t)$ are given discretely on the interval $[0, T]$. This leads to a nonlinear estimation problem, or a so-called deconvolution problem, which deals with estimating the number of components N , the decay rates λ_i , $i = 1, 2, \dots, N$, and the corresponding amplitudes $\alpha_i > 0$, $i = 1, 2, \dots, N$ in the model:

$$x(t) = \sum_{i=1}^N \alpha_i \exp(\lambda_i t), \quad 0 \leq t < \infty, \quad (1.1)$$

$$f(t) = \int_0^t e(s) x(t-s) ds, \quad 0 \leq t < \infty, \quad (1.2)$$

from the given discrete data $f(t_i)$, $e(t_i)$, $i = 1, 2, \dots, q$, [1]-[2].

Deconvolution problems are important in a number of fields. One of these is the analysis of fluorescence decay experimental data [3].

Many numerical methods are available for approaching this problem.

The Isenberg method of moments (or IMOM for short) was found to be the best among these methods in most cases [2]-[7]. A direct computational method which is essentially based on the general moment theory has been developed [7]. This method which was called the direct computational method (or DCM for short) alleviates some long standing difficulties arising from IMOM. Two of these are the cut-off error incurred when integration is required to infinite time and lack of flexibility in weighting the data.

In this paper, the theory is presented with the specification of the admissible weighting functions and the estimation of the amplitudes α_i using the linear least squares method as an option. The filtering scheme, which can be used to filter out one known decay term and increase the accuracy of the estimation, is presented. Numerical examples are provided.

II. DCM (Direct Computational Method)

We consider a model of the form

$$x(t) = \sum_{i=1}^N \alpha_i \exp(\lambda_i t), \quad 0 \leq t \leq T, \quad (2.1)$$

$$f(t) = \int_0^t e(s)x(t-s)ds, \quad 0 \leq t \leq T, \quad (2.2)$$

instead of Equations (1.1)-(1.2). The $\{\lambda_i\}$ are distinct and the $\{\alpha_i\}$ are greater than zero. It has been shown in [7] that by DCM the given discrete data $f(t_i)$, $e(t_i)$, $i = 1, 2, \dots, q$ are mapped onto a moment sequence $\{s_k\}$, $k = 0, 1, 2, \dots, m$, and s_k can be calculated by

$$s_k = (-1)^k \int_0^T f(t) h^{(k)}(t) dt, \quad k = 0, 1, \dots, m \quad (2.3)$$

where $h(t)$ is an admissible weighting function and satisfying: (i) h is m times continuously differentiable where m is sufficiently large so that $m \geq 2N + 1$ (m can be chosen so that $m \geq 2N - 1$ in the cases of estimating λ_i and α_i), (ii) $h^{(k)}(T) = 0$ and

$$\int_0^T e(t) h^{(k)}(t) dt = 0, \quad k = 0, 1, \dots, m-1. \quad (2.4)$$

It also has been shown in [7] that the moment sequence $\{s_k\}$ can be expressed by

$$s_k = \sum_{i=1}^N \alpha_i \lambda_i^k, \quad k = 0, 1, \dots, m \quad (2.5)$$

where

$$\alpha_i = \alpha_i \int_0^T \exp(\lambda_i t) \left[\int_0^{T-t} e(s) h(s+t) ds \right] dt. \quad (2.6)$$

Once the moment sequence $\{s_k\}$ is obtained, the model parameters N , λ_i , and α_i can be estimated by applying the general moment theory, See [2], [7] - [10].

It is natural to choose the admissible weighting function $h(t)$ in the form

$$h(t) = (T-t)^m (t^m + d_{m-1} t^{m-1} + \dots + d_0). \quad (2.7)$$

Substitution of Equation (2.7) into Equation (2.4) gives rise to an $m \times m$ system of equations for $d_0, d_1, \dots, d_{m-1}$, any solution of which yields an admissible weighting function $h(t)$ [7].

III. Filtering Scheme

Consider the case with one known λ_i in Equation (2.1). Without

loss of generality, we assume that λ_N is known. Let $\{s_k\}$ be a moment sequence defined by Equation (2.5). Then

$$\begin{aligned} s_{k+1} - \lambda_N s_k &= \sum_{i=1}^N a_i \lambda_i^{k+1} - \lambda_N \sum_{i=1}^N a_i \lambda_i^k \\ &= \sum_{i=1}^{N-1} a_i (\lambda_i - \lambda_N) \lambda_i^k \\ &= \sum_{i=1}^{N-1} b_i \lambda_i^k \end{aligned}$$

Where $b_i = a_i (\lambda_i - \lambda_N)$, $i = 1, 2, \dots, N-1$. Let $s_{k+1} - \lambda_N s_k$ be denoted by $\hat{s}_k$ then $\{\hat{s}_k\}$ forms a moment sequence. Again, by applying the general moment theory, N , the remaining λ_i , and all a_i can be estimated. Thus we have the following result:

Theorem 1. Let $\{s_k\}$ be a moment sequence defined by (2.5), and let $\hat{s}_k = s_{k+1} - \lambda_N s_k$ then $\{\hat{s}_k\}$ forms a moment sequence. Moreover, let $\{\hat{G}_k\}$ denote the Hankel matrices

$$\hat{G}_k = \begin{pmatrix} \hat{s}_0 & \hat{s}_1 & \dots & \hat{s}_{k-1} \\ \hat{s}_1 & \hat{s}_2 & \dots & \hat{s}_k \\ \vdots & \vdots & & \vdots \\ \hat{s}_{k-1} & \hat{s}_k & \dots & \hat{s}_{2k-2} \end{pmatrix}$$

and let $|\hat{G}_k|$ denote the determinant of $\hat{G}_k$, then the following statements are true:

$$1. \quad \begin{aligned} |\hat{G}_{N-1}| &\neq 0 \\ |\hat{G}_k| &= 0 \quad \text{if } k > N-1. \end{aligned}$$

In the event that $\{\hat{a}_k\}$ is positive we have further:

$$|\hat{G}_k| > 0 \quad \text{if } k \leq N-1.$$

2. The numbers λ_k , $k = 1, 2, \dots, N-1$, are the roots of a polynomial

$$P_{N-1}(\lambda) = \lambda^{N-1} + c_{N-2}\lambda^{N-2} + \dots + c_0$$

where the coefficients c_i are given by

$$\begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-2} \end{pmatrix} = -\hat{G}_{N-1}^{-1} \begin{pmatrix} \hat{s}_{N-1} \\ \hat{s}_N \\ \vdots \\ \hat{s}_{2N-3} \end{pmatrix}.$$

3. The numbers b_i are given by

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{N-1} \end{pmatrix} = V_{N-1}^{-1} \begin{pmatrix} \hat{s}_0 \\ \hat{s}_1 \\ \vdots \\ \hat{s}_{N-2} \end{pmatrix}$$

where V_{N-1} is the Vandermonde matrix

$$V_{N-1} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_{N-1} \\ \vdots & \vdots & & \vdots \\ \lambda_1^{N-2} & \lambda_2^{N-2} & \dots & \lambda_{N-1}^{N-2} \end{pmatrix}.$$

4. The numbers α_i are given by

$$\alpha_i = \frac{b_i}{\lambda_i - \lambda_N}, \quad i = 1, 2, \dots, N-1,$$

and

$$\alpha_N = \frac{s_k - \sum_{i=1}^{N-1} \alpha_i \lambda_i^k}{\lambda_N^k}.$$

Theorem 1 is essentially an algorithm for purposed filtering scheme. If we know one of λ_i , we can filter out λ_i from the calculated moment sequence, and form a new moment sequence. Then by applying the moment theory, we can estimate N , the remaining λ_i (i.e., $\lambda_1, \lambda_2, \dots, \lambda_{N-1}$), and all α_i (i.e., $\alpha_1, \alpha_2, \dots, \alpha_N$). Practically, this leads to increase the accuracy of the estimation. Numerical examples about filtering scheme are presented in Sec. 5.

IV. Linear Least Squares Method for Estimating α_i

Rewriting Equations (2.1) - (2.2), we have

$$\begin{aligned}
f(t) &= \int_0^t e(t-s) \left(\sum_{i=1}^N \alpha_i \exp(\lambda_i s) \right) ds \\
&= \sum_{i=1}^N \alpha_i \left[\int_0^t e(t-s) \exp(\lambda_i s) ds \right], \\
f(t) &= \sum_{i=1}^N \alpha_i \phi_i(t)
\end{aligned} \tag{4.1}$$

Where $\phi_i(t) = \int_0^t e(t-s) \exp(\lambda_i s) ds$. Once λ_i are known, Equation (4.1)

becomes a linear model, then the linear least squares method (or LLSM for short) can be used to estimate α_i as an option [11]. Numerical examples about LLSM for estimating α_i as an option are presented in next section.

V. Numerical Examples

In this section we present a number of examples using computer simulated data as illustrations of the performance by DCM and of the various points raised in this paper. For comparison, we also present the results obtained by analysis of the data by IMOM. All computer programs used in this study were written in FORTRAN-10 and have been performed on the DEC System-10 and System-20 in time-sharing conversational fashion. All computations were performed in double precision.

The model is of the form as in Equations (2.1) - (2.2) with $-1/\tau_i$ instead of λ_i , and let $e(t)$ be chosen so that

$$\begin{aligned}
e(t) &= \beta \sin \pi t / \hat{T}, \quad 0 \leq t \leq \hat{T}, \\
&= 0, \quad \hat{T} < t.
\end{aligned}$$

The data sets for $f(t)$ with noise distributed in a Gaussian fashion and $e(t)$ were generated by choosing $\beta = 10^5$ and $\hat{T} = 30$ as described in [2]. Many different analyses were carried out for the purpose of comparing the results of various combinations of parameters (N , α_i , τ_i , and T). The admissible weighting functions are chosen as Equation (2.7) for all analyses in this section. Table I gives a comparison of the expected parameters and the calculated results obtained by analysis of the data by DCM for different sampling intervals and a variety of cases. Table II gives a comparison of the accuracy of the estimated parameters with the expected values; the relative errors (in percent) of the estimated parameters are presented along with the results obtained by IMOM. Observe that DCM not only provides nice results for the sampling interval $[0, 500]$, but also provides acceptable results even in the cases of smaller sampling intervals. In most cases, especially in the cases of larger N , the larger the sampling interval, the greater the accuracy. In the cases of larger N we need more samples to get acceptable results. Table III shows the performance of the filtering scheme; the expected values and the relative errors (in percent) of the calculated results, are presented. Study of Table III reveals that the filtering scheme increases the accuracy of the estimation in most cases. Table IV gives a comparison of the accuracy of the estimated parameters α_i with the expected values using DCM and LLSM. Evidently, the error committed by DCM is far less than LLSM option.

Comments and conclusion about DCM and the comparison between DCM and IMOM can be found in [7]. Although DCM as described in [7] has already

shown itself to be an effective and efficient method for solving nonlinear estimation problems as well as system identification problems, the new scheme and option described in this paper enhance its applicability.

TABLE I

Comparison of the Estimated Parameters with the Expected Values by DCM

		Expected Values	[0,100]	[0,200]	Interval [0,300]	[0,400]	[0,500]
1	α_1	1.00	1.046	1.013	1.010	1.009	1.008
	τ_1	7.00	7.256	7.051	7.022	7.011	7.005
	α_2	1.00	0.952	0.990	0.994	0.995	0.996
	τ_2	18.50	18.835	18.547	18.525	18.518	18.515
2	α_1	1.00	1.016	1.000	0.999	0.999	0.999
	τ_1	15.00	15.252	15.008	14.998	14.995	14.996
	α_2	0.50	0.481	0.499	0.499	0.499	0.499
	τ_2	75.00	78.133	75.094	75.019	75.013	75.013
3	α_1	0.50		0.512	0.506	0.505	0.505
	τ_1	5.00		5.204	5.116	5.088	5.074
	α_2	0.25		0.250	0.245	0.245	0.246
	τ_2	20.00		21.420	20.458	20.305	20.242
	α_3	0.10		0.084	0.097	0.098	0.099
	τ_3	45.00		47.232	45.337	45.173	45.125
		Expected Values	[0,500]	[0,600]	Interval [0,700]	[0,800]	[0,900]
4	α_1	0.10	0.105	0.104	0.103	0.103	0.103
	τ_1	5.00	5.211	5.164	5.137	5.118	5.105
	α_2	0.034	0.034	0.033	0.033	0.033	0.033
	τ_2	15.00	17.117	16.380	16.033	15.834	15.706
	α_3	0.017	0.013	0.014	0.015	0.016	0.016
	τ_3	30.00	33.266	31.399	30.824	30.564	30.422
	α_4	0.0084	0.0078	0.0082	0.0083	0.0083	0.0084
	τ_4	60.00	60.677	60.198	60.090	60.050	60.031

TABLE II

Comparison of the Accuracy of the Parameters as Estimated by Different Methods
(Relative error in percentage, cpu in seconds, expected values given in Table I)

		[0,100]	[0,200]	DCM [0,300]	[0,400]	[0,500]	IMOM [0,500]
1	$\Delta\alpha_1$	4.64	1.27	0.98	0.88	0.84	7.14
	$\Delta\tau_1$	3.66	0.73	0.32	0.16	0.08	1.13
	$\Delta\alpha_2$	4.77	0.98	0.61	0.48	0.42	2.22
	$\Delta\tau_2$	1.81	0.26	0.13	0.10	0.08	0.03
2	$\Delta\alpha_1$	1.65	0.00	0.06	0.05	0.06	0.74
	$\Delta\tau_1$	1.68	0.05	0.02	0.03	0.02	0.37
	$\Delta\alpha_2$	3.82	0.23	0.11	0.10	0.10	0.12
	$\Delta\tau_2$	4.18	0.13	0.03	0.02	0.02	0.03
3	$\Delta\alpha_1$		2.37	1.20	0.98	0.97	11.31
	$\Delta\tau_1$		4.07	2.33	1.77	1.47	3.90
	$\Delta\alpha_2$		0.04	2.03	1.84	1.66	16.94
	$\Delta\tau_2$		7.10	2.29	1.52	1.21	1.66
	$\Delta\alpha_3$		15.68	3.15	1.75	1.29	11.21
	$\Delta\tau_3$		4.96	0.75	0.39	0.28	0.42
		[0,500]	[0,600]	DCM [0,700]	[0,800]	[0,900]	
4*	$\Delta\alpha_1$	4.77	3.81	3.28	2.93	2.69	
	$\Delta\tau_1$	4.21	3.28	2.73	2.37	2.10	
	$\Delta\alpha_2$	0.08	3.12	3.93	4.10	4.07	
	$\Delta\tau_2$	14.12	9.20	6.88	5.56	4.71	
	$\Delta\alpha_3$	25.57	14.90	10.12	7.57	6.04	
	$\Delta\tau_3$	10.89	4.66	2.75	1.88	1.41	
	$\Delta\alpha_4$	6.81	2.22	1.11	0.68	0.48	
	$\Delta\tau_4$	1.13	0.33	0.15	0.08	0.05	

*IMOM for four components case is not available.

TABLE III

Comparison of the Accuracy of the Parameters as Estimated Without/With Filtering Scheme (Relative error in percent, * indicates the filtering out decay term)

		Expected Values		Without Filtering	1	With Filtering		4
					2	3		
1	α_1	1.00	$\Delta\alpha_1$	0.32	0.06	0.04		
	τ_1	15.00	$\Delta\tau_1$	0.13	*	0.09		
	α_2	1.00	$\Delta\alpha_2$	0.38	0.18	0.07		
	τ_2	30.00	$\Delta\tau_2$	0.07	0.03	*		
2	α_1	1.00	$\Delta\alpha_1$	2.82	0.21	0.78	1.67	
	τ_1	7.50	$\Delta\tau_1$	2.09	*	0.46	1.17	
	α_2	0.50	$\Delta\alpha_2$	4.68	0.82	1.71	3.06	
	τ_2	18.50	$\Delta\tau_2$	2.79	0.69	*	1.10	
	α_3	0.24	$\Delta\alpha_3$	2.76	1.73	0.90	0.44	
	τ_3	34.80	$\Delta\tau_3$	0.38	0.34	0.21	*	
3	α_1	0.10	$\Delta\alpha_1$	4.77	0.17	1.25	2.20	2.77
	τ_1	5.00	$\Delta\tau_1$	4.21	*	1.00	1.89	2.45
	α_2	0.034	$\Delta\alpha_2$	0.08	5.45	5.52	5.04	4.48
	τ_2	15.00	$\Delta\tau_2$	14.12	3.50	*	3.37	5.62
	α_3	0.017	$\Delta\alpha_3$	25.57	10.37	3.77	2.88	7.43
	τ_3	30.00	$\Delta\tau_3$	10.89	4.02	2.12	*	1.60
	α_4	0.0084	$\Delta\alpha_4$	6.81	3.09	1.96	6.86	0.28
	τ_4	60.00	$\Delta\tau_4$	1.13	0.57	0.38	0.16	*

TABLE IV

Comparison of the Accuracy of the Estimated Parameters α_i with the Expected Values by Different Methods

(Relative error in Percentage)

		Expected Values	Estimated Values	DCM Relative Errors	Estimated Values	LLSM Relative Errors
1	α_1	1.00	1.002	0.24	1.026	2.60
	τ_1	15.00	14.993			
2	α_1	1.00	1.000	0.00	1.059	5.89
	τ_1	30.00	30.027			
	α_2	1.00	0.998			
	τ_2	60.00	60.014			
3	α_1	1.00	1.000	0.01	1.061	0.61
	τ_1	15.00	14.997			
	α_2	0.45	0.449			
	τ_2	45.00	45.017			
4	α_1	0.10	0.101	0.62	0.074	25.59
	τ_1	5.00	4.974			
	α_2	0.017	0.0170			
	τ_2	30.00	30.063			
	α_3	0.0084	0.0084			
	τ_3	60.00	60.049			

REFERENCES

- [1] A. P. Sage and J. L. Melsa, *System Identification*, Academic Press, New York, 1971.
- [2] J. Eisenfeld, S. R. Bernfeld, and S. W. Cheng, *System Identification Problems and the Method of Moments*, Math. Biosciences, 36 (1977), pp. 199-211.
- [3] C. Ford and J. Eisenfeld, *Some Unresolved Questions Pertaining to the Mathematical Analysis of Fluorescence Decay Data*, *Applied Nonlinear Analysis*, V. Lakshmikantham, Ed., Academic Press, New York, to appear.
- [4] I. Isenberg and R. D. Dyson, *The Analysis of Fluorescence Decay by a Method of Moments*, Biophys. J., 9 (1969), pp. 1337-1350.
- [5] I. Isenberg, R. D. Dyson, and R. Hanson, *Studies on the Analysis of Fluorescence Decay Data by the Method of Moments*, Biophys. J., 13 (1973), pp. 1090-1115.
- [6] E. W. Small and I. Isenberg, *The use of Moment Index Displacement in Analyzing Fluorescence Time-Decay Data*, Biopolymers, 15 (1976), pp. 1093-1100.
- [7] J. Eisenfeld and S. W. Cheng, *General Moment Methods for a Class of Nonlinear Models*, submitted for publication.
- [8] F. R. Gantmacher, *The Theory of Matrices*, Vol. I & II, Chelsea, New York, 1959.
- [9] J. Eisenfeld and B. Soni, *Linear Algebraic Computational Procedure for System Identification Problems*, *Proceedings of the First International Conference on Math. Modeling* (Vol. I), X.J.R. Avula, Ed., Univ. of Missouri Press, 1977, pp. 551-561.
- [10] J. Hallmark and J. Eisenfeld, *Separation and Monotonicity Results for the Roots of the Moment Problem*, *Applied Nonlinear Analysis*, V. Lakshmikantham, Ed., Academic Press, New York, to appear.
- [11] N. R. Draper and H. Smith, Jr., *Applied Regression Analysis*, Wiley, New York, 1967.