International Journal of Pure and Applied Mathematics
Volume 28 No. 4 2006, 487-501

CERTAIN LINEAR AND RADICAL MODELS OF DISCRETE TIME SERIES

Xiaona Pan[1], Fucheng Liao[2], Aihua Li[3 §]
[1,2]University of Science and Technology of Beijing
Beijing, P.R. CHINA
[1]e-mail: panxiaona@163.com
[2]e-mail: fuchengliao@163.com
[3]Department of Mathematical Sciences
Montclair State University
1 Normal Avenue, Upper Montclair, NJ 07043, USA
e-mail: lia@mail.montclair.edu

Abstract: This paper studies algebraic methods in modeling discrete time series. In [6], methods of constructing linear models for certain time series are discussed. We give methods of constructing linear models for other types of time series not covered in [6]. In case linear models do not exist, we introduce methods of constructing "linear-like" models. One type of such models can be obtained by finding linear models of a relevant time series whose elements are certain powers of the original one. Examples are provided to illustrate the process.

AMS Subject Classification: 39A10, 37N25
Key Words: discrete time series, linear models, radical models, linear transformation

1. Introduction

Many real world problems can be formulated as certain time series problems. Motivated by the increasing trend in genetic studies and fast growing computer technologies, the study of discrete time series is gaining more and more attention. Recently, a group of researchers have studied discrete time series using

Received: April 27, 2006
§Correspondence author

various algebraic methods [4], [3], [2], [6], [5]. Briefly, a discrete time series S with $m+1$ time steps and n points over the real number field is a set of $m+1$ n-vectors in $\mathbf{R}^n$, that is, $S = \{A_1, A_2, \cdots, A_{m+1}\}$, where each $A_i \in \mathbf{R}^n$. Such a time series can be obtained from a data table with $m+1$ rows and n columns, for example, data achieved from scientific experiments or social activity surveys. Our goal is to find a vector function $\mathbf{f}$ from $\mathbf{R}^n$ to itself, which sends every row of the data (except for the last one) to the next row. In another word, $\mathbf{f}(A_i) = A_{i+1}$ for $i = 1, 2, \ldots, m$. Such a function is called a model of S which can be used to analyze the data and to predict future behavior of the time series. When every component of the model is a polynomial, we call it a polynomial model. It was studied in [6] the conditions for S to have linear models and how to construct such models. In this paper, we develop methods of constructing linear models for some other types of time series not covered in [6]. Furthermore, we introduce methods of constructing "linear-like" models whose components are linear combinations of the r-th radical of homogeneous polynomials of degree r for some positive integer r. This is one approach to constructing non-linear models when linear models do not exist. One type of such models can be obtained by finding linear models of a relevant time series whose elements are certain powers of the original one. Examples are provided to illustrate the process. We focus on time series over the field $\mathbf{R}$ of real numbers. Throughout, we assume the time series in consideration have no identical rows and $m \leq n$. We first adobt several definitions from [6].

Definition 1. A time series S (over a field $\mathbf{R}$) with n points and $m+1$ time steps is a set of $m+1$ distinct row vectors in $\mathbf{R}^n$, say $S = \{A_1, A_2, \ldots, A_{m+1}\}$. A model of the time series S is a function $\mathbf{f} = (f_1, f_2, \ldots, f_n) : \mathbf{R}^n \to \mathbf{R}^n$ such that $f(A_i) = A_{i+1}$, for $i = 1, 2, \ldots, m$. A model $\mathbf{f}$ is of polynomial (linear, quadratic, etc.) type if all of its components are polynomials (linear, quadratic, etc.)

For a polynomial model, it is natural to consider the total degree and the individual degrees of its components.

Definition 2. Let S be a time series and $\mathbf{f} = (f_1, \ldots, f_n)$ be a model of S of polynomial type. We define the degree of $\mathbf{f}$ by

$$\deg(\mathbf{f}) = \max\{\text{total degrees of } f_i \;\big|\; i = 1, 2, \ldots, n\}.$$

Definition 3. (Matrix Form of a Time Series) For the time series S as above, we denote

$$\mathbf{S}=\left[\begin{array}{c} A_1\\ A_2\\ A_3\\ \vdots\\ A_m\\ \hline A_{m+1}\end{array}\right]=\left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\\ \hline a_{(m+1)1} & a_{(m+1)2} & \cdots & a_{(m+1)n}\end{array}\right].$$

Let

$$\mathrm{M}_S=\begin{bmatrix} A_1\\ A_2\\ \vdots\\ A_m\end{bmatrix} \quad \text{and} \quad \overline{\mathrm{M}}_S=\begin{bmatrix} A_2\\ A_3\\ \vdots\\ A_{m+1}\end{bmatrix}.$$

We call M_S the associate matrix to S and $\overline{\mathrm{M}}_S$ the target matrix of S.

Furthermore, for any function $\mathbf{g}$ from $\mathbf{R}^n$ to $\mathbf{R}^n$, we naturally identify it with the induced function from $M_{m\times n}(\mathbf{R}) \to M_{m\times n}(\mathbf{R})$ meaning

$$\mathbf{g}(\mathrm{M}_S)=\begin{bmatrix} \mathbf{g}(A_1)\\ \vdots\\ \mathbf{g}(A_m)\end{bmatrix}.$$

Thus models of a time series can be represented by the corresponding matrix forms. That is, $\mathbf{g}$ is a model of S if and only if $\mathbf{g}(\mathrm{M}_S)=\overline{\mathrm{M}}_S$.

Let $\mathbf{T}:\mathbf{R}^n\to\mathbf{R}^n$ be an invertible linear transformation represented by an $(n\times n)$ invertible real matrix T. For any time series $S=\{A_1,A_2,\ldots,A_{m+1}\}$, a new time series $S'=\{A_1\mathrm{T},A_2\mathrm{T},\ldots,A_{m+1}\mathrm{T}\}$ is produced by $\mathbf{T}$. If $\mathbf{g}=(g_1,g_2,\ldots,g_n)$ is a polynomial model for S', then $\mathbf{T}^{-1}\circ\mathbf{g}\circ\mathbf{T}$ is a polynomial model for S. Here for each $\underline{x}=(x_1,x_2,\ldots,x_n)\in\mathbf{R}^n$,

$$\mathbf{T}^{-1}\circ\mathbf{g}\circ\mathbf{T}(\underline{x})=(g_1(\underline{x}\mathrm{T}),\ldots,g_n(\underline{x}\mathrm{T}))\,\mathrm{T}^{-1}.$$

We cite the relavant result here.

Theorem 4. (see [6] Theorem 2.1) *Let S, $\mathbf{T}$, and S' be as above. Then the invertible linear transformation $\mathbf{T}$ induces a one-to-one correspondence between the polynomial models of S and those of S', which preserves degrees. In particular, if $\mathbf{g}$ is a polynomial model of the time series S', then $\mathbf{f}=\mathbf{T}^{-1}\circ\mathbf{g}\circ\mathbf{T}$ is a polynomial model of the time series S with $\deg(\mathbf{f})=\deg(\mathbf{g})$. Consequently, $\mu(S)=\mu(S')$.*

For convenience, we introduce the following definition.

Definition 5. Let $A = (a_{ij})_{m\times n}$ be an $(m \times n)$-matrix over $\mathbf{R}$, where $m \leq n$.

1. We say A is positive (negative, non-positive, or non-negative) if all of its elements are positive (negative, non-positive, or non-negative).

2. A is left-diagonal if the left most $m \times m$ square matrix is diagonal. A is left-diagonalizable if its reduced column-echelon form is left-diagonal, i.e., A can be reduced to a left-diagonal matrix by finitely many column operations.

3. For a positive integer r, we define $A^{(r)} = (a_{ij}^r)$.

Consider a time series S with the associated matrix M_S. As definded above, for a positive integer r, $S^{(r)}$ refers to the resulting time series in which each entry of the matrix respectively takes the r-th ordinary power of the original entry. Obviously, $\mathrm{M}_S^{(r)}$ is the associated matrix of $S^{(r)}$. In [6], it was shown that a time series S has linear models if the associated matrix is left-diagonalizable. In this paper, we provide a construction method to build linear models for certain time series which are not left-diagonalizable. More generally, in case linear models do not exist, we give a method to construct *radical models*, of which all the components are linear combinations of the r-th radicals of certain homogeneuous polynomials with the uniform degree r. Such models are "linear-like" regarding to the total degree.

2. Models of Radical Type

The existence of linear models was discussed in [6]. Assume a time series S has more than two points and no identical points. It was shown in [6] that if the associated matrix M_S is left-diagonalizable, then S has linear models; that is, $\mu(S) = 1$. In particular, S has linear models when M_S is of full rank. Methods of constructing linear models, when exist, were discussed. In this section, we develop methods to construct radical models. The motivation is that when linear models do not exist, we can provide resources to obtain models "linear-like" (concerning the degree issue).

Definition 6. Let $\mathbf{f} = (f_1, \ldots, f_n)$ be a model of a time series S over a field $\mathbf{R}$. If each component of $\mathbf{f}$ is a linear combination of the r-th radical of a homogeneous polynomial of degree r, we call $\mathbf{f}$ a radical model of degree r.

Note that if $\mathbf{f}$ is such a model, each f_i is a linear combination of radicals in the form of $\sqrt[r]{a_{j1}x_1^r + \cdots + a_{jn}x_n^r}$. Below, we discuss the existence of radical models in several cases.

Theorem 7. *Let S be a time series with the associated matrix* M_S *and fix a positive integer r. Assume $S^{(r)}$ has a linear model. We claim that:*

1. If r is odd, then S has a radical model of degree r.

2. If r is even, then S has a radical model of degree r under the condition that M_S *can be transformed by finitely many column operations to a matrix in which each column is either non-positive or non-negative.*

Proof. Assume $\mathbf{f} = (f_1, f_2, \ldots, f_n)$ is a linear model of $S^{(r)}$, where

$$f_j = k_{j1}x_1 + k_{j2}x_2 + \cdots + k_{jn}x_n, \quad j = 1, 2, \ldots, n.$$

If r is odd, we build a new function $\mathbf{h} = (h_1, h_2, \ldots, h_n)$ from $\mathbf{R}^n \longrightarrow \mathbf{C}^n$ by

$$h_j = \sqrt[r]{k_{j1}x_1^r + k_{j2}x_2^r + \cdots + k_{jn}x_n^r}, \quad j = 1, 2, \ldots, n.$$

When r is odd, this function $\mathbf{h}$ is a model of S because

$$\mathbf{f}(P_i^{(r)}) = P_{i+1}^{(r)} \Longrightarrow \mathbf{h}(P_i) = P_{i+1},$$

for all $i = 1, 2, \ldots, m$. Note that for each i, j

$$h_j(P_i) = \sqrt[r]{k_{j1}a_{i1}^r + k_{j2}a_{i2}^r + \cdots + k_{jn}a_{in}^r} = \sqrt[r]{(a_{(i+1)j})^r} = a_{(i+1)j},$$

if r is odd.

When r is even, without loss of generality, we assume every column of M_S is non-positive or non-negative. We apply the same procedure as above to obtain the new function $\mathbf{h} = (h_1, h_2, \ldots, h_n)$. This time, $h_j(P_i) = |a_{(i+1)j}|$. Since the j-th column of M_S is non-negative or non-positive, we can make the following adjustments on the function $\mathbf{h}$:

$$h'_j = \begin{cases} h_j, & \text{if the } j\text{-th column is non-negative,} \\ -h_j, & \text{if the } j\text{-th column is non-positive.} \end{cases}$$

Obviously this new function $\mathbf{h} = (h_1, h_2, \ldots, h_n)$ is a radical model of degree r for S. □

The column-echelon form of $\mathrm{M}_S^{(r)}$ can be viewed as the associated matrix of a time series $S^{(r)}$. We can apply the techniques given in [6] to find a linear model for $S^{(r)}$ and then use it to build a radical model for S. The radical models obtained in this way may have more than one radical term for some component. However, all components will be linear combinations of single radicals.

Corollary 8. *Let S be a time series with the associated matrix* M_S*.*

1. If M_S *is left-diagonalizable, then for every positive integer* r*,* S *has a radical model of degree* r*.*

2. If $\mathrm{M}_S^{(r)}$ *is left-diagonalizable for an odd integer* r*, then* S *has a radical model of degree* r*.*

Proof. By the results from [6], linear models exist for time series with left-diagonalizable associated matrices. Then the above statement follows directly from Theorem 7(1). □

We now adopt the procedures of building linear models given by [6] to construct radical models of certain degrees. The method may apply the invertible linear transformations to simplify the computation. We briefly describe the process below.

Steps of Building Radical Models for Certain Time Series (SBRM)

Let r be a positive odd integer. Assume M_S is the associated matrix for a time series S and $\mathrm{M}_S^{(r)}$ is left-diagonalizable.

1. Perform the column-echlon form on $\mathrm{M}_S^{(r)}$ to obtain the left-diagonal matrix $\mathrm{M}_{S'}^{(r)}$, viewed as the associated matrix for the new resulting time series $(S')^{(r)}$. Record the invertable matrix $\mathbf{T}$ for the linear transformation.

2. Construct a linear model $\mathbf{g}$ for $(S')^{(r)}$. One option for this process is to use the formulas given in [6].

3. Find a linear model for $S^{(r)}$ by reversing the process, that is, compute

$$\mathbf{f} = \mathbf{T}^{-1} \circ \mathbf{g} \circ \mathbf{T}.$$

4. Construct a radical model of degree r for S based on Theorem 7(1).

Remark 9. The above process can be easily modified to obtain radical models of degree r when r is even and each column of M_S is either non-negative or non-positive, as described in the proof of Theorem 7(2).

Our method of constructing radical models provides an alternative way to find models for a time series. It is useful especially when linear models do not exist. If a time series S does not have linear models, the related time series $S^{(r)}$ may have linear models for some positive integer r. Thus we can build radical models using the above procedure *SBRM*. The case is simpler when the number r is odd. When r is even, we need the matrix to be specific, that is, which can be reduced by column operations to a matrix with only non-negative or non-positive columns. Although this condition is not always satisfied for a fixed even integer $r > 0$, in many cases a matrix can be transformed to another matrix of that type by column operations for some r. Since column operations preserve the existence of linear models, we can use the new matrix to obtain radical models. The following examples show how it works.

Example 10. A time series $\mathbf{S}$ consists of four points: $P_1 = [1, -1, 0, 1, 2]$, $P_2 = [0, 2, 2, 2, 2]$, $P_3 = [1, -2, -1, 0, 1]$, and $P_4 = [-1, 2, -2, 1, 0]$.

The matrix form is:

$$S = \left[\begin{array}{ccccc} 1 & -1 & 0 & 1 & 2 \\ 0 & 2 & 2 & 2 & 2 \\ 1 & -2 & -1 & 0 & 1 \\ \hline -1 & 2 & -2 & 1 & 0 \end{array}\right], \quad \text{with} \quad \mathrm{M}_S = \begin{bmatrix} 1 & -1 & 0 & 1 & 2 \\ 0 & 2 & 2 & 2 & 2 \\ 1 & -2 & -1 & 0 & 1 \end{bmatrix}.$$

It is easy to see that M_S is not left-diagonalizable, but $\mathrm{M}^{(3)}$ is so:

$$\mathrm{M}_S^{(3)} = \begin{bmatrix} 1 & -1 & 0 & 1 & 8 \\ 0 & 8 & 8 & 8 & 8 \\ 1 & -8 & -1 & 0 & 1 \end{bmatrix},$$

which is the associated matrix for the new time series $S^{(3)}$, where

$$S^{(3)} = \left[\begin{array}{ccccc} 1 & -1 & 0 & 1 & 8 \\ 0 & 8 & 8 & 8 & 8 \\ 1 & -8 & -1 & 0 & 1 \\ \hline -1 & 8 & -8 & 1 & 0 \end{array}\right].$$

By applying a linear transformation $\mathbf{T}$, represented by an invertible matrix T, we obtain a new time series $S'^{(3)}$ whose associated matrix is in the reduced column-echelon form:

$$S'^{(3)} = \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline \frac{3}{2} & \frac{-21}{16} & \frac{-5}{2} & 10 & 1 \end{array}\right] \quad \text{and} \quad T = \begin{bmatrix} \frac{7}{6} & \frac{-1}{48} & \frac{-1}{6} & -1 & -9 \\ \frac{1}{6} & \frac{-1}{48} & \frac{-1}{6} & 0 & -1 \\ \frac{-1}{6} & \frac{7}{48} & \frac{1}{6} & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Using the techniques in [6], we can easily construct a linear model $\mathbf{g}$ for $S'^{(3)}$, which is shown below:

$$\mathbf{g} = \left(\tfrac{3}{2}x_1, x_1 - \tfrac{21}{16}x_3, x_2 - \tfrac{5}{2}x_3, x_3 + 10x_4, x_4 + x_5 \right).$$

Then we can obtain a linear model $\mathbf{f} = T^{-1}\mathbf{g}T = (f_1, f_2, f_3, f_4, f_5)$ for $S^{(3)}$, where $f_1 = \frac{7}{48}x_1 + \frac{7}{48}x_2 - \frac{1}{48}x_3$, $f_2 = \frac{49}{6}x_1 + \frac{1}{6}x_2 - \frac{7}{6}x_3$, $f_3 = \frac{171}{16}x_1 + \frac{43}{16}x_2 - \frac{45}{16}x_3$, $f_4 = \frac{55}{6}x_1 + \frac{7}{6}x_2 - \frac{7}{6}x_3$, $f_5 = \frac{149}{16}x_1 + \frac{21}{16}x_2 - \frac{19}{16}x_3$. Using Theorem 7(1), we further build a radical model $\mathbf{h} = (h_1, h_2, \dots, h_5)$ of degree 3 for $\mathbf{S}$, where

$$\begin{cases} h_1 = \sqrt[3]{\frac{7}{48}x_1^3 + \frac{7}{48}x_2^3 - \frac{1}{48}x_3^3}, \\ h_2 = \sqrt[3]{\frac{49}{6}x_1^3 + \frac{1}{6}x_2^3 - \frac{7}{6}x_3^3}, \\ h_3 = \sqrt[3]{\frac{171}{16}x_1^3 + \frac{43}{16}x_2^3 - \frac{45}{16}x_3^3}, \\ h_4 = \sqrt[3]{\frac{55}{6}x_1^3 + \frac{7}{6}x_2^3 - \frac{7}{6}x_3^3}, \\ h_5 = \sqrt[3]{\frac{149}{16}x_1^3 + \frac{21}{16}x_2^3 - \frac{19}{16}x_3^3}. \end{cases}$$

One can easily check that $\mathbf{h}(P_i) = P_{i+1}$ for $i = 1, 2, 3$.

Example 11. Consider the time series $\mathbf{S}$ consisting of four points: $P_1 = [1, -1, 3, 0, -1], P_2 = [2, -4, 4, -2, -4]$, $P_3 = [2, -5, 3, -3, -5]$, and $P_4 = [1, 0, 4, 1, 0]$.

The matrix form is as below:

$$S = \left[\begin{array}{rrrrr} 1 & -1 & 3 & 0 & -1 \\ 2 & -4 & 4 & -2 & -4 \\ 2 & -5 & 3 & -3 & -5 \\ \hline 1 & 0 & 4 & 1 & 0 \end{array}\right] \quad \text{with} \quad \mathrm{M}_S = \left[\begin{array}{rrrrr} 1 & -1 & 3 & 0 & -1 \\ 2 & -4 & 4 & -2 & -4 \\ 2 & -5 & 3 & -3 & -5 \end{array}\right].$$

Here the columns of S are not all negative or all positive. By applying a linear transformation $\mathbf{T}$, represented by an invertible matrix T, we obtain a new time series S whose associated matrix is positive:

$$S' = \left[\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 4 & 6 & 8 \\ 2 & 1 & 3 & 5 & 7 \\ \hline 1 & 3 & 4 & 5 & 6 \end{array}\right] \quad \text{with} \quad \mathrm{M}_{S'} = \left[\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 4 & 6 & 8 \\ 2 & 1 & 3 & 5 & 7 \end{array}\right].$$

It is easy to see that $\mathrm{M}_{S'}$ is not left-diagonalizable, but $\mathrm{M}_{S'}^{(2)}$ is so, which is the associated matrix for $S'^{(2)}$:

$$\mathrm{M}_{S'}^{(2)} = \left[\begin{array}{rrrrr} 1 & 4 & 9 & 16 & 25 \\ 4 & 4 & 16 & 36 & 64 \\ 4 & 1 & 9 & 25 & 49 \end{array}\right].$$

Thus we can easily construct a linear model for $S^{(2)}$. Then we use the techniques described in Theorem 7(2) to obtain a radical model of degree 2 for $\mathrm{M}_{S'}^{(2)}$. Finally, we need reverse the linear transformation process to obtain a radical model of degree 2 for S itselt. It is given by $\mathbf{h} = (h_1, h_2, h_3, h_4, h_5)$, where

$$\begin{cases} h_1 = 3x_1 + x_2, \\ h_2 = -9x_1 - 3x_2 + \sqrt{\frac{2690}{24}x_1^2 + \frac{1590}{24}x_1x_2 + \frac{265}{24}x_2^2 - \frac{47}{8}x_3^2}, \\ h_3 = \sqrt{\frac{1750}{8}x_1^2 + \frac{1050}{8}x_1x_2 + \frac{175}{8}x_2^2 - \frac{83}{8}x_3^2}, \\ h_4 = -12x_1 - 4x_2 + \sqrt{\frac{8242}{24}x_1^2 + \frac{4998}{24}x_1x_2 + \frac{833}{24}x_2^2 - \frac{119}{8}x_3^2}, \\ h_5 = -18x_1 - 6x_2 + \sqrt{\frac{11666}{24}x_1^2 + \frac{7134}{24}x_1x_2 + \frac{1189}{24}x_2^2 - \frac{155}{8}x_3^2}. \end{cases}$$

This $\mathbf{h}$ satisfies that $\mathbf{h}(P_i) = P_{i+1}$ for $i = 1, 2, 3$.

3. Linear Models for Certain Time Series

From [6], a time series has linear models if its associated matrix is left-diagonalizable. In this section, we discuss a class of time series whose associated matrix is not left-diagonalizable but the reduced column-echelon form is of certain special type. We construct linear models for the time series represented by the following types of matrices.

Lemma 12. *Assume a time series S is represented by the matrix* $\mathbf{S}$.
1. If $\mathbf{S}$ *has the following form (the first row is a zero row):*

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0_{m\times m} & \cdots & 0_{m\times n} \\ c_1 & c_2 & c_3 & \cdots & c_{m-1} & c_m & \cdots & c_n \end{bmatrix},$$

then a model of S is given by $\mathbf{g} = (g_1, g_2, \cdots, g_n)$, *where*

$$g_i = \begin{cases} c_i x_{m-1} + \left(1 - \sum_{j=1}^{m-1} x_j\right), & \text{if} \quad i = 1; \\ x_{i-1} + c_i x_{m-1}, & \text{if} \quad 2 < i \le m-1; \\ c_i x_{m-1}, & \text{if} \quad m \le i \le n. \end{cases}$$

2. If $\mathbf{S}$ *has the following form (the zero row is the k-th row,* $2 \le k \le m-1$*):*

$$\begin{bmatrix} 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0_{(k-1)\times(k+1)} & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0_{k\times(k+1)} & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0_{(k+1)\times(k+1)} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0_{m\times m} & \cdots & 0 \\ c_1 & \cdots & c_{k-1} & c_k & c_{k+1} & \cdots & c_{m-1} & c_m & \cdots & c_n \end{bmatrix},$$

then we can construct a model $\mathbf{g} = (g_1, g_2, \cdots, g_n)$ *for* S*;*

$$g_i = \begin{cases} c_k x_{m-1} + \left(1 - \sum_{j=1}^{j=m} x_j\right), & \text{if} \quad i = k; \\ x_{i-1} + c_i x_{m-1}, & \text{if} \quad 2 \le i \le m-1, \quad \text{but} \quad i \ne k; \\ c_i x_{m-1}, & \text{if} \quad i = 1 \quad \text{or} \quad m \le i \le n. \end{cases}$$

Proof. The proof is straightforward. □

Theorem 13. *If the column-echelon form of a time series has one of the above forms, then it has linear models.*

Proof. This is directly from the above lemma and the reversing linear transformation process (Theorem 4). □

The following examples show the construction process discussed above.

Example 14. *Caes 1.* The first row is the zero row.

A time series S consisting of five points: $P_1 = [0, 0, 0, 0, 0, 0], P_2 = [1, 10, 3, -5, -2, -1], P_3 = [10, 10, -6, -14, -2, -10], P_4 = [-22, -22, 15, 38, 17, 13]$, and $P_5 = [-14, 4, 30, -2, 28, 32]$. We construct a linear model for S (so $\mu(S) = 1$).

The matrix form of S is:

$$\mathbf{S} = \left[\begin{array}{rrrrrr} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 10 & 3 & -5 & -2 & -1 \\ 10 & 10 & -6 & -14 & -2 & -10 \\ -22 & -22 & 15 & 38 & 17 & 13 \\ \hline -14 & 4 & 30 & -2 & 28 & 32 \end{array}\right].$$

By applying a linear transformation **T**, represented by an invertible matrix T, we obtain a new time series S' whose associated matrix is in its reduced column-echelon form. Here

$$\mathbf{S}' = \left[\begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 3 & -1 & 4 & 2 & 5 & 1 \end{array}\right] \text{ and } \mathrm{T} = \frac{1}{18}\left[\begin{array}{cccccc} 1 & 1 & -1 & 2 & 2 & 3 \\ 2 & 1 & 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & -1 & 2 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 1 & -1 & 0 & 2 & 1 & 1 \end{array}\right].$$

By Theorem 12, we can easily construct a linear model $\mathbf{g}$ for S':

$$\mathbf{g} = \left(\ 1 - x_1 - x_2 + 2x_3 - x_4, x_1 - x_3, x_2 + 4x_3, 2x_3, 5x_3, x_3\ \right).$$

We then use the methods provided by Theorem 4 to obtain a linear model $\mathbf{f}$ for S:

$$\mathbf{f} = \mathbf{T}^{-1} \circ \mathbf{g} \circ \mathbf{T} = (f_1, f_2, f_3, f_4, f_5, f_6),$$

where

$$\begin{cases} f_1 = 1 - \frac{1}{18}x_1 - \frac{1}{3}x_2 + \frac{11}{6}x_3 - \frac{77}{18}x_5 + \frac{5}{3}x_6, \\ f_2 = 10 - \frac{23}{9}x_1 - \frac{7}{3}x_2 + \frac{7}{3}x_3 - 2x_4 - \frac{43}{9}x_5 + \frac{2}{3}x_6, \\ f_3 = 3 - \frac{5}{3}x_1 - \frac{1}{2}x_2 - x_3 - \frac{3}{2}x_4 + \frac{25}{6}x_5 - \frac{3}{2}x_6, \\ f_4 = -5 + \frac{41}{18}x_1 + \frac{5}{3}x_2 - \frac{19}{6}x_3 + 2x_4 + \frac{97}{18}x5 - \frac{7}{3}x_6, \\ f_5 = -2 - \frac{7}{18}x_1 + \frac{7}{6}x_2 - \frac{7}{6}x_3 - \frac{1}{2}x_4 + \frac{50}{9}x_5 - \frac{5}{6}x_6, \\ f_6 = -1 - \frac{13}{9}x_1 - \frac{1}{6}x_2 - \frac{4}{3}x_3 - \frac{3}{2}x_4 + \frac{95}{18}x_5 - \frac{7}{6}x_6. \end{cases}$$

Example 15. The third row is the zero row.

A time series S consists of five points: $P_1 = [1, 1, -1, 2, 2, 3], P_2 = [2, 1, 0, 1, -1, 0], P_3 = [0, 0, 0, 0, 0, 0], P_4 = [1, -1, 0, -1, 2, 1]$, and $P_5 = [-1, -12, -1, 2, 14, 7]$. We construct a linear model for S.

The matrix form of S and the reduction are as below:

$$\mathbf{S} = \left[\begin{array}{cccccc} 1 & 1 & -1 & 2 & 2 & 3 \\ 2 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & -1 & 2 & 1 \\ \hline -1 & -12 & -1 & 2 & 14 & 7 \end{array}\right] \longrightarrow \mathbf{S}' = \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 1 & -3 & 2 & -2 & -1 & 4 \end{array}\right].$$

The linear transformation **T** for the reduction is given by:

$$\mathrm{T} = \frac{1}{18}\begin{bmatrix} 1 & 10 & 3 & -5 & -2 & -1 \\ 10 & 10 & -6 & -14 & -2 & -10 \\ -22 & -22 & 15 & 38 & 17 & 13 \\ 6 & 6 & -9 & -12 & -3 & 3 \\ 18 & 18 & -9 & -36 & -9 & -9 \\ -21 & -30 & 18 & 51 & 15 & 12 \end{bmatrix}.$$

Similarly, a linear model **g** for S' is obtained by Lemma 12: $\mathbf{g} = (x_3, x_1 - 3x_3, 1 - x_1 - x_2 + x_3 - x_4, -2x_3, -x_3, 4x_3)$. A linear model **f** for S is obtained as below: $\mathbf{f} = \mathbf{T}^{-1} \circ \mathbf{g} \circ \mathbf{T} = (f_1, f_2, f_3, f_4, f_5, f_6)$, where

$$\begin{cases} f_1 = 1 - \frac{2}{9}x_1 + \frac{7}{9}x_2 - \frac{19}{9}x_3 + \frac{2}{3}x_4 + 2x_5 - \frac{7}{3}x_6\,, \\ f_2 = -1 - \frac{13}{9}x_1 + \frac{41}{9}x_2 - \frac{193}{18}x_3 + \frac{35}{6}x_4 + \frac{13}{2}x_5 - \frac{73}{6}x_6\,, \\ f_3 = -\frac{5}{6}x_1 + \frac{5}{3}x_2 - \frac{25}{6}x_3 + \frac{5}{2}x_4 + \frac{5}{2}x_5 - 5x_6\,, \\ f_4 = -1 + \frac{19}{18}x_1 - \frac{4}{9}x_2 + \frac{16}{9}x_3 - \frac{5}{3}x_4 - x5 + \frac{17}{6}x_6\,, \\ f_5 = 2 + \frac{7}{9}x_1 - \frac{38}{9}x_2 + \frac{169}{18}x_3 - \frac{29}{6}x_4 - \frac{11}{2}x_5 + \frac{61}{6}x_6\,, \\ f_6 = 1 + x_1 - 3x_2 + 7x_3 - 4x_4 - 4x_5 + 8x_6. \end{cases}$$

4. An Application

In this section, we apply our reseults to model a time series from laboratory measurements of transcription of specific genes. We start with a set of real laboratory data and construct a linear model and two radical models and then use them to predict future values. Comparison is made between these models and the actual laboratory result for the next value.

Example 16. Analyzing a laboratory data for transcription of specific genes.

Here the j-th column represents the transcription data for the j-th individual gene over time and the i-th row shows the i-th measurement for all six genes. The real data P_6 from the laboratory is

$$P_6 = [0.1, 0.2, 0.25, 0.27, 0.3, 0.35].$$

Gene	P_1	P_2	P_3	P_4	P_5	P_6
$t=1$	1.7	1.0	1.3	1.7	2.5	3.5
$t=2$	0.21	0.4	0.5	0.2	0.2	0.2
$t=3$	0.11	0.13	0.14	0.11	0.13	0.16
$t=4$	0	0	0.05	0.1	0.4	0.9
$t=5$	0.11	0.15	0.17	0.19	0.23	0.27

Table 1: Transcription of genes

By the method discussed in Section 3, we construct a linear model $\mathbf{f} = (f_1, f_2, \ldots, f_6)$ for this discrete time series, where

$$\begin{cases} f_1 = \frac{-12828}{1391}x_1 + \frac{29515}{13910}x_2 - \frac{180231}{13910}x_3 - \frac{63291}{27820}x_4 + \frac{68531}{27820}x_5, \\ f_2 = \frac{-24202}{1391}x_1 + \frac{56328}{1391}x_2 - \frac{69821}{2782}x_3 - \frac{21061}{5564}x_4 + \frac{24807}{5564}x_5, \\ f_3 = \frac{-25613}{1391}x_1 + \frac{316752}{6955}x_2 - \frac{398503}{13910}x_3 - \frac{127471}{27820}x_4 + \frac{143317}{27820}x_5, \\ f_4 = \frac{-27103}{1391}x_1 + \frac{363419}{6955}x_2 - \frac{232418}{6955}x_3 - \frac{85421}{13910}x_4 + \frac{86067}{13910}x5, \\ f_5 = \frac{-6280}{1391}x_1 + \frac{231008}{6955}x_2 - \frac{328207}{13910}x_3 - \frac{206819}{27820}x_4 + \frac{149753}{27820}x_5, \\ f_6 = \frac{28874}{1391}x_1 + \frac{4357}{6955}x_2 - \frac{46659}{6955}x_3 - \frac{64824}{6955}x_4 + \frac{26733}{6955}x_5. \end{cases}$$

From this model, we can forecast the next value $P_6^{(1)} = [0.207, 0.4002, 0.4134, 0.3459, -0.0997, -0.8359]$. Using the method discribed in Section 2, a radical model $\mathbf{g} = (g_1, g_2, \ldots, g_6)$ of degree two is obtained:

$$\begin{cases} g_1 = \sqrt{\frac{-16554407}{22503101}x_1^2 - \frac{478851081}{2250310100}x_2^2 + \frac{113224752}{562577525}x_3^2 + \frac{1590627701}{2250310100}x_4^2}, \\ g_2 = \sqrt{\frac{-31673441}{22503101}x_1^2 - \frac{17481877}{45006202}x_2^2 + \frac{46014889}{135018606}x_3^2 + \frac{188754641}{135018606}x_4^2}, \\ g_3 = \sqrt{\frac{-213189930}{22503101}x_1^2 + \frac{1015358313}{225031010}x_2^2 - \frac{3001451597}{1125155050}x_3^2 + \frac{10755327087}{1125155050}x_4^2}, \\ g_4 = \sqrt{\frac{-134280033}{22503101}x_1^2 + \frac{380724447}{225031010}x_2^2 - \frac{1049000653}{1125155050}x_3^2 + \frac{6684311363}{1125155050}x_4^2}, \\ g_5 = \sqrt{\frac{-1201226337}{22503101}x_1^2 + \frac{7128456303}{225031010}x_2^2 - \frac{21820857717}{1125155050}x_3^2 + \frac{60504214507}{1125155050}x_4^2}, \\ g_6 = \sqrt{\frac{-5724493430}{22503101}x_1^2 + \frac{36450061603}{225031010}x_2^2 - \frac{336736650251}{3375465150}x_3^2 + \frac{866448766571}{3375465150}x_4^2}. \end{cases}$$

This model predicts $P_6^{(2)} = [0.1327, 0.1857, 0.5049, 0.3917, 1.2032, 2.6363]$ to be the next value. Similarly, we construct a radical model $\mathbf{h} = (h_1, h_2, \ldots, h_6)$ of degree 3. It is given by

$$\begin{cases} h_1 = \sqrt[3]{\frac{-320945831}{259313957}x_1^3 - \frac{46315813}{740897020}x_2^3 + \frac{14507098}{259313957}x_3^3 + \frac{3182489167}{2593139570}x_4^3}, \\ h_2 = \sqrt[3]{\frac{-841814142}{259313957}x_1^3 - \frac{5560082}{37044851}x_2^3 + \frac{33481146}{259313957}x_3^3 + \frac{838141964}{259313957}x_4^3}, \\ h_3 = \sqrt[3]{\frac{-1461327747}{259313957}x_1^3 + \frac{47909759}{37044851}x_2^3 - \frac{151293049}{259313957}x_3^3 + \frac{1467319353}{259313957}x_4^3}, \\ h_4 = \sqrt[3]{\frac{-1927373533}{259313957}x_1^3 + \frac{35284196}{37044851}x_2^3 - \frac{103488379}{259313957}x_3^3 + \frac{1923801196}{259313957}x_4^3}, \\ h_5 = \sqrt[3]{\frac{-15318638702}{259313957}x_1^3 + \frac{2738473746}{37044851}x_2^2 - \frac{9678110382}{259313957}x_3^3 + \frac{15745173772}{259313957}x_4^3}, \\ h_6 = \sqrt[3]{\frac{-144469488561}{259313957}x_1^3 + \frac{31324615796}{37044851}x_2^3 - \frac{111079112787}{259313957}x_3^3 + \frac{149511331576}{259313957}x_4^3}. \end{cases}$$

Based on this model, the next value should be $P_6^{(3)} = [0.19040.2630, 0.3217, 0.3498, 0.7433, 1.5914]$.

Compare the three predicted results $P_6^{(1)}$, $P_6^{(2)}$, $P_6^{(3)}$ with the real P_6 from laboratory, we observe that:

1. The radical models predict all positive entries and the trend is incremental, which is as the same as the given time series.

2. Let d_j be the Euclidean distance between P_6 and $P_6^{(j)}$, then $d_1 = 1.2846$, $d_2 = 2.4747$, and $d_3 = 1.3271$. It is clear that the linear model is the best one when predicting the next value, if the 'goodness' is based on the closeness under the Euclidean metric.

It remains open on how to determine in what situations linear models are better. In a biology labortory, performing laboratory measurements of transcription of specific genesis is very expensive. Usually, only about 10 such measurements are affordable. Our algebraic approach provides a way to construct mathematics models for the time series produced from the laboratory data and can be used to predict future values. The challenge here is to know what properties of the functions involved will well reflect the required biological meaning and how to recognize and search "good" models which will help to understand the data such as that from the signal transduction pathways of genes. Much work need to be done toward this direction.

References

[1] D. Cox, J. Little, D. O'Shea, *Ideals, Varieties, and Algorithms*, Second Edition, Springer Verlag, New York (1996).

[2] R. Laubenbacher, J. Shah, B. Stigler, A computational algebra approach to the identification of gene regulatory networks, *Proc. the Third Intl. Congress on System Biology*, Stockholm (2002).

[3] R. Laubenbacher, J. Shah, B. Stigler, Simulation of polynomial systems, In: *Simulation in the Health and Medical Sciences* (Ed-s: J.G. Anderson and Katzper), Soc. for Modeling and Simulation Intl., San Diego, CA (2003).

[4] R. Laubenbacher, B. Stigler, A computational algebra approach to the reverse engineering of gene regulatory networks, *Journal of Theoretic Biology*, **229** (2004), 523-537.

[5] Aihua Li, An algebraic approach to building interpolating polynomials, *Discrete and Continuous Dynamical System* (2005), 597-604.

[6] Aihua Li, Chuang Peng, *Linear Transformations On Polynomial Models of Time Series*, **17**, No. 2 (2004), 235-248.